POINCARÉ-WEYL'S PREDICATIVITY: GOING BEYOND Γ_0

ARNON AVRON

Abstract. On the basis of Poincaré and Weyl's view of predicativity as invariance, we develop an extensive framework for predicative, type-free first-order set theory in which Γ_0 and much bigger ordinals can be defined as von-Neumann ordinals. This refutes the accepted view of Γ_0 as the "limit of predicativity".

§1. Introduction.

1.1. What Predicativism, and Why? In [43] the basic historic problem of the research in foundations of mathematics (FOM) is formulated as follows:

> How to reconstruct mathematics on a secure basis, one maximally immune to rational doubts.

The predicativist program ([24, 11, 12, 48, 50]) has been one of the attempts to solve this basic problem of FOM. It seeks to establish certainty in mathematics in a constructive way, but without revolutionizing it or changing its underlying classical logic (as the intuitionistic program does). The program was initiated by Poincaré [35, 36, 37, 38]. Its viability was demonstrated by Weyl, who seriously developed it for the first time in his famous small book "Das Kontinuum" ([51, 53]. Weyl, and then Feferman ([22, 25], have shown that a very large part of classical analysis can be developed within their predicative systems. Feferman further argued that predicative mathematics in fact suffices for developing all the mathematics that is actually indispensable to present-day natural sciences. Hence the predicativist program has been successful in solving the basic problem of FOM. (In my opinion it is the only one about which this can truly be said.)

Poincaré's predicativism started as a reaction to the set-theoretical paradoxes. However, in the writings of both Poincaré and Weyl, predicativity derives not so much from the need to avoid paradoxes, but from their definitionist view that infinite objects, such as sets or functions, exist only in so far as they are introduced through *legitimate definitions*:

> "No one can describe an infinite set other than by indicating properties which are characteristic of the elements of the set. And no one can establish a correspondence among infinitely many things without indicating a rule, i.e., a relation, which connects the corresponding objects with one another. The notion that an infinite set is a 'gathering' brought together by infinitely many individual arbitrary acts of

1991 *Mathematics Subject Classification.* 03B30,03E10,03E30,03F99,00A30,03A05,0303.
Key words and phrases. predicativity, set theory, foundations of mathematics, proof theory.

> selection, assembled and then surveyed as a whole by consciousness, is nonsensical." ([53], P.23)

The implications of the above principle concerning infinite objects depend of course in a crucial way on the question: What definitions should be accepted as 'legitimate'? Therefore it is no wonder that 'predicativism' (like 'constructivism') becomes a name of a group of approaches to mathematics and its foundations ([24, 12]).[1] We emphasize that in this paper we reserve this name solely to the program as it was initiated by Poincaré and pursued by Weyl. That program is known nowadays as 'predicativity given the natural numbers', since in addition to the definitionist principle mentioned above, it also accepts the collection N of the natural numbers as a well understood mathematical concept that constitutes a set. Moreover, it views the idea of iterating an operation or a relation a finite number of times as fundamental, and accepts induction on the natural numbers as a universally valid method of proof.[2] Still, even with this restriction, the word 'predicative' has two different interpretations, corresponding to Poincaré's "two distinct diagnoses of the source of the paradoxes" ([24]; see also [11]). We call them 'Russell's predicativity' and 'Poincaré-Weyl predicativity'. This paper is devoted to the second one. However, since it is the first which is usually identified with predicativism, we discuss it first.

1.2. Russell's predicativity. Adopting the analysis indicated in Richard's paper [39], Poincaré's first diagnosis was that the definition of Richard's paradoxical real number is circular: it uses the totality of all definitions, to which it already belongs. The corresponding Vicious Circle Principle, VCP, was adopted by Russel in [40] and in *Principia Mathematica* [54]. According to the latter, a vicious circle arises when we assume that "a collection of objects may contain members which can only be defined in terms of the collection itself". [3] A clearer (and stronger) formulation of the VCP has been given by Kreisel in [30]:

> "A predicative definition of a set (say, of natural numbers) is required to use quantification only over 'previously defined' totalities; the set of natural numbers themselves is supposed to be given, or the notion of 'finite' is supposed to be well-defined."

Kreisel then went on and note that

> "The traditional way of making the idea of a predicative definition explicit is by introducing a ramified hierarchy."

The idea of ramified hierarchy was introduced and used by Russell in *Principia Mathematica*. Later it was generalized by Wang ([47]) and Kreisel. In the second-order context the generalization is explained in [24] as follows:

> "The basic step in that hierarchy consists in passing from a collection D of subsets of N to a new collection $D^\star$ by putting a set S in $D^\star$ just in case there is a formula $\varphi(x)$ of second-order arithmetic such

[1] A particularly extreme case can be found in [34]. As noted in [24], what is called there 'predicative arithmetic' is actually strictly finitist arithmetic.

[2] For reasons that will be clarified in Section 3, true predicativity is in our opinion necessarily 'predicativity given the natural numbers'.

[3] Chapter II.1 of the second edition; the explanation in the first edition is similar. The term 'predicative' is used in *Principia Mathematica* in a technical different sense.

> that for all n, $n \in S \leftrightarrow (\varphi(n))^D$, where the superscript 'D' indicates that all second-order quantifiers in φ are relativized to range over D. Then we can define the collections R_α for arbitrary ordinals α by $R_0 = \emptyset, R_{\alpha+1} = (R_\alpha)^\star,$ and for limit $\alpha, R_\alpha = \bigcup_{\beta<\alpha} R_\beta$."

This description raises the question: What ordinals α can serve for the purpose of constructing this ramified hierarchy of R_αs? To answer this, Kreisel proposed in [29] an autonomous process, where a well-ordering becomes available at some stage only if it has been defined and recognized (as a well-ordering of ω) at an earlier stage. Without the 'recognition' criterion, which introduces proof-theoretic considerations, we are left with a purely semantic condition that allows to go up to every well-ordering $< \omega_1^{CK}$ (Church-Kleene's first non-recursive ordinal). With the recognition condition, Feferman [15] and Schütte [41] independently replaced ω_1^{CK} by the much smaller Γ_0 (the Feferman-Schütte ordinal). Following their work, the hierarchies of formal systems up to Γ_0 which were developed by them (on the basis of the intuitive semantics of the R_αs) has become the "canonical reference: one considers predicative any formal system which can be reduced to a system in that hierarchy" ([11]). Accordingly. Γ_0 is almost universally accepted as the 'ordinal of predicativity'. An example of the implications of this is given in [13]: "The fact that the proof-theoretic strength of theories of inductive definitions exceeds the strength of the whole ramified hierarchy is taken as clear indication that generalized inductive definitions involve impredicativity."

Up to now, the only mathematician to reject the 'Γ_0-thesis' of Feferman and Schütte has been Weaver, who forcefully attacked this thesis in [49]. Unfortunately, his (in my opinion quite justified) criticism has been almost totally ignored by the logical community. Nevertheless, as a true predicativist (which is what I am taking myself to be), it is clear to me that the identification of predicativity with the ramified systems of Feferman and Schütte cannot be correct. A first, very simple, problem with it is that no predicativist (and for that matter — no mathematician) thinks in terms of ramified systems. Moreover: Feferman admits that "ramified theories are unsuitable as a framework for the development of analysis" ([24]). Another problem, that was repeatedly noted by Feferman himself, is that the general notions of well-ordering and ordinal on which they are based are not predicatively acceptable.[4] Third, and most important: just the description and understanding of the ramified hierarchy rely on principles that are not included in the theories in [15] and [42] which are based on this hierarchy. This is rather clear in case α is a limit ordinal: R_α is actually defined in this case as $\bigcup\{R(x) \mid x \in \{\beta \mid \beta < \alpha\}\}$. Hence it is based on accepting

[4]Feferman explicitly wrote in [24] that in his view, a system considered adequate for predicativity "should not be taken to involve the notions of ordinal or well-ordering in any way that is not already contained in the basic concepts of predicativity". This is the reason why over the years he has developed three different characterizations of predicativity that do not rely on the notion of ordinal. (See [19, 21, 26].) Γ_0 is still the proof-theoretical ordinal of the three corresponding systems. However, this is achieved in each case by imposing unjustified constrains on the applications of some of the principles on which the systems are based. This was first observed by Weaver in [49]. In the case of the characterization given in [19] (about which Feferman admits in [21] that it "may still be considered more persuasive" than the one given in that paper) this will be clearly shown in the sequel.

at least some instances of **ZF**'s axioms of union and replacement.[5] However, if we let $R_0 = N$ (which is the predicativist natural starting point, rather then $R_0 = \emptyset$), then similar problems exist even if we do not use transfinite ordinals, but only the natural numbers (as Russell did). Thus already in constructing elements of R_2 we allow quantification over R_1. This should mean that R_1 is taken as a "complete totality". But R_1 is *not* obtained using just "quantification only over previously defined totalities", and it is actually unclear what R_1 is at the first place. Usually it is identified with the collection (set?) of arithmetical subsets of N. If so, then each element of R_1 is indeed obtained using just quantification only over N. But this does not mean that so is R_1 itself. (For example: if we identify ordinals with von Neumann's ordinals, then Γ_0 is not predicatively definable according to the Γ_0-thesis, even though each element of Γ_0 is.) In fact, the only reasonable definition of the collection of arithmetical sets I can think of is the following:

$$\{A \mid \exists n \in N(predicae(n) \wedge \forall k \in N(k \in A \leftrightarrow n(\overline{k}) \in \mathbf{True})\}$$

where **True** is the set of true sentences in the first-order language of **PA**, and $predicae(n)$ means that n is a formula with exactly one free variable. (As usual, we identify here a formula with its Gödel number.) However, this definition relies on the availability of the set **True**. which is *not* arithmetical. It follows that if R_1 is the collection of arithmetical sets then it can only be defined using a set that at best belongs to some R_i such that $i > 1$, and so is defined in terms of R_1. This is obviously circular. To construct the ramified hierarchy we have therefore either to accept in addition to N infinitely many other sets as 'given' to us, or to realize that some hidden predicatively acceptable principles are involved already in the passage from N to R_1. The first option completely goes against the central ideas of Poincaré and Weyl. So we are left with the second.

NOTE 1. A close inspection shows that in general, the passage from D to $D^\star$ implicitly involves accepting two principles:

- One may use the model-theoretic operation that associates with any formula $\varphi(x)$ of second-order arithmetic the set $\{n \in N \mid N \models (\varphi(n))^D\}$.
- One may take as predicatively valid the instance of the replacement axiom that allows us to construct the image of this model-theoretic operation.

NOTE 2. Actually, some other indications that the 'Γ_0-thesis' is wrong are known. For example, Γ_0 is the proof-theoretical ordinal of ATR_0, which is one of the 'big five' theories which were studied in reverse mathematics ([45]). Therefore ATR_0 is considered to be *locally predicative*, i.e. all its theorems are taken to be predicatively acceptable. In contrast, the proof-theoretical ordinal of the full theory ATR is the much bigger Γ_{ϵ_0}. Hence the Γ_0-thesis implies that ATR is not locally predicative. But the only difference between ATR and ATR_0 is that the single induction *axiom* of the latter is replaced in the former by the full induction *scheme*. However, the induction scheme is universally accepted as

[5] It is rather strange that there is no attempt in e.g. [24] to justify this obvious use of the union operation, even though only a few pages before this use is made, it is argued that the set-theoretical operation of union is not predicatively acceptable. I have never been able to understand this straightforward incoherence.

being predicatively justified, and is actually included in the finitary systems of [15]. Therefore the theorems of *ATR* should be considered as no less predicatively acceptable than those of ATR_0.

1.3. Poincaré and Weyl's View of Predicativity as Invariance. As noted above, a notion of predicativity which is quite different from the Russellian one was introduced by Poincaré in [37]. This was particularly emphasized in [11]:

> "For Poincare impredicative definitions are problematic as they treat as completed infinite classes which are instead open-ended or incomplete by their very nature. Predicative definitions, instead, guarantee that the classes so defined are stable and invariant."

In Poincaré own words:

> "Hence a distinction between two species of classifications, which are applicable to the elements of infinite collections: the predicative classifications, which cannot be disordered by the introduction of new elements; the non predicative classifications, which are forced to remain without end by the introduction of new elements."

This view of predicativity underlies Weyl's great work in [51]. A careful reading of this book and of related papers of Weyl on the subject ([2, 6]) shows that predicativity as invariance is based in his work on the following principles:[6]

1. Sets are 'produced' *genetically*, that is: from applying legitimate operations to sets which are accepted as basic, or had previously been produced.
   > "If we imagine, as is appropriate for an intuitive understanding, that the relations and corresponding sets are 'produced' genetically, then this production will ... occur in merely parallel individual acts (so to speak)." ([53], P.40]
2. Accordingly, the elements of a set are logically *prior* to that set.
3. Sets are extensional, and the identity of a set is fully determined by the identity of its elements — sets that have the same elements are identical.
4. There is no single, complete intended universe $\mathbf{V}$ of sets. The 'universe of sets' is created in stages, and is always *open and growing*. To each stage corresponds what Weyl called a *sphere of operation* (i.e. a definite universe of sets equipped with some (finite) collection of predicates and operations) in which terms and formulas take values.
   > "Thus, contrary to Cantor's proposal, no universal scale of infinite ordinal and cardinal numbers applicable to every sphere of operation can exist." ([53], P.24)

   On the other hand (and in contrast):
   > "The numbers can (in any sphere of operation) be used to determine the cardinality of sets" ([53], P.55)
5. The current sphere of operation may be expanded in the future, e.g. by introducing *new legitimate methods of defining sets*, which in turn might produce new sets. The truth values of formulas may then be changed.

[6]Most, if not all, of these principles have been accepted already by Poincaré himself. See Section 2 of [27] for an excellent analysis of his views on the subject.

> "If we regard the principles of definition as an *'open'* system, i.e., if we reserve the right to extend them when necessary by making additions, then in general the question of whether a given function is continuous must also remain *open* (though we may attempt to resolve any *delimited* question). For a function which, within our current system, is continuous can lose this property if our principles of definition are expanded and, accordingly, the real numbers 'presently' available are joined by others." ([53], P.87)

6. Values of terms and truth-values of formulas are always evaluated *with respect to some definite sphere of operation* — never with respect to the whole open 'universe of sets'. Hence *classical logic is accepted as valid.* For example, $\neg\varphi \vee \varphi$ is valid in any sphere of operation, even though the truth value of φ may change when the current sphere of operation is expanded.
7. In contrast, the identity of already existing objects (and so the value of terms) should remain the same even if the current sphere of operation is expanded. Accordingly, a definition of an object is legitimate (or 'predicative') if, and only if, the identity of the object it defines is *invariant* under extension. (Because of this principle, there are certain constraints in Weyl's system on the use of quantification in *definitions.* However, there are no constrains in that system on the use of quantifiers in building *formulas.*) Weyl called an object so defined "a definite, self-existent object". Similarly, a definition of an operation is legitimate if the results of its applications depend only on the identity of the arguments, but not on the specific sphere of operation in which the application is made.
8. Any *theory* we develop should be true not only in the current sphere of operation, but in *any future one.* Hence our current theories impose constraints on future spheres of operation. Accordingly, expanding our spheres of operation and extending our theories are done simultaneously. Moreover:

   > "Our principles for the formation of derived relations can be formulated as axioms concerning sets and functions; and, in fact, mathematics will proceed in such a way that it draws the logical consequences of these axioms." ([53], p. 44)
9. The predicates of elementhood ($\in$) and equality ($=$) are basic.
10. Using ramification in definitions, or classifying sets according to 'levels', *should be avoided.*

    > "The temptation to pass beyond the first level of construction must be resisted; instead, one should try to make the range of constructible relations as wide as possible by enlarging the stock of basic operations." ([52])

It should be emphasized that according to the last quotation, principle 10 is not due just to the inconvenience which is caused by using ramified systems. There is also a *direct conflict* between Russell's approach and Weyl's approach. The former is based on the view that there are no predicatively legitimate methods of defining subsets of N beyond those allowed in the construction of the hierarchy of R_αs (and so the union of the R_αs includes all the predicative subsets of N). In contrast, Weyl's approach is based on the open-ended of the class of predicatively accepted methods of definitions. Thus already what is called 'R_1' in the Russellian approach is an *open* collection according to Weyl's view. (This point was missed by Feferman in [20]. Therefore when he discovered that Weyl's operation of iteration makes it possible to define non-arithmetical subsets of N even though in the above quotation Weyl explicitly refuses to go beyond 'R_1', Feferman wrongly interpreted this fact as an incoherence in Weyl's work. See [6] for further details.)

NOTE 3. Actually, the invariance criterion has been used by Feferman too in some of his *unramified* systems. Thus, **IR**, the first of the two unramified second-order locally predicative systems which were given in [15], uses the Hyperarithmetic Comprehension Rule Δ^1_1-CR. This is justified in [24] as follows:

> "The motivation for Δ^1_1-CR is the recognizable absoluteness (or invariance) of provably Δ^1_1 definitions, in the following sense. At each stage one has recognized certain closure conditions on the 'open' universe of sets, and the definitions $D(x)$ of sets introduced at the next stage should be independent of what further closure conditions may be accepted. In the words of Poincaré, the definitions used of objects in an incomplete totality should not be "disturbed by the introduction of new elements." Thus if U represents a universe of sets (subsets of N) satisfying given closure conditions and is extended to S' (satisfying the same closure conditions and possibly further ones) one wants D to be provably invariant or absolute in the sense that $(\forall x)[D^U(x) \leftrightarrow D^{U'}(x)]$. This requirement is easily seen to hold for provably Δ^1_1 formulas D."

This passage contains a rather accurate description of the invariance criterion, but ignores the conflict this criterion has with the Russellian approach, on which Feferman's canonical ramified systems are based. (In fact, I have not been able to find in any of Feferman's papers an explanation of the connection between the invariance criterion and the Russellian approach.)

1.4. Predicative Set Theories: Why and How.

1.4.1. *Why.* Feferman's systems (and to a lesser degree also Weyl's system in [51], as formalized in [6]) have one big drawback: they are practically inaccessible to the majority of the mathematical community. We believe that the major reason for this is that those systems do not use the framework of axiomatic set theory, which is almost universally accepted as the basic framework that provides the foundations of mathematics. What is more: Feferman's systems are by far more complicated in comparison to impredicative axiomatic set theories like **ZF**, which are currently used for developing the whole of present day mathematics.

Another reason to prefer the set-theoretical framework is that some of its principles are anyway underlying the constructions on which the second-order ramified systems of Feferman and Schütte are based. Thus we have seen in Section 1.2 that the construction of the hierarchy of R_αs uses instances of the union and replacement axioms of **ZF**. But in what cases is such a use predicatively justified? It seems to me that developing predicative set theories is the only way to answer such questions. What is more, the notion of ordinal, which is crucial for the ramified systems but is also very problematic in their context, is not problematic at all in the set-theoretical one. There one can use the notion of von Neumann ordinals, and those are defined by a simple, absolute formula.

Finally, it is worth noting that predicativism was born as a reaction to the set-theoretical paradoxes, and was intended to provide a satisfactory solution to them. So (at least in my opinion) it should be most natural to develop predicative mathematics in the framework in which it has started.

NOTE 4. Locally predicative (in the sense of being proof-theoretically reducible to the systems of [15]) set theories have already been introduced by Feferman in [16, 18]. We shall say more about them in Section 4.3.

1.4.2. *How.* In the case of pure set theories, the main principle of the predicativity as invariance view of Poincaré and Weyl can be expressed as follows: *a set exists if and only if can be determined by an invariant definition.* Accordingly, the main two features of the system $\mathcal{PW}$ which is developed in the sequel are:

(I) Any existence claim which is made in one of those axioms of $\mathcal{PW}$ whose purpose is to allow the expansion of the sphere of operation is actually an existence and uniqueness claim. In other words: positive occurrences of the quantifier $\exists$ in such an axiom are in the form $\exists!$. (This, of course, rules out the axiom of choice, as well as the axiom of Δ_0-collection of Kripke-Platek set theory as given in [9].)

(II) Following Principles 1 and 7 in Section 1.3, any definition of a set which is made in $\mathcal{PW}$ is *invariant*. This is ensured by employing a syntactically defined *invariance relation* $\succ$ between formulas and finite sets of variables.

NOTE 5. Principle (I) is not applicable to general validities like logically valid formulas or instances of $\in$-induction. A trivial example is given by $\forall x \neg\varphi \vee \exists x\varphi$, where φ is arbitrary. In general theorems of $\mathcal{PW}$ of this sort cannot be used for introducing new sets, or for providing absolute identification of existing ones.

The following other important features of $\mathcal{PW}$ also directly correspond to the principles of Weyl and Poincaré that were described in Section 1.3:

1. Like in **ZF**, and unlike in the systems of Weyl and Feferman, our system has a single type (or 'category', in Weyl's terminology) of objects: sets. (This corresponds to Principle 10 in Section 1.3.)
2. Like in most of Feferman's systems, $\mathcal{PW}$ is practically not really a single theory, but involves many theories, all of them *first-order*. In each stage of working within it, we do have a single theory $\mathbf{T}$, but we have two options how to proceed: We may simply derive new theorems in $\mathbf{T}$, but we may also move to a *strictly stronger* theory $\mathbf{T}^\star$ in an expanded language.[7] (This feature of $\mathcal{PW}$ implements Principles 4,5, and 8 in Section 1.3.)
3. The logic of all theories in $\mathcal{PW}$ is *classical* logic. (Principle 6 in Section 1.3.)
4. The initial language of the system includes just two predicate symbols: $=$ and $\in$ (Principle 9 in Section 1.3) and a constant ω for the set of natural numbers (taken to be the finite von Neumann ordinals). The inclusion of

[7]Practical work with any ordinary first-order theory $\mathbf{T}$ also always involves the use of the procedure of extension by definitions (see e.g. [44]), which also allows moving from $\mathbf{T}$ to an extension $\mathbf{T}^\star$ in an expanded language. However, this $\mathbf{T}^\star$ is a *conservative* extension of $\mathbf{T}$, and is no more than just an equivalent variant of it. This is not the situation in $\mathcal{PW}$.

the latter is actually *not* essential, but it reflects well the central place that the natural numbers have in the predicativism of Poincaré and Weyl.

5. The following axiom and axiom schema are included in all theories in $\mathcal{PW}$:
   - The axiom of extensionality (Principle 3 in Section 1.3);
   - Comprehension for invariant formulas (Principle 7 in Section 1.3);
   - $\in$-induction (which implements the vague Principle 2 in Section 1.3).[8]
6. Our main method of extending a given predicative set theory $\mathbf{T}$ to a stronger predicative set theory $\mathbf{T}^\star$ is by *adding a new symbol* to the signature of $\mathbf{T}$, together with an axiom that defines it. (In addition, we include of course in $\mathbf{T}^\star$ all the instances in the extended language of the axiom schemas of $\mathbf{T}$.) Such an extension is done by applying one of the syntactic methods that $\mathcal{PW}$ provides for this purpose.
7. As usual, extending $\mathbf{T}$ by a operation symbol is allowed only if $\mathbf{T}$ proves some corresponding *existence and uniqueness conditions*. Still, the extension is usually *not* conservative.
8. Adding an n-ary predicate symbol P is allowed only if its defining axiom implies its *absoluteness*. Similarly, adding an n-ary operation symbol F is allowed only if its defining axiom implies that the formula $y = F(x_1, \ldots, x_n)$ is *invariant* with respect to y in case $y, x_1, \ldots, x_n$ are distinct.

NOTE 6. In designing $\mathcal{PW}$ we have adopted two additional principles:

- *Platonists should be able to accept any theory in our framework.* In particular: every theory in our framework is a subsystem (of some extension by definitions) of **ZF**. This principle immediately rules out, e.g. Axiom VIII (Enumerability) of the system $\mathbf{PS_1E}$ from [18] and the Axiom of Countability of the system $\mathbf{ATR_0^{Set}}$ from [45], which say that every set is countable.[9]
- Every rule or axiom of $\mathcal{PW}$ is a very close counterpart of some rule or axiom that was used by Feferman in one of his predicative (or locally predicative) systems. Hence $\mathcal{PW}$ should be accepted as predicative by anyone who accepts those system of Feferman's as predicative.

NOTE 7. Among the axioms of **ZFC**, $\mathcal{PW}$ completely rejects the axiom of powerset and the axiom of choice, and it restricts the axiom schema of separation to the case in which the separating condition is absolute. It also accepts only special cases of the axiom schema of replacement. In our opinion, these are properties that should be shared by any predicative set theory.

1.5. The Structure of the Paper. Section 2 explains our notations and terminology. In Section 3 we establish the predicativity of the set of natural numbers, using a rather weak subsystem of $\mathcal{PW}$. $\mathcal{PW}$ itself is precisely defined, justified, and compared with Feferman's Systems in Section 4. Section 5 includes important examples of the power of $\mathcal{PW}$. Section 6 develops in $\mathcal{PW}$ the fundamentals of the theory of von Neumann ordinals. Section 7 includes the main results of this paper: it shows that $\mathcal{PW}$ provides terms which define Γ_0 and much bigger ordinals, and that it can prove the main properties of those ordinals. We conclude in Section 8 with some remarks and directions for further research.

[8] In Section 4.2 we shall discuss the justification of full $\in$-induction in greater detail.

[9] Another reason to reject these axioms is that the notion of being countable is not absolute. Recall that Weyl rejected the idea of assigning different cardinalities to infinite sets.

§2. Terminology and Notations.

2.1. The Difference Between Operations and Functions. In standard textbooks on first-order theory it is common to refer to the symbols in a signature of a first-order language as 'relation symbols' and 'function symbols'. We cannot use this terminology here, since we reserve the words 'relation' and 'function' to their official meaning in set theory, that is: to *sets* of pairs satisfying certain conditions. Instead, we use the name 'predicate' for any relation that is not a set (like the predicates $\in$ or $=$), and we use the name 'operation' for any 'function' that is not a set (like the operation of union on sets or the operation of addition on ordinals).[10] Accordingly, the symbols of a first-order signature are divided in this paper into predicate symbols, operation symbols, and constants. The latter may actually be viewed as operations with arity 0, except that they should always be interpreted as sets.

2.2. Notations. We use small letters from the beginning of the Latin alphabet as variables in the metalanguage for sets, and i, j, k, l, m, n as special variables for natural numbers (in both the metalanguage and the formal language). t and s will serve as variables (in the metalanguage) for terms, and $\varphi, \psi, \theta, A, B, C$ as variables for formulas. In all cases the variables may be decorated with subscripts or superscripts. We denote by $Fv(\varphi)$ ($Fv(t)$) the set of free variables of φ (of t). When we denote a formula by $\varphi(x_1, \ldots, x_n)$ it means that $\{x_1, \ldots, x_n\} \subseteq Fv(\varphi)$. On the other hand when we write $\varphi(\vec{y}, x)$ it means that $\vec{y} = \langle y_1, \ldots, y_n\rangle$ for some n (whose identity may be obtained from the context or it does not matter); the variables $x, y_1, \ldots, y_n$ are all distinct from each other; and $Fv(\varphi) = \{x, y_1, \ldots, y_n\}$.

The substitution of a term t for a free variable y in a formula φ (a term s) is denoted by $\varphi\{t/y\}$ ($s\{t/y\}$). However, when we denote a formula by $\varphi(y)$ ($\varphi(x, y)$, $\varphi(\vec{x}, y)$) we might simply write $\varphi(t)$ ($\varphi(x, t)$, $\varphi(\vec{x}, t)$) instead.

Given a first-order signature σ, we take a structure for σ to be a pair $\mathcal{M} = \langle \mathcal{D}, I\rangle$, where $\mathcal{D} \neq \emptyset$ is the domain of $\mathcal{M}$ and I is its interpretaton function. If r is one of the symbols in σ we shall usually write r^I instead of $I(r)$. If ν is an assignment of elements of $\mathcal{D}$ to variables of the language, $x_1, \ldots, x_n$ are n distinct variables, and $\vec{a} \in \mathcal{D}^n$, we denote by $\nu\{\vec{x} := \vec{a}\}$ the assignment which is obtained from ν by assigning a_i to x_i ($i = 1, \ldots, n$). We denote by $\nu_{\mathcal{M}}[t]$ the element of $\mathcal{D}$ that ν assigns according to I to the term t of σ. Similarly, if f is an operation symbol of σ then we use in the metalanguage square brackets to denote applications of f^I to arguments in $\mathcal{D}$. Thus, if f is n-ary, and ν is an assignment in $\mathcal{D}$, then $\nu_{\mathcal{M}}[f(t_1, \ldots, t_n)] = f^I[\nu_{\mathcal{M}}[t_1], \ldots, \nu_{\mathcal{M}}[t_n]]$. We write $\mathcal{M}, \nu \models \varphi$ in case ν satisfies in $\mathcal{M}$ the formula φ of σ. If $Fv(\varphi) = \{x_1, \ldots, x_n\}$, and $\nu[x_i] = a_i$ ($i = 1, \ldots, n$) then we might write instead $\mathcal{M} \models \varphi(\langle a_1, \ldots, a_n\rangle)$.

Finally, when we refer in the metalanguage to the collection of things that satisfy a certain condition C we shall denote it by $[x \mid C(x)]$, reserving the notation $\{x \,|\, \varphi(x)\}$ for being used in our formal system. Moreover, in case there is a danger of confusion, we shall use ':' in the metalanguage instead of '$\in$'. (Recall that the latter is a basic symbol of the language of our system.)

[10]Thus, what are usually called 'the rudimentary functions' ([28, 14]) are called here 'the rudimentary operations'.

§3. The Predicativity of the Natural Numbers. Our system $\mathcal{PW}$ includes a constant ω for the natural numbers. Before presenting $\mathcal{PW}$ in the next section, we would like to justify this inclusion. We do that by:

1. Providing a *bounded* formula $N(x)$ in the language of set theory that defines when x is a natural number (i.e. a finite von Neumann ordinal).
2. Present a basic predicative set theory **VBS**, in which one can show that $N(x)$ is adequate for the task. This is done by proving in it all the properties that one expects from a formula that defines the natural numbers;
3. Give an intuitive proof in the metalanguage that the formula $N(x)$ is invariant, and so it may be used for defining a new set.

We start by presenting **VBS**.[11] The axioms of this system include the Extensionality axiom [Ext] and the $\in$-induction axiom schema [$\in$-ind] from Section 4.1.4 below, as well as the following four elementary instances of the general predicative comprehension scheme:

1. Empty Set ([Em]):
$$\exists Z \forall x(x \in Z \leftrightarrow x \neq x)$$
2. Pairing ([Pa]):
$$\forall x \forall y \forall \exists Z \forall w(w \in Z \leftrightarrow w = x \vee w = y)$$
3. Union ([U]):
$$\forall x \exists Z \forall w(w \in Z \leftrightarrow \exists y(y \in x \wedge w \in y))$$
4. Difference ([D]):
$$\forall x \forall y \exists Z \forall w(w \in Z \leftrightarrow w \in x \wedge w \notin y)$$

Note 8. It is easy to see that the structure $\langle \mathsf{HF}, \in \rangle$, where HF (which is identical to V_ω) is the set of the hereditarily finite sets, forms the minimal model of **VBS**. Moreover, $\langle V_\alpha, \in \rangle$ is a model of **VBS** whenever α is a limit ordinal.[12]

In order to present the formula $N(x)$ it is convenient (though not really necessary) to use the usual procedure of extension by definitions, and develop **VBS** in an enriched language in which the four axioms above are replaced by:[13]

1. [Em]: $\forall x(x \notin \emptyset)$
2. [Pa]: $\forall x \forall y \forall w(w \in \{x, y\} \leftrightarrow w = x \vee w = y)$
3. [U]: $\forall x \forall w(w \in \bigcup x \leftrightarrow \exists y(y \in x \wedge w \in y))$
4. [D]: $\forall x \forall y \forall w(w \in x - y \leftrightarrow w \in x \wedge w \notin y)$

Note 9. In general, one should be careful when applying the extension by definitions procedure to theories with axioms schemas, since the extension of such a schema to the expanded language involves the addition of infinitely many new axioms besides those that are allowed by the procedure. This is not a problem in cases like we have here, where every axiom schema is pure in the sense that no

[11]For the material of this section a much weaker system would suffice. However, we need the full power of **VBS** for developing the basic theory of ordinals in Section 6.

[12]From our predicativist point of view, V_α is not a set in case $\alpha > \omega$, but only a class. By this we mean that there is an absolute formula $\varphi(\alpha, x)$ that defines the predicate '$x \in V_\alpha$'.

[13]These applications of the extension by definitions procedure are permissible also according to the restricted version of this procedure which is allowed in $\mathcal{PW}$.

constraint is imposed on the formulas to which it may be applied. Therefore we may assume that every instance of $[\in\!-ind]$ in the expanded language is an axiom of **VBS**. However, one should be cautious about the issue of being conservative when the procedure is applied in more complicated cases, like in case we have $[\in\!-ind]$ restricted to some class of formulas (e.g. bounded formulas).

PROPOSITION 1. *For every $n \geq 0$:*

$$\vdash_{\mathbf{VBS}} \forall x_0 \forall x_1 \in x_0 \forall x_2 \in x_1 \cdots \forall x_n \in x_{n-1}. x_0 \notin x_n$$

Proof. We do the case $n = 2$ (which trivially implies the cases $n = 0$ and $n = 1$). The proof for any other n is similar.

Given x_0, to show that $\forall x_1 \in x_0 \forall x_2 \in x_1. x_0 \notin x_2$, we may assume (using $[\in\!-ind]$) that $(\star)$ $\forall y_0 \in x_0 \forall y_1 \in y_0 \forall y_2 \in y_1. y_0 \notin y_2$. Suppose now that there are x_1 and x_2 such that: $x_1 \in x_0 \wedge x_2 \in x_1 \wedge x_0 \in x_2$ Since $x_1 \in x_0$, we may apply $(\star)$ with $y_0 = x_1$, $y_1 = x_2$, and $y_2 = x_0$, to get that $x_1 \notin x_0$, which is a contradiction. So no such x_1 and x_2 exist. $\square$

We are ready to introduce our definition of the notion of a natural number:

DEFINITION 1.

1. $S(x) := x \cup \{x\}$ (where $\{x\} = \{x, x\}$).
2. $N(x) := \forall y \in S(x)(y = \emptyset \vee \exists z \in x. y = S(z))$

NOTE 10. Officially, $N(x)$ is the following formula:

$$\forall y((y \in x \vee y = x) \to (\forall z(z \notin y) \vee \exists z \in x \forall w(w \in y \leftrightarrow (w \in z \vee w = z))))$$

PROPOSITION 2. *Let $0 := \emptyset$. The following are provable in* **VBS***:*

1. $N(0)$.
2. $\forall x(N(x) \leftrightarrow N(S(x)))$.
3. $S(x) \neq 0$
4. $S(x) = S(y) \to x = y$

Proof. 1. and 3. are trivial. 4. easily follows from Proposition 1. We show 2.

- Suppose that $N(x)$. We show that $N(S(x))$, i.e. that

  $$\forall y \in S(S(x))(y = 0 \vee \exists z \in S(x). y = S(z))$$

  So let $y \in S(S(x))$. We show that $y = 0 \vee \exists z \in S(x). y = S(z)$. Since we assume $N(x)$, this is obvious in case $y \in S(x)$, because every $z \in x$ is also in $S(x)$. There remains the case $y = S(x)$, but this case is trivial, since $x \in S(x)$ (and so x is an element z in $S(x)$ such that $y = S(z)$).
- Suppose that $N(S(x))$, i.e.

  $$\forall y \in S(S(x))(y = 0 \vee \exists z \in S(x). y = S(z))$$

  We show that $N(x)$. So let $y \in S(x)$. We show that $y = 0 \vee \exists z \in x. y = S(z)$. Suppose that the first disjunct fails, i.e. $y \neq 0$. Since $S(x) \subseteq S(S(x))$, our assumption implies that there is $z \in S(x)$ such that $y = S(z)$. It is impossible that $z = x$, since in this case we would get that $S(x) \in S(x)$. Hence $z \in x$, and we are done. $\square$

PROPOSITION 3. $\vdash_{\mathbf{VBS}} \varphi\{0/x\} \wedge \forall x(\varphi \to \varphi\{S(x)/x\}) \to \forall x(N(x) \to \varphi)$

Proof. Assume (I) $\varphi\{0/x\} \wedge \forall x(\varphi \to \varphi\{S(x)/x\})$. To show $\forall x(N(x) \to \varphi)$, it suffices by $[\in\!-ind]$ to show that $\forall x(\forall y \in x(N(y) \to \varphi\{y/x\}) \to (N(x) \to \varphi))$. So assume that (II) $\forall y \in x(N(y) \to \varphi\{y/x\})$ and (III) $N(x)$. We show φ. If $x = 0$ then this is implied by (I). If not, then it follows from (III) that there is $z \in x$ such that $x = S(z)$. Hence (II) entails that $N(z) \to \varphi\{z/x\}$. But $N(z)$ follows from (III) by the second item of Proposition 2. Therefore $\varphi\{z/x\}$. From this φ (which is equivalent to $\varphi\{S(z)/x\}$) follows by (I). □

PROPOSITION 4. *Let $<:=\in$. The following are provable in* **VBS***:*

1. $x \not< 0$
2. $x < S(y) \leftrightarrow x < y \vee x = y$
3. $N(x) \wedge y < x \to N(y)$

Proof. The first two items are trivial. We prove the third by induction on x (i.e. by using Proposition 3). So let $\varphi := \forall y(N(x) \wedge y < x \to N(y))$. Obviously, $\vdash_{\mathbf{VBS}} \varphi\{0/x\}$. We show that $\vdash_{\mathbf{VBS}} \forall x(\varphi \to \varphi\{S(x)/x\})$. So assume $\varphi(x)$ and that $N(S(x))$ and $y < S(x)$. By Proposition 2, the first assumption implies that $N(x)$. Hence $y = x$ implies that $N(y)$. Otherwise $y < x$, and so the induction hypothesis implies that $N(y)$. □

Finally we show that $N(x)$ (intuitively) defines an invariant collection. Since this is an intuitive theorem in the meta-language, the proof is intuitive too. Still, it employs only predicatively acceptable principles.

PROPOSITION 5. *Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be transitive models of* **VBS** *(in the usual sense of set theory) such that $\mathcal{M}_1 \subseteq \mathcal{M}_2$. Then:*

$$\omega_2 = [x : \mathcal{M}_2 \mid \mathcal{M}_2 \models N(x)] = [x : \mathcal{M}_1 \mid \mathcal{M}_1 \models N(x)] = \omega_1$$

Proof. The fact that $N(x)$ is bounded (and so absolute) implies that $\omega_1 \subseteq \omega_2$. For the converse, we show that $\forall a \in \mathcal{M}_2((\mathcal{M}_2 \models N(a)) \to (a \in \mathcal{M}_1 \wedge \mathcal{M}_1 \models N(a)))$. So let $a \in \mathcal{M}_2$, and assume that $\mathcal{M}_2 \models N(a)$. By $[\in\!-ind]$ we may assume also that $\forall b \in a((\mathcal{M}_2 \models N(b)) \to (b \in \mathcal{M}_1 \wedge \mathcal{M}_1 \models N(b)))$. Since $\mathcal{M}_2 \models N(a)$, either $a = \emptyset$, or there exists $b \in a$ such that $\mathcal{M}_2 \models a = S(b)$. In the first case $a \in \mathcal{M}_1$ holds trivially. So assume the second case. Then item 2 of Proposition 2 implies that $\mathcal{M}_2 \models N(b)$. Hence the induction hypothesis implies that $b \in \mathcal{M}_1 \wedge \mathcal{M}_1 \models N(b)$. Again by Proposition 2, this entails that $S(b) \in \mathcal{M}_1 \wedge \mathcal{M}_1 \models N(S(b))$. Hence $a \in \mathcal{M}_1 \wedge \mathcal{M}_1 \models N(a)$. □

§4. The System $\mathcal{PW}$.

4.1. A Description of the System. The language of the system $\mathcal{PW}$ is a one-sorted first-order language with equality. Like in any first-order language, it has infinitely many variables, taken here to be $v_0, v_1, \ldots$, with letters from the end of the alphabet to vary over them. All other components of the system (terms, formulas, the invariance relation $\succ$, Σ-formulas, predicate and operation symbols, axioms, rules and proofs) are simultaneously generated as described below. (Note that in the formulation of the last rule [Unif] there is a use of the formula *Fun*, the operation *Dom*, and the term $f(x)$. They are all introduced in Section 5.1 without using [Unif]. $Fun(f)$ says that f is a function. $Dom(f)$ and $f(x)$ have their usual meaning.)

4.1.1. *Predicate Symbols and Operation Symbols.*

1. $=$ and $\in$ are binary predicate symbols.
2. If $\varphi \succ \emptyset$ and $Fv(\varphi) = \{v_0, \ldots, v_n\}$ then P_φ is an $n+1$ predicate symbol.
3. If φ is Σ and $Fv(\varphi) = \{v_0, \ldots, v_n\}$ then F_φ is an n-ary operation symbol.

4.1.2. *Terms.*

1. Every variable is a term.
2. The constant ω is a term. (A constant is a 0-ary operation symbol.)
3. $F(t_1, \ldots, t_n)$ is a term if F is an n-ary operation, and $t_1, \ldots, t_n$ are terms.

4.1.3. *Formulas.*

1. $P(t_1, \ldots, t_n)$ is a formula if P is an n-ary predicate, $t_1, \ldots, t_n$ are terms.
2. The formulas are closed under $\neg$, $\wedge$, $\vee$ and $\rightarrow$.
3. If φ is a formula and x is a variables, then $\exists x\varphi$ and $\forall x\varphi$ are formulas.

4.1.4. *Axioms.*

1. [Fol]: Every formula which is valid in first-order logic with equality.[14]
2. [Ext] (Extensionality): $\forall z(z \in x \leftrightarrow z \in y) \rightarrow x = y$.
3. [Comp] ($\succ$-Comprehension): $\exists! Z\forall x(x \in Z \leftrightarrow \varphi)$, provided that $\varphi \succ \{x\}$.
4. [$\in$-ind] ($\in$-induction): $(\forall x(\forall y(y \in x \rightarrow \varphi\{y/x\}) \rightarrow \varphi)) \rightarrow \forall x\varphi$.
5. [Inf] (Infinity): $\forall x(x \in \omega \leftrightarrow N(x))$, where $N(x)$ is presented in Note 10.
6. [PrI] $P_\varphi(v_0, \ldots, v_n) \leftrightarrow \varphi$ (provided that φ is as in 4.1.1).

4.1.5. *Rules.*

1. [MP]: From φ and $\varphi \rightarrow \psi$ infer ψ.
2. [Gen]: From $\vdash_{\mathcal{PW}} \varphi$ infer $\vdash_{\mathcal{PW}} \forall x\varphi$.
3. [OpI] From $\vdash_{\mathcal{PW}} \forall v_1 \cdots \forall v_n \exists! v_0\varphi$ infer $\vdash_{\mathcal{PW}} F_\varphi(v_1, \ldots, v_n) = v_0 \leftrightarrow \varphi$ (provided that φ is as in 4.1.1, i.e. φ is Σ and $Fv(\varphi) = \{v_0, \ldots, v_n\}$).
4. [Unif] (Unification Rule for Operations):

$$\frac{\vdash_{\mathcal{PW}} \forall y_1\forall y_2(\varphi\{y_1/y\} \wedge \varphi\{y_2/y\} \rightarrow y_1 = y_2)}{\vdash_{\mathcal{PW}} \forall x \in Z\exists y\varphi \rightarrow \exists! f(Fun(f) \wedge Dom(f) = Z \wedge \forall x \in Z\varphi\{f(x)/y\})}$$

   Provided φ is Σ, x and y are distinct variables in $Fv(\varphi)$, and $Z \notin Fv(\varphi)$.

4.1.6. *The invariance relation $\succ$.*

1. $t \in s \succ \emptyset$ and $t = s \succ \emptyset$.
2. $\varphi \succ \{x\}$ if $\varphi \in \{x \neq x, x = t, t = x, x \in t\}$ and $x \notin Fv(t)$.
3. $P_\varphi(t_0, \ldots, t_n) \succ X$ if $\varphi\{t_0/v_0, \ldots, t_n/v_n\} \succ X$.
4. $\neg\varphi \succ \emptyset$ if $\varphi \succ \emptyset$.
5. $\varphi \vee \psi \succ X$ if $\varphi \succ X$ and $\psi \succ X$.
6. $\varphi \wedge \psi \succ X \cup Y$ if $\varphi \succ X$, $\psi \succ Y$ and $Y \cap Fv(\varphi) = \emptyset$.
7. $\exists y\varphi \succ X - \{y\}$ if $y \in X$ and $\varphi \succ X$.

4.1.7. *Σ-formulas.*

1. If $\varphi \succ \emptyset$ then φ is Σ. Such formulas are called *absolute*.
2. If φ and ψ are Σ then so are $\varphi \vee \psi$ and $\varphi \wedge \psi$.
3. If φ is Σ then so is $\exists x\varphi$.
4. If $\varphi \succ \{y_1, \ldots, y_k\}$, and ψ is Σ, then $\forall y_1 \cdots \forall y_k(\varphi \rightarrow \psi)$ is Σ.

[14]Instead, we can of course choose any standard axiomatization of this logic.

Note 11. In the formulation above we have used the quantifier $\exists!$ in [Comp] and in [Ext]. This was done according to Principle (I) of Section 1.4.2. However, in both cases we can actually use the simpler connective $\exists$. In the case of [Comp] this is due to the axiom [Ext], while in the case of [Unif] it follows from the premise of that rule.

Note 12. Like in the system **PZF** from [5], the predicativity of definitions of sets is ensured $\mathcal{PW}$ by using an appropriate syntactic *invariance relation* $\succ$ between a formula φ and subsets of $Fv(\varphi)$. Relations of this sort provide a common generalization of the set-theoretical notion of absoluteness ([14]), and the notion of domain independence used in database theory ([46, 1]). They have originally been introduced in [3, 4] in order to provide a unified theory of constructions and operations as they are used in different branches of mathematics and computer science, including set theory, computability theory, and database theory. Further important theorems about them can be found in [8]. [15]

Note 13. It is easy to show by induction that if $\varphi \succ X$ and $Y \subseteq X$ then $\varphi \succ Y$. Therefore we have not included this important condition from [8] in our present definition of $\succ$, but we shall use it freely in what follows. Another important condition which we shall treat as if it included in the definition of $\succ$ is that $\forall x_1 \cdots \forall x_n(\varphi \rightarrow \psi) \succ \emptyset$ if $\varphi \succ \{x_1, \ldots, x_n\}$ and $\psi \succ \emptyset$. The reason is that every consequence of this condition can easily be derived without it (because $\forall x_1 \cdots \forall x_n(\varphi \rightarrow \psi)$ is logically equivalent to $\neg\exists x_1 \cdots \forall x_n(\varphi \wedge \neg\psi)$.)

Note 14. The clauses in the definition of Σ-formulas are taken from [4]. This definition is a straightforward generalization of the usual definition of Σ-formulas. In particular: it includes all the formulas which are called 'essentially existential HF-formulas' in [17]. From Theorem 4.1 of [17] it follows that every persistent formula is equivalent in $\mathcal{PW}$ to a Σ-formula, but we shall not use this fact here.

Note 15.

1. The fact that [PrI] applies only to absolute formulas ensures (together with Note 13) that $P(t_1, \ldots, t_n) \succ \emptyset$ whenever P is a predicate.
2. Actually, the use of [PrI] does not really increase the power of $\mathcal{PW}$, and so we can omit it from $\mathcal{PW}$. However, it is very convenient to include it.

Note 16. It should be emphasized that [MP] is the only rule of derivation of $\mathcal{PW}$. All the rest are only *rules of proof.* This means that they can be applied *only in assumptions-free deductions* (i.e. pure deductions from axioms). It follows that the deduction theorem obtains for $\mathcal{PW}$: to show $\Gamma \vdash_{\mathcal{PW}} (\varphi \rightarrow \psi)$ it suffices to prove that $\Gamma, \varphi \vdash_{\mathcal{PW}} \psi$. However, while using this theorem in order to show that $\Gamma \vdash_{\mathcal{PW}} (\varphi \rightarrow \psi)$, one should be careful not to rely on a proof of ψ from Γ, φ in which [Gen],[OpI] or [Unif] is applied to a formula which depends on an assumption in Γ, φ. (As usual, in the case of [Gen] one may actually infer in a derivation $\forall x \varphi$ from φ in case x is not free in any of the assumptions on which φ depends in that derivation.)

[15] Following standard terminology in database theory ([46]), we have used in our previous papers the name "safety relations" for this type of relations.

4.2. Justification of the System. From the non-logical axioms and rules of $\mathcal{PW}$, [Ext] needs no justification, [PrI] was justified in Note 15, and [Inf] in Section 3. It remains to justify the other non-logical axioms and *rules* of $\mathcal{PW}$.

4.2.1. [Comp] *and* [OpI]. Since invariance is our major criterion for predicativity, we start by showing that $\mathcal{PW}$ satisfies this criterion according to a precise notion of invariance which is adequate for the present context.

Obviously, talking about invariance of definitions of sets and operations, when we expand one sphere of operation $\mathcal{M}_1$ to a bigger one $\mathcal{M}_2$, can make sense only if the identities of the elements of $\mathcal{M}_1$ are preserved in $\mathcal{M}_2$. Since in the context of set theory we take the identity of a set to be fully determined by the identity of its elements, this means that the same objects should belong to an element a of $\mathcal{M}_1$ in both $\mathcal{M}_1$ and $\mathcal{M}_2$. Accordingly we define:

Definition 2. Let $\mathcal{M}_1 = \langle \mathcal{D}_1, I_1\rangle$ and $\mathcal{M}_2 = \langle \mathcal{D}_2, I_2\rangle$ be structures for signatures that contain $\in$, and let both be models of [Ext]. $\mathcal{M}_2$ is an $\in$-*extension* of $\mathcal{M}_1$ if $\mathcal{D}_1 \subseteq \mathcal{D}_2$, and the following holds for every element a of $\mathcal{D}_1$:

$$[x : \mathcal{D}_1 \mid x \in^{I_1} a] = [x : \mathcal{D}_2 \mid x \in^{I_2} a]$$

Note 17. From a platonic point of view, if $\in^{I_1}$ and $\in^{I_2}$ are both the 'real' $\in$ of V, then $\mathcal{M}_2$ is an $\in$-*extension* of $\mathcal{M}_1$ iff $\mathcal{D}_1$ is a transitive subset of $\mathcal{D}_2$.

Definition 3. Let e be a term or a formula or an operation symbol of $\mathcal{PW}$.

1. σ_e is the minimal signature σ that satisfies the following conditions:
   - It includes $\in$ and $=$.
   - It includes all the predicate symbols and operation symbols (including constants) that occur in e;
   - If either P_ψ or F_ψ is in σ, then so are all the predicate symbols and operation symbols that occur in ψ.
2. e is *legal* if the premise of [OpI] obtains whenever F_φ is in σ_e.
3. For legal e, $\mathcal{PW}_e$ is the set of all theorem of $\mathcal{PW}$ in the language of σ_e.

Definition 4. Let e be legal. A structure $\mathcal{M} = \langle \mathcal{D}, I\rangle$ is *adequate* for e if:

- $\mathcal{M}$ is a structure for a signature that contains σ_e;
- $\mathcal{M}$ is a model of $\mathcal{PW}_e$;

Note 18. Actually, we would have liked to include in the definition of adequacy the demand that $\in^I$ is well-founded, in the strong intuitive sense of this notion. By Proposition 5, this would force the structure $\langle [a \in \mathcal{D} \mid \mathcal{M} \models N(a)], \in^I\rangle$ to be isomorphic to $\langle \mathcal{N}, <\rangle$, where $\mathcal{N}$ is the structure of the natural numbers. Unfortunately, defining either 'well-founded' or 'isomorphic' in a predicatively appropriate way is problematic. (We return to this issue in Section 4.2.2.) To avoid the complications involved, we keep the above weak notion of adequacy, but explicitly include in the next definition an unproblematic weak corollary of demanding $\langle [a \in \mathcal{D} \mid \mathcal{M} \models N(a)], \in^I\rangle$ to be isomorphic to $\langle \mathcal{N}, <\rangle$.

Definition 5. For legal e, $\langle \mathcal{M}_1, \mathcal{M}_2\rangle$ is an *e-pair* if:

- $\mathcal{M}_1$ and $\mathcal{M}_2$ are adequate for e.
- $\mathcal{M}_2$ is an $\in$-*extension* of $\mathcal{M}_1$.
- $[a \in \mathcal{D}_2 \mid \mathcal{M}_2 \models N(a)] = [a \in \mathcal{D}_1 \mid \mathcal{M}_1 \models N(a)]$

Definition 6.

1. A legal term t of $\mathcal{PW}$ is *invariant* if $\nu_{\mathcal{M}_1}[t] = \nu_{\mathcal{M}_2}[t]$ whenever $\langle \mathcal{M}_1, \mathcal{M}_2\rangle$ is a t-pair, and ν is an assignment in $\mathcal{M}_1$.
2. A legal n-ary operation F of $\mathcal{PW}$ is *invariant* if
$$F^{\mathcal{M}_1}[a_1,\ldots,a_n] = F^{\mathcal{M}_2}[a_1,\ldots,a_n]$$
whenever $\langle \mathcal{M}_1, \mathcal{M}_2\rangle$ is an F-pair, and $a_1,\ldots,a_n$ are elements of $\mathcal{M}_1$.
3. A legal formula φ of $\mathcal{PW}$ such that $\{x_1,\ldots,x_n\} \subseteq Fv(\varphi)$ is *invariant* with respect to $\{x_1,\ldots,x_n\}$ if the following holds for every assignment ν in $\mathcal{M}_1$:
$$[\vec{a} : \mathcal{D}_2^n \mid \mathcal{M}_2, \nu\{\vec{x} := \vec{a}\} \models \varphi] = [\vec{a} : \mathcal{D}_1^n \mid \mathcal{M}_1, \nu\{\vec{x} := \vec{a}\} \models \varphi]$$
4. A legal formula φ of $\mathcal{PW}$ is *persistent* if $\mathcal{M}_1, \nu \models \varphi$ implies that $\mathcal{M}_2, \nu \models \varphi$ as well whenever $\langle \mathcal{M}_1, \mathcal{M}_2\rangle$ is a φ-pair, and ν is an assignment in $\mathcal{M}_1$.

Note 19. The definition of invariability of formulas can also be formulated as follows: a legal formula φ of $\mathcal{PW}$ such that $Fv(\varphi) = \{x_1,\ldots,x_n, y_1,\ldots,y_k\}$ is *invariant* with respect to $\{x_1,\ldots,x_n\}$ if the following holds whenever $\langle \mathcal{M}_1, \mathcal{M}_2\rangle$ is a φ-pair, and $c_1 \ldots , c_k$ are are elements of $\mathcal{M}_1$:
$$[\vec{a} : \mathcal{D}_2^n \mid \mathcal{M}_2 \models \varphi(\vec{a}^\frown\vec{c})] = [\vec{a} : \mathcal{D}_1^n \mid \mathcal{M}_1 \models \varphi(\vec{a}^\frown\vec{c})]$$
(where $\langle a_1,\ldots,a_n\rangle^\frown\langle c_1,\ldots,c_k\rangle = \langle a_1,\ldots,a_n, c_1,\ldots,c_k\rangle$.)

Note 20. Identifying $\langle\rangle$ with $\emptyset$, we get that in case $n = 0$, the collection $[\vec{a} : \mathcal{D}^n \mid \mathcal{M} \models \varphi(\vec{a}^\frown\vec{c})]$ is either 1 ($= \{\emptyset\}$) or 0 ($= \emptyset$), depending on whether $\mathcal{M} \models \varphi(\vec{c})$ or not. Hence invariance with respect to $\emptyset$ is simply *absoluteness*.

Convention. From now on, when we talk about terms, we shall mean legal terms. The same convention applies to formulas and operation symbols.

Theorem 1.

1. *Every term t of $\mathcal{PW}$ is invariant.*
2. *Every operation F of $\mathcal{PW}$ is invariant.*
3. *If $\varphi \succ \{x_1,\ldots,x_n\}$ in $\mathcal{PW}$ then φ is invariant with respect to $\{x_1,\ldots,x_n\}$.*
4. *Every Σ-formula of $\mathcal{PW}$ is persistent.*

Proof. We prove *all* parts *simultaneously*, using induction on the complexity of e, where e is a term or a formula or an operation symbol of $\mathcal{PW}$. Nevertheless, to facilitate reading and understanding, we split the various induction steps into groups that correspond to the four parts of the theorem.

In what follows we assume that for every e we consider, $\langle \mathcal{M}_1, \mathcal{M}_2\rangle$ is an e-pair ($\mathcal{M}_1 = \langle \mathcal{D}_1, I_1\rangle$, $\mathcal{M}_2 = \langle \mathcal{D}_2, I_2\rangle$), and ν is an assignment in $\mathcal{D}_1$.

operations: Suppose that e is the n-ary operation symbol F_φ. For convenience of presentation, assume that $n = 1$. Let $a : \mathcal{D}_1$, and let $b = F_\varphi^{I_1}[a]$. Since $\mathcal{M}_1$ is a model of $\mathcal{PW}_e$, the legality of F_φ and the rule [OpI] imply that $\mathcal{M}_1 \models \varphi(\langle a, b\rangle)$. Since φ should be a Σ-formula, the induction hypothesis for φ implies that $\mathcal{M}_2 \models \varphi(\langle a, b\rangle)$ too. Since $\mathcal{M}_2$ is a model of $\mathcal{PW}_e$, this in turn implies (with the help of [OpI]) that $b = F_\varphi^{I_2}[a]$ as well.

Terms:

- The case where e is a variable is trivial.

- The case $e = \omega$ is immediate from the third item of Definition 5.
- Suppose that e is the term $F_\varphi(s_1, \ldots, s_n)$. Then

$$\nu_{\mathcal{M}_1}[e] = F_\varphi^{I_1}[\nu_{\mathcal{M}_1}[s_1], \ldots, \nu_{\mathcal{M}_1}[s_n]] \quad \nu_{\mathcal{M}_2}[e] = F_\varphi^{I_2}[\nu_{\mathcal{M}_2}[s_1], \ldots, \nu_{\mathcal{M}_2}[s_n]]$$

  Now from the induction hypothesis for φ we get:

$$F_\varphi^{I_1}[\nu_{\mathcal{M}_1}[s_1], \ldots, \nu_{\mathcal{M}_1}[s_n]] = F_\varphi^{I_2}[\nu_{\mathcal{M}_1}[s_1], \ldots, \nu_{\mathcal{M}_1}[s_n]]$$

  while from the induction hypotheses for $s_1, \ldots, s_n$ we get that

$$F_\varphi^{I_2}[\nu_{\mathcal{M}_1}[s_1], \ldots, \nu_{\mathcal{M}_1}[s_n]] = F_\varphi^{I_2}[\nu_{\mathcal{M}_2}[s_1], \ldots, \nu_{\mathcal{M}_2}[s_n]]$$

  It follows that $\nu_{\mathcal{M}_1}[e] = \nu_{\mathcal{M}_2}[e]$.

The invariance relation $\succ$:

- – To show the absoluteness (invariance with respect to $\emptyset$) of $s \in t$, let ν be an assignment in $\mathcal{M}_1$. Assume first that $\mathcal{M}_1, \nu \models s \in t$. Then $\nu_{\mathcal{M}_1}[s] \in^{I_1} \nu_{\mathcal{M}_1}[t]$. It follows by the induction hypotheses for s and t and the fact that $\mathcal{M}_2$ is an $\in$-extension of $\mathcal{M}_1$ that $\nu_{\mathcal{M}_2}[s] \in^{I_2} \nu_{\mathcal{M}_2}[t]$ too. Hence $\mathcal{M}_2, \nu \models s \in t$. The proof of the converse (i.e. that if $\mathcal{M}_2, \nu \models s \in t$ then $\mathcal{M}_1, \nu \models s \in t$) is similar.
  – We leave the simpler proof that $s = t$ is absolute to the reader.
- – That $x \neq x$ is invariant with respect to x follows from the fact that for every $\mathcal{M} = \langle \mathcal{D}, I \rangle$ and ν, $[a : \mathcal{D} \mid \mathcal{M}, \nu\{x := a\} \models x \neq x] = \emptyset$.
  – Let e be the formula $x = t$, where $x \notin Fv(t)$. Then the collection $[a \in \mathcal{D}_1 \mid \mathcal{M}_1, \nu\{x := a\} \models x = t]$ is the singleton of $\nu_{\mathcal{M}_1}[t]$, while $[a \in \mathcal{D}_2 \mid \mathcal{M}_2, \nu\{x := a\} \models x = t]$ is the singleton of $\nu_{\mathcal{M}_2}[t]$. Hence the induction hypothesis for t implies that the two sets are equal.
  – Let e be the formula $x \in t$, where $x \notin Fv(t)$. Then

$$[a \in \mathcal{D}_1 \mid \mathcal{M}_1, \nu\{x := a\} \models x \in t] = [a \in \mathcal{D}_1 \mid a \in^{\mathcal{M}_1} \nu_{\mathcal{M}_1}[t]]$$

$$[a \in \mathcal{D}_2 \mid \mathcal{M}_2, \nu\{x := a\} \models x \in t] = [a \in \mathcal{D}_2 \mid a \in^{\mathcal{M}_1} \nu_{\mathcal{M}_2}[t]]$$

  Since $\nu_{\mathcal{M}_1}[t] = \nu_{\mathcal{M}_2}[t]$ by the induction hypothesis for t, these two equations and the fact that $\mathcal{M}_2$ is an $\in$-extension of $\mathcal{M}_1$ imply:

$$[a \in \mathcal{D}_1 \mid \mathcal{M}_1, \nu\{x := a\} \models x \in t] = [a \in \mathcal{D}_2 \mid \mathcal{M}_2, \nu\{x := a\} \models x \in t]$$

- The case $e = P_\varphi(t_1, \ldots, t_n)$ is immediate from the induction hypothesis for φ, and the φ-instance of [PrI] (which belongs to $\mathcal{PW}_e$, of course).
- The proof that if φ is absolute then so is $\neg\varphi$ is left to the reader.
- Let e be $\varphi \vee \psi$, where $\varphi \succ \{x_1, \ldots, x_n\}$ and $\psi \succ \{x_1, \ldots, x_n\}$. Then the collection $[\vec{a} \in \mathcal{D}_1^n \mid \mathcal{M}_1, \nu\{\vec{x} := \vec{a}\} \models \varphi \vee \psi]$ is the union of $[\vec{a} \in \mathcal{D}_1^n \mid \mathcal{M}_1, \nu\{\vec{x} := \vec{a}\} \models \varphi]$ and $[\vec{a} \in \mathcal{D}_1^n \mid \mathcal{M}_1, \nu\{\vec{x} := \vec{a}\} \models \psi]$. A similar equation holds for $[\vec{a} \in \mathcal{D}_2^n \mid \mathcal{M}_2, \nu\{\vec{x} := \vec{a}\} \models \varphi \vee \psi]$ The equality of these two collections follows therefore from the induction hypotheses for φ and ψ.
- Let e be $\theta = \varphi \wedge \psi$, where $\varphi \succ X$, $\psi \succ Y$, and $Y \cap Fv(\varphi) = \emptyset$. To simplify notation, assume that $Fv(\varphi) = \{x, z\}$, $Fv(\psi) = \{x, y, z\}$, $X = \{x\}$, $Y = \{y\}$. For $c \in \mathcal{D}_1$, let $Z(c) = [a \in \mathcal{D}_2 \mid \mathcal{M}_2 \models \varphi(a, c)\}$. Since $\varphi \succ X$, $Z(c) = [a \in \mathcal{D}_1 \mid \mathcal{M}_1 \models \varphi(a, c)]$ as well (by the induction

hypothesis for φ). Hence $Z(c) \subseteq \mathcal{D}_1$. By the induction hypothesis for ψ, this and the fact that $\psi \succ Y$ imply that if $d \in Z(c)$ then

$$[b \in \mathcal{D}_2 \mid \mathcal{M}_2 \models \psi(d, b, c)] = [b \in \mathcal{D}_1 \mid \mathcal{M}_1 \models \psi(d, b, c)]$$

Denote this set by $W(c, d)$. Now both $[\langle a, b\rangle \in \mathcal{D}_2^2 \mid \mathcal{M}_2 \models \theta(a, b, c)]$ and $[\langle a, b\rangle \in \mathcal{D}_1^2 \mid \mathcal{M}_1 \models \theta(a, b, c)]$ equal $\bigcup_{d\in Z(c)}\{d\} \times W(c, d)$. Hence these two sets are the same (for every $c \in \mathcal{D}_1$).

- We leave the case $e = \exists y\varphi$ where $\varphi \succ X$ to the reader.

Σ-formulas:

- Invariance implies persistence. Hence if φ is absolute then it is persistent by what we have shown above.
- It is well known that persistence of formulas is closed under disjunction, conjunction, and existential quantification, so we leave to the reader the standard proofs of these cases.
- Suppose that θ is $\forall x(\varphi \to \psi)$, where $\varphi \succ \{x\}$ and ψ is a Σ-formula. By the induction hypothesis for φ and ψ, φ is invariant with respect to x, and ψ is persistent. We show that θ is persistent too. So let ν be an assignment in $\mathcal{M}_1$ such that $\mathcal{M}_1, \nu \models \theta$. We show that $\mathcal{M}_2, \nu \models \theta$. So let $a \in \mathcal{D}_2$. We should show that $\mathcal{M}_2, \nu\{x := a\} \models \varphi \to \psi$. This is certainly true in case $\mathcal{M}_2, \nu\{x := a\} \not\models \varphi$. So assume that $\mathcal{M}_2, \nu\{x := a\} \models \varphi$. Then the invariability of φ with respect to x implies that $a \in \mathcal{D}_1$, $\nu\{x := a\}$ is an assignment in $\mathcal{D}_1$, and $\mathcal{M}_1, \nu\{x := a\} \models \varphi$. Since $\mathcal{M}_1, \nu \models \theta$, also $\mathcal{M}_1, \nu\{x := a\} \models \varphi \to \psi$. It follows that $\mathcal{M}_1, \nu\{x := a\} \models \psi$. This implies that $\mathcal{M}_2, \nu\{x := a\} \models \psi$, since ψ is persistent. It follows that $\mathcal{M}_2, \nu\{x := a\} \models \varphi \to \psi$ in this case too. □

Clearly, Theorem 1 directly justifies [Comp]. As for [OpI], its premise ensures that F_φ can be introduced using the usual procedure of extension by definitions. That the resulting operation is invariant follows again from Theorem 1.

4.2.2. [∈-ind]. That [∈-ind] is predicatively valid should be obvious: In any sphere of operation, the empty set is the starting building block of all sets, every other set A is formed from the elements of A, and those elements are logically prior to A. Therefore in any sphere of operation, a property that the empty set has and is inherited by a set from its elements should necessarily hold for all sets. Hence [∈-ind] is valid in any acceptable sphere of operation.

Another way to look at the matter is by asking what properties the basic predicate ∈ should have according to the predicativist view of sets. Since predicatively accepted sets are constructed bottom-up, it should be clear that the 'well-foundedness' of ∈ should be one of those properties. But only the general, open-ended *scheme* of [∈-ind] fully exploits what we really have in mind when we say that ∈ should be well founded. (In contrast, the intuitive absoluteness of the well-foundness of ∈ is not fully captured by any of the standard formal definitions of this notion.[16] For example: whether any subset of a given set A has a minimal element depends on what subsets of A are available, while the existence of a descending ∈-chain of elements of A depends on what sequences of elements of A are available.)

[16]As observed by Weaver in [48], some of those definitions are not predicatively equivalent.

Another point of view that might be taken here is that of a modest platonist who looks at the predicative spheres of operation. As was emphasized in Note 6, we restrict ourselves to predicative subsystems (of some extension by definitions) of **ZF**. In the latter (and even in **ZFC**) it is impossible to *define* a set which have elements that form an infinite descending chain with respect to $\in$. Hence this certainly cannot be done in more constrained predicative systems.

Here it is worth noting that the predicativist conception of sets that underlies $\mathcal{PW}$ and the platonist cumulative one have a lot in common. According to both, all sets are created in stages, where the latter can be taken to be the von Neumann's ordinals. (According to predicativists the "creation" is done by using legitimate definitions, while for platonists this can also be done by methods that go beyond what actual people can use.) Moreover: as we are going to see, both predicativists and platonists associate with every set A which is available to the former *the same* stage (called $rank(A)$) in which it first becomes available (even though the platonist *set* V_α is available to predicativists only for $\alpha \leq \omega$). It seems to me obvious that this view of sets that predicativists and platonists share dictates for both the validity of [$\in$-ind].

Two related notes:

1. Gerhard Jäger has raised two objections against accepting full [$\in$-ind] as a predicatively accepted principle:
   - [$\in$-ind] *is a minimality condition on the universe and thus leads to an inherent vicious circle.*
     I do not share this view. In my opinion, [$\in$-ind] is *not* a minimality condition, but only a *constraint* that our spheres of operation should respect. These are two different things. What is more: I believe (though this belief is not reflected in the present paper) that not all forms of 'circularity' should be rejected from the invariance point of view. Thus, although one might claim that all sorts of recursive definitions are inherently circular, some forms of them, like primitive recursive definitions, are certainly predicatively acceptable.
     In any case, the same argument can be raised against accepting the scheme of induction on the natural numbers. However this general *scheme* was accepted and used Weyl in [51] (see [2, 6]), and in fact it is accepted as predicative by almost everyone interested in the subject.
   - *In the form* $\exists x\varphi[x] \to \exists x(\varphi[x] \wedge (\forall y \in x)(\neg\varphi[y])$ *the scheme of* [$\in$-ind] *claims the existence of a set, without presenting an explicit definition of this set, and without ensuring that its identity is invariant.*
     This issue has already been dealt with above in Note 5. As was emphasized there, every predicative system which is based on classical logic proves pure existential propositions of this sort, and there is nothing impredicative about that, as long as such propositions are understood and used in an appropriate way.
2. Although this is not really an acceptable argument for the predicative validity of [$\in$-ind], it is still interesting to note that Feferman himself included this scheme in the predicative set theories he constructed in [18]. This clearly implies that he saw this scheme as predicatively valid. In response

one may argue that Feferman justified his set theories in [18] only by reducing them to what he did in [15]. However, Feferman did not take his set theories out of the blue; actually he too was led to them by pursuing the invariance criterion. (See also [17].) Hence the fact that he included [$\in$-ind] already in the system $\mathbf{PS}_0$, which is by far the weakest system studied in [18], is telling. Moreover: Feferman explicitly said about the system **PS** that he studied in [16] (Note the name of that paper!) that an ideal predicativist can recognize as correct any particular axiom and rule of inference of that system. Since any instance of [$\in$-ind] can be derived from the axioms and rules of **PS**, this means that according to Feferman a predicativist can recognize as correct any instance of [$\in$-ind].

A final remark: personally, I have no doubt that the answer to the question whether [$\in$-ind] is predicatively valid is positive, and that the arguments given above for this answer should be convincing (and would be accepted as such by Weyl and even Feferman). However, this question is not a mathematical one, and so some people might have different views on this point. Since I see debates on the *exact* meaning of a given word as useless, I simply take such conflicting views as indicating that like 'predicativity' in general, also 'predicativity as invariance' is a family of approaches to the foundations of mathematics. I would be quite happy to call mine 'predicativity given that $\in$ is well-founded'.

4.2.3. [Unif]. Let φ be Σ. For convenience, let $Fv(\varphi) = \{x, y\}$. Suppose that at a certain point of working in $\mathcal{PW}$ we have reached a stage s in which the premise of [Unif] has been derived. Then from that point on it is valid in any sphere of operation that we reach. Let a be an object in some such sphere of operation. If $\forall x \in Z \exists y \varphi$ is false for $Z := a$ in every sphere of operation that includes a which is reached at stage s or later, then certainly the corresponding conclusion of [Unif] is valid in all such spheres. Otherwise there is a stage s' in which $\forall x \in Z \exists y \varphi$ is true for $Z := a$. Since φ is Σ, so is $\forall x \in Z \exists y \varphi$. Hence Theorem 1 implies that $\forall x \in Z \exists y \varphi$ remains true for $Z := a$ at any stage from s' on. This and *the validity of the premise of* [Unif] *in the sphere of operation* $\mathcal{M}$ *of each such stage* imply that the collection of pairs $\langle c, d\rangle$ such that $c \in a$ and $\langle c, d\rangle$ satisfies φ in $\mathcal{M}$ forms ('in V') a function f on a. From the fact that φ is Σ it again follows that as a collection of pairs f remains *invariant*. Therefore f is entitled to be added at some stage to the sphere of operation of that stage, and from that point the corresponding instance of [Unif] remains valid.

Note 21. The only justification given by Feferman in [19] for the version of [Unif] which is used in his auxiliary system $\exists/P$ is that he does not believe there can be a real dispute about it. This argument is not very convincing. I believe that the least Feferman should have done here is to explain why the rule of proof [Unif] cannot simply be replaced by an axiom stating that the premise of the rule implies its conclusion (both without the '$\vdash_{\mathcal{PW}}$' in front). Why could there be a real dispute about this implication, but not about the rule? The answer is clear from the argument given above: For the truth of the conclusion of the rule for some a in some $\mathcal{M}$ such that $\forall x \in Z \exists y \varphi$ holds in $\mathcal{M}$ for $Z := a$, it is crucial that the premise of the rule (i.e. the unicity condition) *remains valid* in any sphere of operation that contains $\mathcal{M}$; its truth just at $\mathcal{M}$ itself is insufficient.

4.3. Comparison with Feferman's Systems. The design of $\mathcal{PW}$ has a lot in common (and was greatly influenced) by the second-order system $P + \exists/P$ for predicative analysis that Feferman has developed in [19]. In particular:

- In practice, neither $\mathcal{PW}$ nor $P + \exists/P$ has a signature which is fixed in advance, and the method of repeatedly extending the language and adding new corresponding axioms is an essential component of the work in both.
- The definitions of both systems involve a *simultaneous* recursive construction of their sets of symbols, terms, formulas, axioms and rules.
- The two special rules of $\mathcal{PW}$, [OpI] and [Unif], which (as is shown below) give this system its strong power (far beyond that of the systems $\mathbf{RST}\omega$ and **PZF** investigated in [5, 7, 10]), respectively generalize to set theory the following two rules from [19]:

  Functional defining rule: This rule allows to infer
  $$\vdash_P \forall v_1 \forall v_0(\varphi(v_1, v_0) \leftrightarrow v_0 = F_\varphi(v_1))$$
  from the premises $\vdash_P \forall v_1 \forall v_2 \forall v_3(\varphi(v_1, v_2) \wedge \varphi(v_1, v_3) \to v_2 = v_3)$ and $\vdash_{\exists/P} \forall v_1 \exists v_0 \varphi(v_1, v_0)$, provided that φ is a formula in which quantification is made only over the lowest type $\mathbf{N}$ (implying that φ is absolute).

  Unification rule: This rule allows to infer
  $$\vdash_{\exists/P} \forall v_1 \in \mathbf{N} \exists v_0 \varphi(v_1, v_0) \to \exists f \forall v_1 \in \mathbf{N} \varphi(v_1, f(v_1))$$
  from $\vdash_P \forall v_1 \in \mathbf{N} \forall v_2 \forall v_3(\varphi(v_1, v_2) \wedge \varphi(v_1, v_3) \to v_2 = v_3)$, provided that φ is a formula in which quantification is made only over $\mathbf{N}$.

  It should be noted that a preliminary version, called [F], of a *combination* of [OpI] and [Unif] appears already in the last section of [15], where predicativity at higher types is discussed. In case φ is a Σ-formula, it allows to infer $\exists f \forall x \forall y(\varphi(x, y) \leftrightarrow y = f(x))$ from the formula $\forall x \exists! y \varphi(x, y)$ in case φ is a Σ-formula. (Note that unlike in the present context of type-free set theory, in [F] the variable f is of type higher than that of x.) The rule [F] was then implicitly split in [19] into the two rules described above.

Nevertheless, there are also important differences between $\mathcal{PW}$ and $P + \exists/P$:

1. $\mathcal{PW}$ is a single system. In contrast, $P + \exists/P$ is a combination of two different ones: P and $\exists/P$. P is taken to be the major one, while the stronger system $\exists/P$ is taken to be only an auxiliary system, which is needed for the precise definition of P. The connection between these two systems involves some choices for which no justification is given in [19]. (See Weaver's criticism, with which I fully agree, in Section 1.6 of [49].) We believe, in fact (though we have not tried to show), that with different (but still predicatively justified) choices, $P + \exists/P$ can be strengthened to a single theory which is as strong as $\mathcal{PW}$.
2. P and $\exists/P$ are based on an extensive system of types, which has types of all finite levels (but actually uses only those of levels 0, 1 and 2). $\mathcal{PW}$, in contrast, is a type-free, single-sorted *set theory.*
3. Unlike $\mathcal{PW}$, which is is purely first-order, P and $\exists/P$ have second-order variables in addition to the first-order ones. In the case of $\exists/P$ it is even allowed (under certain conditions) to quantify on them. Moreover, both systems employ specific second-order rules, like substitution of terms, and

even (again under certain conditions) of formulas, for second-order variables. (No rule of substitution is needed in the case of $\mathcal{PW}$.)

4. The set of natural numbers is taken as given in P. In $\mathcal{PW}$ it is defined.
5. In general, $\mathcal{PW}$ is significantly simpler then $P + \exists/P$, having simpler language and less rules.

Other systems of Feferman which are obviously related to $\mathcal{PW}$ are the pure set theories $\mathbf{PS_i}$ and $\mathbf{PS_iE}$ ($i = 0, 1$) which were introduced in [18]. (See Note 4.) Like $\mathcal{PW}$, they too are intuitively motivated by *the invariance criterion*, and so Σ-formulas and absolute formulas play an important role in their formulation. Accordingly, those systems are less restricted than the second-order systems of [15]. Thus the union operation is allowed in them, although it is taken as impredicative in most of Feferman's papers. On the other hand, unlike $\mathcal{PW}$, none of those systems respects our principle (I) from Section 1.4.2, since their axiomatizations include purely existential principles. An example is given by the Σ-reflection rule, which allows to infer $\exists a\psi^{(a)}$ from ψ in case ψ is a Σ-formula (where $\psi^{(a)}$ is obtained from ψ by restricting each quantifier in it to a). $\mathbf{PS_iE}$ ($i = 0, 1$) violates Principle (II) (from Section 1.4.2) as well, since it includes an axiom that says that every set is enumerable. Another very significant point of difference is due to the fact (shown in [18]) that $\mathbf{PS_1}$ and $\mathbf{PS_1E}$ has a minimum model consisting of all sets constructible before Γ_0. In contrast, the results of Section 7 below imply than any transitive model of $\mathcal{PW}$ should contain Γ_0. It follows that although the predicativity of the systems in [18] is dubious, in some important sense they are all weaker than $\mathcal{PW}$. Another drawback of them is that the choice of some of their axioms and rules seems to be rather ad-hoc.[17]

§5. Some Examples of the Power of $\mathcal{PW}$.

5.1. Abstraction Terms and RST. Let $Fv(\varphi) = \{x, v_1, \ldots, v_n\}$ ($n \geq 0$), and suppose that $\varphi \succ \{x\}$. Using [Comp], this entails:

$$\vdash_{\mathcal{PW}} \forall v_1 \cdots \forall v_n \exists! v_0 \forall x(x \in v_0 \leftrightarrow \varphi)$$

Since the formula $\varphi^\star := \forall x(x \in v_0 \leftrightarrow \varphi)$ is Σ,[18] an application of [OpI] yields:

$$\vdash_{\mathcal{PW}} F_{\varphi^\star}(v_1, \ldots, v_n) = v_0 \leftrightarrow \forall x(x \in v_0 \leftrightarrow \varphi(x, v_1, \ldots, v_n))$$

Henceforth we shall write $\{x \mid \varphi(x, v_1, \ldots, v_n)\}$ instead of $F_{\varphi^\star}(v_1, \ldots, v_n)$ (where $\varphi^\star$ is defined from φ as above), and call this an abstraction term. Obviously, $\vdash_{\mathcal{PW}} \forall x(x \in \{x \mid \varphi\} \leftrightarrow \varphi)$. It follows that $\mathcal{PW}$ is an extension of the theory **RST** from [5]. By the results of that paper, this implies that every rudimentary operation is definable in $\mathcal{PW}$. Here are some examples of terms available in **RST**, and so in $\mathcal{PW}$:

- $\emptyset =_{Df} \{x \mid x \neq x\}$
- $\{t_1, \ldots, t_n\} =_{Df} \{x \mid x = t_1 \vee \ldots \vee x = t_n\}$
- $\langle t, s\rangle =_{Df} \{\{t\}, \{t, s\}\}$
- $\{x \in t \mid \varphi\} =_{Df} \{x \mid x \in t \wedge \varphi\}$, provided $\varphi \succ \emptyset$
- $\{t(x) \mid x \in s\} =_{Df} \{y \mid \exists x. x \in s \wedge y = t\}$

[17] More information about the relations between the axioms and rules of $\mathcal{PW}$ and those of $\mathbf{PS_1}$ is given in Note 22.

[18] More precisely: the logically equivalent formula $\forall x \in v_0 \varphi(x) \wedge \forall x(\varphi(x) \to x \in v_0)$ is Σ.

- $\bigcup t =_{Df} \{x \mid \exists y.y \in t \wedge x \in y\}$
- $s \times t =_{Df} \{x \mid \exists a \exists b.a \in s \wedge b \in t \wedge x = \langle a, b\rangle\}$
- $\iota x\varphi =_{Df} \bigcup\{x \mid \varphi\}$ (provided $\varphi \succ \{x\}$)
- $\lambda x \in s.t =_{Df} \{\langle x, t\rangle \mid x \in s\}$
- $f(x) =_{Df} \iota y.\exists z \exists v(z \in f \wedge v \in z \wedge y \in v \wedge z = \langle x, y\rangle)$
- $Dom(f) = \{x \in \bigcup f \mid \exists y \in \bigcup\bigcup f.\langle x, y\rangle \in f\}$
- $Im(f) = \{x \in \bigcup f \mid \exists y \in \bigcup\bigcup f.\langle y, x\rangle \in f\}$
- $f \restriction s =_{Df} \{\langle x, f(x)\rangle \mid x \in s\}$ (where x is new)

The following are examples of easy related theorems of **RST**:

- $\exists! x\varphi(x) \to \forall x(\varphi(x) \leftrightarrow x = \iota x\varphi(x))$ (if $\varphi \succ \{x\}$)
- $u \in s \to (\lambda x \in s.t)u = t\{u/x\}$ (if u is free for x in t)
- $Fun(f) \to (\langle x, y\rangle \in f \leftrightarrow y = f(x))$, where $Fun(f)$ is the following absolute formula (which says that f is a function):

$$\forall z \in f \exists x \exists y(z = \langle x, y\rangle) \wedge \forall x \forall y_1 \forall y_2(\langle x, y_1\rangle \in f \wedge \langle x, y_2\rangle \in f \to y_1 = y_2)$$

5.2. Explicit Definitions and the Extended [OpI].

5.2.1. *Explicit Definitions.* Explicit definitions of operations are particularly simple case of applying [OpI]:

PROPOSITION 6. *Let t be a term of $\mathcal{PW}$ such that $Fv(t) = \{v_1, \ldots, v_n\}$. Then $\mathcal{PW}$ has an operation F such that $\vdash_{\mathcal{PW}} \forall v_1 \cdots \forall v_n F(v_1, \ldots, v_n) = t$.*

Proof. $F = F_\varphi$, where φ is $v_0 = t$. □

5.2.2. *The Extended* [OpI]. The use of [OpI] can be made more effective with the help of the following proposition:

PROPOSITION 7. *For $1 \le i \le k$, let φ_i and ψ_i be formulas of $\mathcal{PW}$ such that $\varphi_i \in \Sigma$, $Fv(\varphi_i) = \{v_0, v_1, \ldots, v_n\}$, $\psi_i \succ \emptyset$, and $Fv(\psi_i) \subseteq \{v_1, \ldots, v_n\}$. Suppose that for every $1 \le i, j \le k$ such that $i \ne j$ we have that $\vdash_{\mathcal{PW}} \neg(\psi_i \wedge \psi_j)$ and that $\vdash_{\mathcal{PW}} \forall v_1, \ldots, \forall v_n(\psi_i \to \exists! v_0\varphi_i)$. Then $\mathcal{PW}$ has an operation F such that for every $1 \le i \le k$: $\vdash_{\mathcal{PW}} \forall v_1 \cdots \forall v_n(\psi_i \to \varphi_i\{F(v_1, \ldots, v_n)/v_0\})$.*

Proof. $F = F_\theta$, where $\theta = (\psi_1 \wedge \varphi_1) \vee \cdots \vee (\psi_k \wedge \varphi_k) \vee (\neg(\psi_1 \wedge \cdots \psi_k) \wedge v_0 = 0)$.□

We shall call the principle stated in Proposition 7 the *extended* [OpI].

EXAMPLE 1. Let φ be the following Σ-formula (where we write y, x, n instead of v_0, v_1, v_2, respectively).

$$\exists f(Fun(f) \wedge Dom(f) = S(n) \wedge f(0) = x \wedge f(n) = y \wedge \forall i \in n.f(i+1) = \bigcup f(i)$$

An induction on n shows that $\vdash_{\mathcal{PW}} \forall x \forall n(n \in \omega \to \exists! y\varphi)$. Hence the Extended [OpI] implies that $\vdash_{\mathcal{PW}} \forall x \forall n(n \in \omega \to \varphi\{H(x, n)/y\})$ for some operation H of $\mathcal{PW}$. Let $RTC(x) = \cup\{H(n, x) \mid n \in \omega\}$. Then $RTC(x)$ represents in $\mathcal{PW}$ the reflexive-transitive closure of x, and so $TC(x) = RTC(x) - \{x\}$ represents the transitive closure of x. Hence the relation $\in^\star$ (the transitive closure of $\in$) is definable in $\mathcal{PW}$ by $y \in^\star x =_{Df} y \in TC(x)$.

5.3. Transitive Closures and PZF. Let ψ be a formula of $\mathcal{PW}$ such that $\{x,y\} \subseteq Fv(\psi)$, and $\psi \succ \{y\}$. Define φ like in Example 1, replacing the conjunct $\forall i \in n.f(i+1) = \bigcup f(i)$ by $\forall i \in n.f(i+1) = \{y \mid \exists z \in f(i).\psi\{z/x\}\}$. Following exactly the same procedure as in Example 1, we construct in $\mathcal{PW}$ an operation TC_ψ such that $TC_\psi(x,y) := y \in TC_\psi(x)$ is true iff there is a finite ψ-chain that connects y to x. Hence this formula is semantically equivalent to the formula $(TC_{x,y}\psi)(x,y)$ of **PZF**. Moreover: like the latter, $TC_\psi(x,y) \succ \{y\}$. It is also easy to prove in $\mathcal{PW}$ that $TC_\psi(x,y)$ has all the properties that $(TC_{x,y}\psi)(x,y)$ has in **PZF** according to Section 3.3 of [5]. It follows that $\mathcal{PW}$ contains **PZF**.

5.4. The Use of Δ-formulas.

Definition 7. A formula φ is a Δ-*formula* if both φ and $\neg\varphi$ are equivalent in $\mathcal{PW}$ to Σ-formulas.

Proposition 8. *Let φ be a Δ-formula, $Fv(\varphi) = \{v_1,\ldots,v_n\}$. Then $\mathcal{PW}$ has an n-ary predicate R_φ such that $\vdash_{\mathcal{PW}} R_\varphi(v_1,\ldots,v_n) \leftrightarrow \varphi$.*

Proof. Let $\psi := (\varphi \wedge v_0 = 1) \vee (\neg\varphi \wedge v_0 = 0)$. Obviously, $\vdash_{\mathcal{PW}} \forall v_1 \cdots \forall v_n \exists! v_0 \psi$, and ψ is equivalent in $\mathcal{PW}$ to a Δ-formula $\psi^\star$. Therefore by [OpI] we get:

$$\vdash_{\mathcal{PW}} \forall v_1 \cdots \forall v_n(\varphi \wedge F_{\psi^\star}(v_1,\ldots,v_n) = 1) \vee (\neg\varphi \wedge F_{\psi^\star}(v_1,\ldots,v_n) = 0)$$

It follows that $\vdash_{\mathcal{PW}} \varphi \leftrightarrow F_{\psi^\star}(v_1,\ldots,v_n) = 1$. Let $\varphi^\star := F_{\psi^\star}(v_1,\ldots,v_n) = 1$. Then $\varphi^\star$ is absolute. Hence we may apply [PrI] to it. Take R_φ to be $P_{\varphi^\star}$. □

Corollary 1. *$\mathcal{PW}$ has a predicate Tr which defines in it truth in $\mathcal{N}$ of formulas in the first-order language of* **PA**.

Proof. It is well known that truth in $\mathcal{N}$ for that language is definable by a Δ-formula, and it is not difficult to see that the main properties of truth are derivable in $\mathcal{PW}$ for the corresponding predicate that Proposition 8 provides. □

Corollary 2. *Let φ be a Δ-formula, $Fv(\varphi) = \{v_1,\ldots,v_n\}$. Then:*

$$\vdash_{\mathcal{PW}} \forall a \exists Z \forall v_1 \cdots \forall v_n(\langle v_1,\ldots,v_n\rangle \in Z \leftrightarrow \langle v_1,\ldots,v_n\rangle \in a \wedge \varphi)$$

Note 22. The stronger form of [PrI] given by Proposition 8 appears as an admissible procedure in [18]. (See (ii) on p. 18 there). However, its direct use for getting new instances of nonlogical schemas of the theory $\mathbf{PS}_1$ considered there is explicitly *forbidden* — in sharp contrast to its use here. Corollary 2 provides what is called in [16] the 'predicative separation rule'. Both the procedure and the rule are obvious counterparts of [HCR] (also called [Δ^1_1-CR] in, e.g., [24]) — the hyperarithmetic comprehension rule, which is used in [15] as the basis of the progression of theories $\mathbf{HC}_\alpha$, as well as in the single arithmetic second-order theory **IR**, which is introduced there, and is supposed to prove exactly the second order arithmetic formulas which can be proved by predicative means.

5.5. Predicative Set-theoretic Recursion.

Theorem 2. *Let F be an $(n+2)$-ary operation of $\mathcal{PW}$. Then $\mathcal{PW}$ has an $(n+1)$-ary operation G such that*

$$\vdash_{\mathcal{PW}} \forall z_1 \cdots \forall z_n \forall x.G(z_1,\ldots,z_n,x) = F(z_1,\ldots,z_n,x,G\restriction TC(x))$$

(where by $G\restriction y$ we mean the function *$\lambda u \in y.G(z_1,\ldots,z_n,u)$).*

Proof. For simplicity of the presentation, we prove the case $n = 0$. We freely use (here and later) facts about provability in $\mathcal{PW}$ that can easily be seen.

Define a formula φ such that $Fv(\varphi) = \{x, f\}$ by:

$$\varphi := Fun(f) \wedge Dom(f) = RTC(x) \wedge \forall z \in RTC(x).f(z) = F(z, f\restriction TC(z))$$

Next we show that φ has the following properties.

1. φ is absolute.

   This easily follows from Example 1.
2. $\vdash_{\mathcal{PW}} \varphi(x, f) \wedge z \in RTC(x) \rightarrow \varphi(z, f\restriction RTC(z))$

   This follows from $\vdash_{\mathcal{PW}} z \in x \rightarrow RTC(z) \subseteq RTC(x)$ and φ's definition.
3. $\vdash_{\mathcal{PW}} \varphi(x, f_1) \wedge \varphi(x, f_2) \rightarrow f_1 = f_2$

   The proof is by $\in$-induction on x, using the previous item and the fact that $\vdash_{\mathcal{PW}} RTC(x) = \{x\} \cup \bigcup\{RTC(z) \mid z \in x\}$.
4. $\vdash_{\mathcal{PW}} \varphi(x_1, f_1) \wedge \varphi(x_2, f_2) \wedge z \in RTC(x_1) \cap RTC(x_2) \rightarrow f_1(z) = f_2(z)$

   This follows from the previous two items and the definition of φ.
5. $\vdash_{\mathcal{PW}} \forall x \exists f \varphi(x, f)$

   The proof is by an $\in$-induction (in $\mathcal{PW}$) on x. So assume (in $\mathcal{PW}$) that $\forall z \in TC(x) \exists g \varphi(z, g)$. Using an application of [Unif] (which is justified by items 1 and 3), it follows from this assumption that there is a function g such that $Dom(g) = TC(x)$ and $\varphi(z, g(z))$ holds for every $z \in TC(x)$. Let $f^\star = \bigcup\{g(z) \mid z \in TC(x)\}$. From item 4 it follows that $f^\star$ is a function whose domain is $TC(x)$. Let $f = f^\star \cup \{\langle x, F(x, f^\star)\rangle\}$. It is straightforward to verify that $\varphi(x, f)$.

Define $\psi(x, y) = \exists f.\varphi(x, f) \wedge f(x) = y$. From items 3 and 5 it follows that $\vdash_{\mathcal{PW}} \forall x \exists! y \psi(x, y)$, while item 1 implies that ψ is a Σ-formula. Therefore by applying [OpI] we get an operation G such that $\vdash_{\mathcal{PW}} \psi\{G(x)/y\}$. Obviously, G has the required property. □

Example 2. From Theorem 2 it follows that $\mathcal{PW}$ has a unary operation $rank$ such that $\vdash_{\mathcal{PW}} rank(x) = \bigcup\{S(rank(y)) \mid y \in x\}$. Obviously, Platonists assign to any set which predicativists construct the same rank as the predicativists do.

§6. Ordinals in $\mathcal{PW}$.

The notion of an ordinal as a type of some well-order R is totally impredicative. Therefore like in **ZF** (and unlike Feferman or Schütte), we identify here the notion of an ordinal with that of a von Neumann's ordinal.

6.1. Basic Theory of Ordinals.

Definition 8.

- $Tra(x) := \forall y \in x \forall z \in y.z \in x$ (x is transitive).
- $Lin(x) := \forall y \in x \forall z \in x.y \in z \vee y = z \vee z \in y$ (x is linear).
- $On(x) := Tra(x) \wedge Lin(x)$ (x is an ordinal).

As usual, we use small Greek letters to vary over ordinals (writing e.g. $\exists \alpha \varphi$ instead of $\exists x(On(x) \wedge \varphi)$ and $\forall \alpha \varphi$ instead of $\forall x(On(x) \rightarrow \varphi)$). We shall also frequently write $\alpha < \beta$ instead of $\alpha \in \beta$ and $\alpha \leq \beta$ instead of $\alpha \in \beta \vee \alpha = \beta$.

Proposition 9. *The following are provable already in* **VBS***:*

1. $On(\alpha) \wedge \beta \in \alpha \to On(\beta)$
2. $On(\emptyset)$
3. $On(\alpha) \leftrightarrow On(S(\alpha))$
4. $\alpha \leq \beta \leftrightarrow \alpha \subseteq \beta$
5. $\alpha \leq \beta \leftrightarrow \alpha \in S(\beta)$
6. $\alpha = \emptyset \vee \emptyset \in \alpha$.
7. $\beta \in \alpha \to (\alpha = S(\beta) \vee S(\beta) \in \alpha)$.
8. $\beta \in \alpha \vee \alpha = \beta \vee \alpha \in \beta$
9. *Every transitive set of ordinals is an ordinal.*
10. *Every set A of ordinals has a supremum $sup A$.*
11. *If $\alpha \neq 0$ then $\alpha = sup\, \alpha \vee \alpha = S(sup\, \alpha)$ (and not both).* In the first case α is called a limit ordinal, in the second — a successor.
12. *Every non-empty set A of ordinals has a minimal element $min\, A$.*

Proof. All the proofs are standard, and are left for the reader. We just note that none of the proofs requires the full power of the $\in$-induction schema of **VBS**; $\in$-induction limited to absolute formulas suffices. (The latter principle is equivalent to the foundation axiom of **ZF**.) □

From now on, we leave to the readers most of the proofs of claims about provability in $\mathcal{PW}$ in case the proofs in $\mathcal{PW}$ are practically just the standard ones, and their availability in $\mathcal{PW}$ can easily be checked. (This does not include, of course, any proof which makes use of [PrI], [OpI], or [Unif].)

Proposition 10. *The principle of* transfinite induction on ordinals *is available in $\mathcal{PW}$: $\vdash_{\mathcal{PW}} \forall\alpha(\forall\beta < \alpha\varphi(\beta) \to \varphi(\alpha)) \to \forall\alpha\varphi(\alpha)$.*

6.2. Operations on Ordinals.

Theorem 3. *Let F be an $(n+2)$-ary operation in $\mathcal{PW}$. Then $\mathcal{PW}$ has an $(n+1)$-ary operation G such that*

$$\vdash_{\mathcal{PW}} \forall\vec{z}\forall\alpha.G(\vec{z},\alpha) = F(\vec{z},\alpha,\lambda\xi \in \alpha.G(\vec{z},\xi))$$

Proof. Immediate from Theorem 2, since $\vdash_{\mathcal{PW}} On(\alpha) \to TC(\alpha) = \alpha$. □

Once we have the ability to use transfinite recursion on ordinals, we can introduce the standard binary operations of addition $(\alpha + \beta)$, multiplication $(\alpha \times \beta)$ and exponentiation (α^β) in the usual way, and prove their main properties using transfinite induction. One particularly important such property is given in the next definition and proposition.

Definition 9. An ordinal α is *additive principal* if $\xi + \alpha = \alpha$ for every $\xi < \alpha$.

Proposition 11. $\vdash_{\mathcal{PW}}$ *α is additive principal iff $\alpha = \omega^\xi$ for some $\xi < \alpha$.*

In the sequel we shall also need the following theorem about ω-sequences.

Theorem 4. *Suppose that* $(*)$ $\vdash_{\mathcal{PW}} \psi(\alpha,\beta_1) \wedge \psi(\alpha,\beta_2) \to \beta_1 = \beta_2$, *where $\psi(\alpha,\beta)$ is a Σ-formula. Then the following is a theorem of $\mathcal{PW}$:*

$$\forall\alpha\exists\beta\psi(\alpha,\beta) \to \forall\gamma\exists! f(Fun(f) \wedge Dom(f) = \omega \wedge f(0) = \gamma \wedge \forall n \in \omega\psi(f(n), f(n+1)))$$

Proof. The proof of the uniqueness of f is standard, and is left to the reader. For the existence of f, let $\theta(\vec{z}, x, n, h)$ be the following Σ-formula:

$$Fun(h) \wedge Dom(h) = S(n) \wedge h(0) = \gamma \wedge \forall k < n[On(h(S(k))) \wedge \psi(\vec{\xi}, h(k), h(S(k)))]$$

Using $(*)$, an easy induction on k shows that

$$\vdash_{\mathcal{PW}} \theta(\vec{\xi}, \gamma, n, h_1) \wedge \theta(\vec{\xi}, \gamma, n, h_2) \to \forall k < S(n)\ h_1(k) = h_2(k)$$

Hence $(**)$ $\vdash_{\mathcal{PW}} \theta(\vec{\xi}, \gamma, n, h_1) \wedge \theta(\vec{\xi}, \gamma, n, h_2) \to h_1 = h_2$.

Since θ is a Σ-formula, it follows from $(**)$ by [Unif] that

$$\vdash_{\mathcal{PW}} \forall n \in \omega \exists h\ \theta(\vec{\xi}, \gamma, n, h) \to \exists g[Fun(g) \wedge Dom(g) = \omega \wedge \forall n \in \omega \theta(\vec{\xi}, \gamma, n, g(n))]$$

On the other hand, it is straightforward to prove in $\mathcal{PW}$ by induction on n that $\forall \alpha \exists \beta \psi(\vec{\xi}, \alpha, \beta) \to \forall n \in \omega \exists h\ \theta(\vec{\xi}, \gamma, n, h)$. Given γ, the assumption that $\forall \alpha \exists \beta \psi(\vec{\xi}, \alpha, \beta)$ implies therefore in $\mathcal{PW}$ that there exists a function g such that $Dom(g) = \omega$ and $\theta(\vec{\xi}, \gamma, n, g(n))$ holds for every $n \in \omega$. It is easy now to verify (assuming $\forall \alpha \exists \beta \psi(\vec{\xi}, \alpha, \beta)$) that $\lambda n \in w.g(n)(n)$ has the required properties. □

6.3. Ordering Functions.

Definition 10. A function f is an *ordering function* of a set B of ordinals, (in symbols: $Ord(f, B)$) if:

- $Dom(f)$ is an ordinal.
- $Im(f) = B$.
- f is (strictly) monotonic: If $\beta < \gamma$ then $f(\beta) < f(\gamma)$.

Proposition 12. *$\mathcal{PW}$ proves that if f is a strictly monotonic function such that $Dom(f)$ is an ordinal, then $\alpha \leq f(\alpha)$ for every $\alpha \in Dom(f)$.*

Proof. By transfinite induction on α. □

Notation. $f[X] =_{Df} \{f(x) \mid x \in X\}$.

Proposition 13. *$\mathcal{PW}$ proves that every set B of ordinals has a unique ordering function f.*

Proof. The proof of uniqueness is standard, and is left for the reader.

To prove the existence of f, we first introduce the following abbreviations:

$$\varphi(X, g) := \forall x \in X.On(x) \wedge Fun(g)$$

$$\xi_0 = max\{S(sup X), S(sup\{x \in Im(g) \mid On(x)\})\}$$

$$\psi := (\xi \in X \vee \xi > sup X) \wedge \xi \notin Im(g) \wedge \forall \beta < \xi((\beta \in X \vee \beta > sup X) \to \beta \in Im(g))$$

It is easy to see that

$$\vdash_{\mathcal{PW}} \varphi(X, g) \to ((\xi_0 \in X \vee \xi_0 > \sup X) \wedge \xi_0 \notin Im(g))$$

Using items 10 and 12 of Proposition 9, this implies:

$$\vdash_{\mathcal{PW}} \forall X \forall g(\varphi(X, g) \to \exists! \xi \psi(X, g, \xi))$$

Since φ is absolute, and ψ is in Σ (and even absolute), an application of the extended [OpI] (Section 5.2.2) provides therefore an operation F such that

$$\vdash_{\mathcal{PW}} \forall X \forall g(\varphi(X, g) \to \psi(X, g, F(X, g)))$$

It easily follows from that by Theorem 3 that $\mathcal{PW}$ has an operation G such that

$$\vdash_{\mathcal{PW}} \forall X \forall \xi ((\forall x \in X. On(x)) \to G(X,\xi) = F(X, \lambda\tau \in \xi. G(X,\tau))$$

Suppose now that B is a set of ordinals. Then it follows from the last two theorems of $\mathcal{PW}$ shown above that $\forall \xi \psi(B, \lambda\tau \in \xi.G(B,\tau), G(B,\xi))$. It is straightforward to show that for every ordinal ξ, $\lambda\tau \in \xi.G(B,\tau)$ is strictly monotonic. It follows by Proposition 12 that $G(B, S(supB)) \geq S(supB)$, and so $G(B, S(supB) \notin B$. Let β be the minimal ordinal such that $G(B,\beta) \notin B$. It is not difficult to see that $\lambda\tau \in \beta.G(X,\tau)$ is an ordering function of B. □

§7. The Operations ϕ and Γ.

Recall (see [42]) that the binary operation ϕ is defined by $\phi(\alpha,\beta) = \phi_\alpha(\beta)$, where $\phi_0(\beta) = \omega^\beta$, and for $\alpha > 0$, $\phi_\alpha(\beta)$ is the βth ordinal γ such that $\phi_\xi(\gamma) = \gamma$ for every $\xi < \alpha$. The unary operation Γ on ordinals is then defined by letting $\Gamma(\beta)$ be the βth ordinal γ such that $\phi(\gamma, 0) = \gamma$. In particular: Γ_0 is the first fixed-point of the operation $\lambda\alpha.\phi(\alpha, 0)$.

7.1. The Binary Operation ϕ.

Definition 11. The absolute formula $\psi_\phi(\delta,\alpha,f)$ is the conjunction of the following absolute formulas:

1. $On(\alpha) \wedge On(\delta)$
2. $Fun(f) \wedge Dom(f) = \alpha$
3. $\forall \xi \in \alpha(Fun(f(\xi)) \wedge On(Dom(f(\xi))) \wedge \forall \tau \in Dom(f(\xi)) On(f(\xi)(\tau)))$
4. $\alpha > 0 \to (Dom(f(0)) = \delta \wedge \forall \tau < \delta(f(0)(\tau) = \omega^\tau))$
5. $\forall \xi (0 < \xi < \alpha \to Ord(f(\xi), \{\tau \in \delta \,|\, \forall \eta < \xi(\tau \in Dom(f(\eta)) \wedge f(\eta)(\tau) = \tau)\})$

Proposition 14. $\vdash_{\mathcal{PW}} \psi_\phi(\delta,\alpha,f_1) \wedge \psi_\phi(\delta,\alpha,f_2) \to f_1 = f_2$

Proof. Suppose that $\psi_\phi(\delta,\alpha,f_1)$ and $\psi_\phi(\delta,\alpha,f_2)$. Then $Dom(f_1) = Dom(f_2) = \alpha$. We show that $f_1(\xi) = f_2(\xi)$ for every $\xi \in \alpha$. For this we use an induction on ξ. The claim is obvious for $\xi = 0$. So assume that $\xi > 0$ and that $f_1(\eta) = f_2(\eta)$ for every $\eta \in \xi$. By the definition of ψ_ϕ, this implies that both $f_1(\xi)$ and $f_2(\xi)$ are the ordering functions of the same set of ordinals. Hence $f_1(\xi) = f_2(\xi)$. □

Lemma 1. $\vdash_{\mathcal{PW}} \xi < \alpha \wedge \psi_\phi(\delta,\alpha,f) \to \psi_\phi(\delta,\xi,f \restriction \xi)$.

Proof. Immediate from the definition of ψ_ϕ. □

Proposition 15. $\vdash_{\mathcal{PW}} \forall\delta\forall\alpha\exists! f\ \psi_\phi(\delta,\alpha,f)$.

Proof. If $\alpha = 0$ then $f = \emptyset$ is the only f such that $\psi_\phi(\delta,\alpha,f)$. So assume that $\alpha > 0$. Since ψ_ϕ is absolute, it follows by [Unif] from Proposition 14 that

$(\star)$ $\vdash_{\mathcal{PW}} \forall \xi \in \alpha \exists f \psi_\phi(\delta,\xi,f) \to \exists! g(Fun(g) \wedge Dom(g) = \alpha \wedge \forall \xi \in \alpha \psi_\phi(\delta,\xi,g(\xi)))$

Now let δ be an ordinal. By Proposition 14, it suffices to prove in $\mathcal{PW}$ that $\forall\alpha\exists f \psi_\phi(\delta,\alpha,f)$. For this we use an induction on α. The claim is obvious in case $\alpha = 0$ or $\alpha = 1$. So assume that $\alpha > 1$ and that $\forall \xi \in \alpha \exists f \psi_\phi(\delta,\xi,f)$. Using $(\star)$, this implies that there is a function g such that $Dom(g) = \alpha$ and $\forall \xi \in \alpha \psi_\phi(\delta,\xi,g(\xi))$. Now we have two cases to consider:

- $\alpha = S(\xi_0)$ for some ξ_0. then we let $f = g(\xi_0) \cup \{\langle \xi_0, o\rangle\}$, where o is the ordering function of $\{\tau \in \delta \,|\, \forall \eta < \xi_0(\tau \in Dom(f(\eta)) \wedge f(\eta)(\tau) = \tau)\}$).
- α is a limit ordinal. Then for $\xi < \alpha$ let $f(\xi) = g(S(\xi))(\xi)$.

Using Lemma 1, it is straightforward to verify in both cases that $\psi_\phi(\delta,\alpha,f)$. $\square$

Proposition 15 allows us to apply [OpI] and introduce in $\mathcal{PW}$ a new operation symbol F_ϕ together with:

AXIOM$_{F_\phi}$: $f = F_\phi(\delta,\alpha) \leftrightarrow \psi_\phi(\delta,\alpha,f)$

In other words: for any two ordinals α and δ, $F_\phi(\delta,\alpha)$ is the unique function that has the properties 2.–5. from Definition 11.

Some other important properties of F_ϕ are given in the next proposition.

Proposition 16. *The following are theorems of $\mathcal{PW}$:*

1. *If $\alpha > 0$ and $\beta < \delta$ then $F_\phi(\delta,\alpha)(0)(\beta) = \omega^\beta$.*
2. *If $\alpha_1 < \alpha_2$ then $F_\phi(\delta,\alpha_1) = F_\phi(\delta,\alpha_2) \restriction \alpha_1$.*
3. *If $\delta_1 < \delta_2$ then for every $\xi < \alpha$: $Dom(F_\phi(\delta_1,\alpha)(\xi))$ is an initial segment of $Dom(F_\phi(\delta_2,\alpha)(\xi))$, and $F_\phi(\delta_1,\alpha)(\xi) = F_\phi(\delta_2,\alpha)(\xi) \restriction Dom(F_\phi(\delta_1,\alpha)(\xi))$.*
4. *If $\xi_1 < \xi_2$ Then for every $\beta_2 \in Dom(F_\phi(\delta,\alpha)(\xi_2))$ there exists $\beta_1 \geq \beta_2$ such that $\beta_1 \in Dom(F_\phi(\delta,\alpha)(\xi_1))$, and $F_\phi(\delta,\alpha)(\xi_2)(\beta_2) = F_\phi(\delta,\alpha)(\xi_1)(\beta_1)$.*
5. *If $\xi < \alpha$ and $\beta \in Dom(F_\phi(\delta,\alpha)(\xi))$ then there exists $\eta < \delta$ such that $F_\phi(\delta,\alpha)(\xi)(\beta) = \omega^\eta$. In particular: $F_\phi(\delta,\alpha)(\xi)(\beta) > 0$.*
6. *If $\xi < \alpha$ and $\beta \in Dom(F_\phi(\delta,\alpha)(\xi))$ then $F_\phi(\delta,\alpha)(\xi)(\beta) \geq max(\xi,\beta)$.*

Proof.

1. Immediate from AXIOM$_{F_\phi}$ and Definition 11.
2. Immediate from AXIOM$_{F_\phi}$ and Lemma 1.
3. Induction on ξ. The claim is obvious for $\xi = 0$ by the first item. Now suppose that $\xi > 0$, and that for every $\eta < \xi$ it holds that $Dom(F_\phi(\delta_1,\alpha)(\eta))$ is an initial segment of $Dom(F_\phi(\delta_2,\alpha)(\eta))$, and $F_\phi(\delta_1,\alpha)(\eta) = F_\phi(\delta_2,\alpha)(\eta) \restriction Dom(F_\phi(\delta_1,\alpha)(\eta))$. For $i = 1,2$ let
   $$A_i = \{\tau \in \delta_i \mid \forall \eta < \xi(\tau \in Dom(F_\phi(\delta_i,\alpha)(\eta)) \land F_\phi(\delta_i,\alpha)(\eta)(\tau) = \tau)\}$$
   Then $A_1 = A_2 \cap \delta_1$ by our induction hypothesis. Since $F_\phi(\delta_i,\alpha)(\xi)$ is the ordering function of A_i $(i = 1,2)$, this implies the claim for ξ.
4. Immediate from the fact that $F_\phi(\delta,\alpha)(\xi_2)$ is the ordering function of a certain subset of $Im(F_\phi(\delta,\alpha)(\xi_1))$.
5. Immediate from the previous item and the fact that $F_\phi(\delta,\alpha)(0)(\eta) = \omega^\eta$.
6. That $F_\phi(\delta,\alpha)(\xi)(\beta) \geq \beta$ follows from Proposition 12 in case $\xi > 0$, and from the fact that $\omega^\beta > \beta$ in case $\xi = 0$.
   That $F_\phi(\delta,\alpha)(\xi)(\beta) \geq \xi$ is shown by induction on ξ. It is certainly true if $\xi = 0$. Suppose that $\xi > 0$. Let $\eta < \xi$. Then item 3 of this proposition and the induction hypothesis for η together imply that $\eta \leq F_\phi(\delta,\alpha)(\xi)(\beta)$. Since this is true for every $\eta < \xi$, it follows that $\xi \leq F_\phi(\delta,\alpha)(\xi)(\beta)$. $\square$

Next we introduce a ternary Σ-formula that expresses (as we show below) the graph of the binary operation ϕ on ordinals.

Definition 12. $\varphi_\phi(\alpha,\beta,\gamma)$ is the following Σ-formula:

$$\exists\delta(\beta \in Dom(F_\phi(\delta,S(\alpha))(\alpha)) \land F_\phi(\delta,S(\alpha))(\alpha)(\beta) = \gamma)$$

Proposition 17. $\vdash_{\mathcal{PW}} \varphi_\phi(\alpha,\beta,\gamma) \to (\alpha \leq \gamma \land \beta \leq \gamma)$.

Proof. This follows from the last item of Proposition 16. □

PROPOSITION 18. $\vdash_{\mathcal{PW}} \varphi_\phi(\alpha,\beta,\gamma_1) \wedge \varphi_\phi(\alpha,\beta,\gamma_2) \to \gamma_1 = \gamma_2$.

Proof. Suppose that $\varphi_\phi(\alpha,\beta,\gamma_1)$ and $\varphi_\phi(\alpha,\beta,\gamma_2)$. Then there exist δ_1 and δ_2 such that $\beta \in Dom(F_\phi(\delta_1, S(\alpha))(\alpha)) \wedge F_\phi(\delta_1, S(\alpha))(\alpha)(\beta) = \gamma_1$, and also $\beta \in Dom(F_\phi(\delta_2, S(\alpha))(\alpha)) \wedge F_\phi(\delta_2, S(\alpha))(\alpha)(\beta) = \gamma_2$. Without loss in generality, we may assume that $\delta_1 \leq \delta_2$. Then it follows from item 3 of Proposition 16 that $F_\phi(\delta_2, S(\alpha))(\alpha)(\beta) = \gamma_1$. Hence $\gamma_1 = \gamma_2$. □

To follow the proof of the next theorem, it would be helpful to remember that the intend meaning of the formula '$\varphi_\phi(\alpha,\beta,\gamma)$' of $\mathcal{PW}$ is '$\phi(\alpha,\beta) = \gamma$' (which at present is a formula only in the metalanguage of $\mathcal{PW}$).

THEOREM 5. $\vdash_{\mathcal{PW}} \forall\alpha\forall\beta\exists!\gamma\varphi_\phi(\alpha,\beta,\gamma)$

Proof. The uniqueness part follows from Proposition 18. We prove the existence part by using an $\in$-induction on α in $\mathcal{PW}$ to simultaneously show:

(a) $\forall\beta\exists\gamma\varphi_\phi(\alpha,\beta,\gamma)$
(b) $\forall\beta\forall\gamma([Fun(f) \wedge Fun(g) \wedge Dom(f) = Dom(g) = \omega \wedge \forall n \in \omega\varphi_\phi(\alpha, f(n), g(n)) \wedge \beta = sup\{f(n) \mid n \in \omega\} \wedge \gamma = sup\{g(n) \mid n \in \omega\}] \to \varphi_\phi(\alpha,\beta,\gamma))$

The case $\alpha = 0$ is easy, since $\vdash_{\mathcal{PW}} \varphi_\phi(0,\beta,\gamma) \leftrightarrow \omega^\beta = \gamma$.

Now fix some $\alpha > 0$, and assume that (a) and (b) are true for every $\xi \in \alpha$. In particular, we have that $\forall\xi \in \alpha\forall\beta\exists\gamma\varphi_\phi(\xi,\beta,\gamma)$. This implies:

$$(1) \quad \forall\beta\forall\xi \in \alpha\exists\gamma\varphi_\phi(\xi,\beta,\gamma)$$

Using Proposition 18, an application of [Unif] yields:

$$(2) \; \vdash_{\mathcal{PW}} \forall\xi \in \alpha\exists\gamma\varphi_\phi(\xi,\beta,\gamma)) \to \exists! f(Fun(f) \wedge Dom(f) = \alpha \wedge \forall\xi \in \alpha\varphi_\phi(\xi,\beta,f(\xi)))$$

From (1) and (2) we get:

$$(3) \quad \forall\beta\exists! f(Fun(f) \wedge Dom(f) = \alpha \wedge \forall\xi \in \alpha\varphi_\phi(\xi,\beta,f(\xi)))$$

Let $\theta(\beta,\tau)$ be the following formula:

$$\theta := \exists f(Dom(f) = \alpha \wedge \forall\eta \in \alpha\varphi_\phi(\eta,\beta,f(\eta)) \wedge \tau = sup\{f(\eta) \mid \eta \in \alpha\})$$

Then θ is in Σ, and (3) implies that the following holds for the given α:

$$(4) \quad \forall\beta\exists!\tau\theta$$

Next we show that for every ordinal β there exists a bigger ordinal γ such that $\varphi_\phi(\xi,\gamma,\gamma)$ holds for every $\xi \in \alpha$. So fix an ordinal β. From (4) it follows by Theorem 4 that there exists an ω-sequence g such that $g(0) = \beta + 1$, and $\theta\{g(n), g(n+1)\}$ for every $n \in \omega$. Let $\gamma = sup\{g(n) \mid n \in \omega\}$. Then $g(n) \leq \gamma$ for every $n \in \omega$. Hence $\beta < \gamma$ (since $\beta < g(0)$). We show that $\varphi_\phi(\xi,\gamma,\gamma)$ for $\xi < \alpha$. So fix $\xi \in \alpha$. (1) implies that $\forall n \in \omega\exists\gamma'\varphi_\phi(\xi, g(n), \gamma')$. Using Proposition 18 and an application of [Unif], this yields a function d such that $\varphi_\phi(\xi, g(n), d(n))$ for every $n \in \omega$. Let $D = \sup\{d(n) \mid n \in \omega\}$. Part (b) of the induction hypothesis entails that $\varphi_\phi(\xi,\gamma,D)$. Hence $\gamma \leq D$ by Proposition 17. On the other hand, the fact that $\theta\{g(n), g(n+1)\}$ implies that there is a function f such that $Dom(f) = \alpha$ and for every $\eta < \alpha$, $\varphi_\phi(\eta, g(n), f(\eta)) \wedge g(n+1) \geq f(\eta)$. In particular: (i) $\varphi_\phi(\xi, g(n), f(\xi))$ and (ii) $g(n+1) \geq f(\xi)$. By Proposition 18,

(i) entails that $f(\xi) = d(n)$. Hence (ii) implies that $d(n) \le g(n+1)$. This is true for every $n \in \omega$. Hence $D \le \gamma$. It follows that $D = \gamma$, implying that $\varphi_\phi(\xi, \gamma, \gamma)$.

Now we prove (a) for α, i.e. that for every β there is γ such that $\varphi_\phi(\alpha, \beta, \gamma)$. We do this by induction on β. So suppose that $\forall \eta < \beta \exists \gamma \varphi_\phi(\alpha, \eta, \gamma)$. Using Proposition 18 and [Unif], this provides a function h with domain β such that $\forall \eta < \beta \varphi_\phi(\alpha, \eta, h(\eta))$. Let γ_0 be the least ordinal γ such that $\varphi_\phi(\xi, \gamma, \gamma)$ for every $\xi \in \alpha$, and $\gamma > sup\{h(\eta) \mid \eta < \beta\}$. ($\gamma_0$ exists by the claim we have just proved.) By definition, the first property of γ_0 means that:

$$(6)\ \ \forall \xi < \alpha \exists \delta (\gamma_0 \in Dom(F_\phi(\delta, S(\xi))(\xi)) \wedge F_\phi(\delta, S(\xi))(\xi)(\gamma_0) = \gamma_0)$$

By the second item of Proposition 16, this implies:

$$(7)\ \ \forall \xi < \alpha \exists \delta (\gamma_0 \in Dom(F_\phi(\delta, S(\alpha))(\xi)) \wedge F_\phi(\delta, S(\alpha))(\xi)(\gamma_0) = \gamma_0)$$

Let $A(\xi, \delta) := \gamma_0 \in Dom(F_\phi(\delta, S(\alpha))(\xi)) \wedge F_\phi(\delta, S(\alpha))(\xi)(\gamma_0) = \gamma_0$, and let $B(\xi, \delta) := A \wedge \forall \delta' < \delta \neg A\{\delta'/\delta\}$. (7) implies that $\forall \xi < \alpha \exists \delta\ B(\xi, \delta)$. Obviously, $\vdash_{\mathcal{PW}} B(\xi, \delta_1) \wedge B(\xi, \delta_2) \to \delta_1 = \delta_2$. Since B is absolute (because A is), we can use [Unif] in order to infer from the last two facts that there is a function b such that $\forall \xi < \alpha B(\xi, b(\xi))$. This, in turn, implies that $\forall \xi < \alpha A(\xi, b(\xi))$. Let $\delta_0 = sup\{b(\xi) \mid \xi < \alpha\}$. By the third item of Proposition 16, we get from the last claim and the definition of A that:

$$(8)\ \ \forall \xi < \alpha (\gamma_0 \in Dom(F_\phi(\delta_0, S(\alpha))(\xi)) \wedge F_\phi(\delta_0, S(\alpha))(\xi)(\gamma_0) = \gamma_0)$$

Now, by AXIOM$_{F_\phi}$ and Definition 11(5), we have that $Im(F_\phi(\delta_0, S(\alpha))(\alpha))$ is $\{\tau \in \delta_0 \mid \forall \eta < \alpha (\tau \in Dom(F_\phi(\delta_0, S(\alpha))(\eta)) \wedge F_\phi(\delta_0, S(\alpha))(\eta)(\tau) = \tau)\}$. Hence (8) implies that $\gamma_0 \in Im(F_\phi(\delta_0, S(\alpha))(\alpha))$. It follows that there exists β' in $Dom(F_\phi(\delta_0, S(\alpha))(\alpha))$ such that $F_\phi(\delta, S(\alpha))(\alpha)(\beta') = \gamma_0$. But for every $\eta < \beta$, $\gamma_0 > h(\eta)$. Since $\forall \eta < \beta \varphi_\phi(\alpha, \eta, h(\eta))$, this and Proposition 18 imply that $\forall \tau \forall \eta < \beta (\varphi_\phi(\alpha, \eta, \tau) \to \tau < \gamma_0)$. By definition of φ_ϕ, this means that $\forall \delta \forall \eta < \beta F_\phi(\delta_0, S(\alpha))(\alpha)(\eta) < \gamma_0$. In particular: $F_\phi(\delta_0, S(\alpha))(\alpha)(\eta) < \gamma_0$ for every $\eta < \beta$. Therefore $\beta' \ge \beta$. It follows that $\beta \in Dom(F_\phi(\delta_0, S(\alpha))(\alpha))$ as well (by the definition of an ordering function). Let $\gamma = F_\phi(\delta_0, S(\alpha))(\alpha)(\beta)$. Then $\varphi_\phi(\alpha, \beta, \gamma)$.

Finally, we prove that α satisfies (b) too. So let f and g be functions whose domain is ω, and suppose that $\forall n \in \omega \varphi_\phi(\alpha, f(n), g(n))$, $\beta = sup\{f(n) \mid n \in \omega\}$ and $\gamma = sup\{g(n) \mid n \in \omega\}$. Then $\forall n \in \omega \exists \delta (F_\phi(\delta, S(\alpha))(\alpha)(f(n)) = g(n))$. With the help of the method used above to infer (8) from (7), we can infer from this that there exists δ such that $\forall n \in \omega (F_\phi(\delta, S(\alpha))(\alpha)(f(n)) = g(n))$. Since $\psi_\phi(\delta, S(\alpha), F_\phi(\delta, S(\alpha)))$ by AXIOM$_{F_\phi}$, and $0 < \alpha < S(\alpha)$, this implies that $Ord(F_\phi(\delta, S(\alpha))(\alpha), \{\tau \in \delta \mid \forall \xi < \alpha F_\phi(\delta, S(\alpha))(\xi)(\tau) = \tau\}$. It follows that for every $n \in \omega$ we have that $\forall \xi < \alpha F_\phi(\delta, S(\alpha))(\xi)(g(n)) = g(n)$. By item 2 of Proposition 16 this implies that $\forall \xi < \alpha F_\phi(\delta, S(\xi))(\xi)(g(n)) = g(n)$, and so also $\forall \xi < \alpha \varphi_\phi(\xi, g(n), g(n))$ (by the definition of φ_ϕ). Hence the induction hypothesis for $\xi < \alpha$ entails that $\forall \xi < \alpha \varphi_\phi(\xi, \gamma, \gamma)$. It follows that $\forall \xi < \alpha \exists \delta' (F_\phi(\delta', S(\xi))(\xi)(\gamma) = \gamma)$. Applying again the method used to infer (8) from (7), we get from that an ordinal δ such that $\forall \xi < \alpha (F_\phi(\delta, S(\xi))(\xi)(\gamma) = \gamma)$. Therefore $\forall \xi < \alpha (F_\phi(\delta, S(\alpha))(\xi)(\gamma) = \gamma)$ (by the second item of Proposition 16). It follows that $\gamma \in Im(c)$, where $c = F_\phi(\delta, S(\alpha))(\alpha)$. Hence $\gamma = c(\beta')$ for some β'. Since c is monotonic, and $\gamma \ge g(n)$ for every $n \in \omega$, $\beta' \ge f(n)$ for every $n \in \omega$, and so $\beta' \geq \beta$. It follows that β is in the domain of c, and so $\gamma = c(\beta') \geq c(\beta)$. On the other hand, $c(\beta) \geq c(f(n)) = g(n)$ for every $n \in \omega$, and so $c(\beta) \geq \gamma$. Hence $c(\beta) = \gamma$, and so $\varphi_\phi(\alpha, \beta, \gamma)$. □

Theorem 5 allows us to apply [OpI] and introduce in $\mathcal{PW}$ a new operation symbol ϕ together with the following axiom:

AXIOM$_\phi$: $\gamma = \phi(\alpha, \beta) \leftrightarrow \phi(\alpha, \beta, \gamma)$

The standard characteristic properties of ϕ are given in the next proposition.

Proposition 19. *The following are theorems of $\mathcal{PW}$:*

1. $\beta_1 < \beta_2 \rightarrow \phi(\alpha, \beta_1) < \phi(\alpha, \beta_2)$
2. $\phi(0, \beta) = \omega^\beta$.
3. $\alpha > 0 \rightarrow \forall\gamma(\exists\beta(\gamma = \phi(\alpha, \beta)) \leftrightarrow \forall\xi < \alpha(\phi(\xi, \gamma) = \gamma))$

Proof.

1. Let $\phi(\alpha, \beta_1) = \gamma_1$, $\phi(\alpha, \beta_2) = \gamma_2$. Then $\varphi_\phi(\alpha, \beta_1, \gamma_1)$ and $\varphi_\phi(\alpha, \beta_2, \gamma_2)$. Hence there exist δ_1 and δ_2 such that $F_\phi(\delta_i, S(\alpha))(\alpha)(\beta_i) = \gamma_i$ for $i = 1, 2$. Without loss in generality, we may assume that $\delta_1 \leq \delta_2$. Then it follows from item 3 of Proposition 16 that $F_\phi(\delta_2, S(\alpha))(\alpha)(\beta_1) = \gamma_1$. Since $F_\phi(\delta_2, S(\alpha))(\alpha)$ is an ordering function, it follows that $\gamma_1 < \gamma_2$.
2. Immediate from the first item of Proposition 16 (and Theorem 5).
3. Suppose first that $\gamma = \phi(\alpha, \beta)$. Then $\varphi_\phi(\alpha, \beta, \gamma)$, and so there exists δ such that $F_\phi(\delta, s(\alpha))(\alpha)(\beta) = \gamma$. Therefore it follows by AXIOM$_{F_\phi}$ and Definition 11 that $F_\phi(\delta, S(\alpha))(\xi)(\gamma) = \gamma$ for every $\xi < \alpha$. Hence $F_\phi(\delta, S(\xi))(\xi)(\gamma) = \gamma$ for every $\xi < \alpha$, by the second item of Proposition 16. It follows that $\varphi_\phi(\xi, \gamma, \gamma)$, and so $\phi(\xi, \gamma) = \gamma$, for every $\xi < \alpha$.

   For the converse, let $\phi(\xi, \gamma) = \gamma$ (i.e. $\varphi_\phi(\xi, \gamma, \gamma)$) for every $\xi < \alpha$. We can show that this implies that there exists δ such that $\gamma \in Im(F_\phi(\delta, s(\alpha))(\alpha)))$ exactly as we show the same implication for γ_0 at the proof of Theorem 5. This means that there is β such that $F_\phi(\delta, s(\alpha))(\alpha))(\beta) = \gamma$. Hence, by definition, $\varphi_\phi(\alpha, \beta, \gamma)$, and so there exists β such that $\phi(\alpha, \beta) = \gamma$. □

Corollary 3. *The following are theorems of $\mathcal{PW}$:*

1. $\phi(\xi, \phi(\alpha, \beta)) = \phi(\alpha, \beta)$ *for every* $\xi < \alpha$.
2. $\alpha > 0 \rightarrow \forall\gamma < \phi(\alpha, \beta)((\forall\xi < \alpha\phi(\xi, \gamma) = \gamma) \rightarrow \exists\beta' < \beta\phi(\alpha, \beta') = \gamma)$.
3. *Let* $\alpha > 0$*, and suppose that the following three conditions are satisfied:*
   (a) $\phi(\xi, \gamma) = \gamma$ *for every* $\xi < \alpha$*;*
   (b) $\phi(\alpha, \beta') < \gamma$ *for every* $\beta' < \beta$*;*
   (c) $\forall\gamma' < \gamma[(\forall\xi < \alpha\phi(\xi, \gamma') = \gamma') \rightarrow \exists\beta' < \beta\phi(\alpha, \beta') = \gamma']$
   Then $\phi(\alpha, \beta) = \gamma$.

Proof. Easily follows from Proposition 19. □

Proposition 20. $\vdash_{\mathcal{PW}} \alpha \leq \phi(\alpha, \beta) \land \beta \leq \phi(\alpha, \beta)$.

Proof. Immediate from Proposition 17.

Corollary 4. *$\mathcal{PW}$ proves that for every ordinal $\alpha > 0$ and for every ordinal γ such that $\forall\xi < \alpha\ \phi(\xi, \gamma) = \gamma$ there exists $\beta \leq \gamma$ such that $\phi(\alpha, \beta) = \gamma$.*

Proof. $\gamma < \phi(\alpha, S(\gamma))$ by Proposition 20. Hence the claim follows from item 2 of Corollary 3. □

Once we have proved Proposition 19, all of ϕ's main properties (as given. e.g., at Chapter V of [42]) can predicatively be derived using the standard proofs. As an example, we present here the full proof of the following well-know result.

Proposition 21. $\vdash_{\mathcal{PW}} \forall A(A \neq \emptyset \to \forall\alpha\, \phi(\alpha, sup\ A) = sup\{\phi(\alpha,\beta) \mid \beta \in A\})$

Proof. The claim is obviously true for $\alpha = \emptyset$. So assume that $\alpha > 0$. Using $\in$-induction on α, we show that $\phi(\alpha, \beta) = \gamma$, where $\gamma = sup\{\phi(\alpha,\tau) \mid \tau \in A\}$ and $\beta = sup\ A$. We do that by showing that the three conditions given in item 3 of Corollary 3 are satisfied.

- Let $\xi < \alpha$. Then $\phi(\xi, \gamma) = sup\{\phi(\xi, \phi(\alpha, \tau)) \mid \tau \in A\}$ by the induction hypothesis for ξ. Hence $\phi(\xi, \gamma) = \gamma$ by item 1 of Corollary 3.
- Let $\beta' < \beta$. Then there exists $\tau \in A$ such that $\beta' < \tau$. It follows by item 1 of Proposition 19 that $\phi(\alpha, \beta') < \phi(\alpha, \tau) \leq \gamma$.
- Let $\gamma' < \gamma$, and suppose that $\phi(\xi, \gamma') = \gamma'$ for every $\xi < \alpha$. We show that there exists $\beta' < \beta$ such that $\phi(\alpha, \beta') = \gamma'$. Since $\gamma' < \gamma$, there exists $\beta^\star \in A$ (and so $\beta^\star \leq \beta$) such that $\gamma' < \phi(\alpha, \beta^\star)$. Hence item 2 of Corollary 3 implies that there exists $\beta' < \beta^\star \leq \beta$ such that $\phi(\alpha, \beta') = \gamma'$. □

7.2. Γ_0 and the Operation Γ. Once the operation ϕ becomes available, it is almost a routine matter to introduce in $\mathcal{PW}$ the operation Γ as well.

Proposition 22. $\vdash_{\mathcal{PW}} \alpha > 0 \to (\phi(\alpha, 0) = \alpha \leftrightarrow \forall\xi < \alpha\phi(\xi, \alpha) = \alpha)$.

Proof. The implication from left to right follows from item 1 of Corollary 3. Its converse follows from the assumption that $0 < \alpha$. □

Theorem 6. $\vdash_{\mathcal{PW}} \forall\alpha\exists\beta(\beta > \alpha \wedge \phi(\beta, 0) = \beta)$.

Proof. Given an ordinal α, we use Theorem 4 to define a function f on ω by letting $f(0) = S(\alpha)$, and $f(n + 1) = \phi(f(n), 0)$ for every $n \in \omega$. Then $f(n) \leq f(n + 1)$ by Proposition 20. Let $\beta = sup\{f(n) \mid n \in \omega\}$. Obviously, $\beta > \alpha$. We show that also $\phi(\beta, 0) = \beta$. By Proposition 22 it suffices to show that $\phi(\gamma, \beta) = \beta$ for every $\gamma < \beta$. So let $\gamma < \beta$. Then there exists $k \in \omega$ such that $\gamma < f(n)$ for every $n > k$. Hence $\phi(\gamma, f(n)) = f(n)$ for every $n > k$ by item 1 of Corollary 3. It follows by Proposition 21 that $\phi(\gamma, \beta) = \beta$. □

Theorem 7. *$\mathcal{PW}$ has an operation Γ such that $\mathcal{PW}$ proves the following:*

1. $\forall\alpha\forall\beta(\beta > \alpha \to \Gamma(\beta) > \Gamma(\alpha))$.
2. $\forall\alpha\ \phi(\Gamma(\alpha), 0) = \Gamma(\alpha)$.
3. $\forall\alpha\forall\gamma(\Gamma(\alpha) > \gamma \wedge \phi(\gamma, 0) = \gamma \to \exists\beta < \alpha\ \gamma = \Gamma(\beta))$.

Proof. Using the basic properties of ordinals, Theorem 6 implies that

$$\vdash_{\mathcal{PW}} \forall\alpha\exists!\beta(\beta > \alpha \wedge \phi(\beta, 0) = \beta \wedge \forall\gamma(\gamma < \beta \wedge \phi(\gamma, 0) = \gamma \to \gamma \leq \alpha))$$

Therefore it follows from [OpI] that $\mathcal{PW}$ has an operation G such that

$$\vdash_{\mathcal{PW}} G(\alpha) > \alpha \wedge \phi(G(\alpha), 0) = G(\alpha) \wedge \forall\gamma(\gamma < G(\alpha) \wedge \phi(\gamma, 0) = \gamma \to \gamma \leq \alpha)$$

We can now use recursion to introduce an operation Γ as follows:

$$\Gamma(\alpha) = \begin{cases} G(0) & \alpha = 0 \\ G(\Gamma(\beta)) & \alpha = S(\beta) \\ sup\{\Gamma(\beta) \mid \beta < \alpha\} & \alpha \text{ is a limit ordinal} \end{cases}$$

It is obvious that Γ has property 1. To show that Γ has property 2, we use an $\in$-induction on α. This is obviously true in case $\alpha = 0$ or α is a successor ordinal. So assume that α is a limit ordinal, By Proposition 22 and Proposition 20, it suffices to show that $\phi(\tau, \Gamma(\alpha)) \leq \Gamma(\alpha)$ for all $\tau < \Gamma(\alpha)$. So let $\tau < \Gamma(\alpha)$. Since α is a limit ordinal, the definition of $\Gamma(\alpha)$ implies that there is $\beta < \alpha$ such that $\tau < \Gamma(\beta)$. By induction hypothesis and Corollary 3, for every such β it holds that $\phi(\tau, \Gamma(\beta)) = \Gamma(\beta)$ and so $\phi(\tau, \Gamma(\beta)) < \Gamma(\alpha)$. By Proposition 21, this implies that $\phi(\tau, \Gamma(\alpha)) \leq \Gamma(\alpha)$.

Finally, we prove that Γ has property 3 by an $\in$-induction on α. So suppose that $\Gamma(\alpha) > \gamma$ and $\phi(\gamma, 0) = \gamma$. There are three cases to consider:

- The case $\alpha = 0$ is trivial, since by definition of G, there is no γ such that $\Gamma(0) > \gamma$ and $\phi(\gamma, 0) = \gamma$.
- Suppose that $\alpha = S(\xi)$ for some ξ. Then $\Gamma(\alpha) = G(\Gamma(\xi))$. It follows that $\gamma < G(\Gamma(\xi))$. By the properties of G, this implies that $\gamma \leq \Gamma(\xi)$. If $\gamma = \Gamma(\xi)$ we are done. Otherwise we apply the induction hypothesis to ξ, and get $\beta < \xi < \alpha$ such that $\gamma = \Gamma(\beta)$.
- Suppose that α is a limit ordinal. Then by definition of Γ, $\gamma < \Gamma(\alpha)$ implies that $\gamma < \Gamma(\xi)$ for some $\xi < \alpha$. By applying the induction hypothesis to ξ we get $\beta < \xi < \alpha$ such that $\gamma = \Gamma(\beta)$. □

COROLLARY 5. $\vdash_{\mathcal{PW}} \phi(\gamma, 0) = \gamma \to \exists\beta(\gamma = \Gamma(\beta))$

Proof. Together with the first item of Theorem 7, Proposition 12 implies that $\gamma \leq \Gamma(\gamma) < \Gamma(S(\gamma))$. Hence the claim follows from item 3 of that theorem. □

COROLLARY 6. *Feferman-Schütte's ordinal* $\Gamma(0)$ *(usually denoted* Γ_0*) is definable by a term of* $\mathcal{PW}$*. So are much bigger ordinals, like* $\Gamma(\Gamma_0)$*.*

NOTE 23. It is not difficult to define in $\mathcal{PW}$ a relation R on ω such that $\mathcal{PW}$ proves that $\langle \omega, R \rangle$ is isomorphic to $\langle \Gamma_0, \in \rangle$. This can be done, e.g., by using the recursive well-ordering of the natural numbers which is constructed in [42] (with the help of notations for the ordinals smaller than Γ_0.)

§8. Conclusion and Further Research. As recalled by Feferman in [24], Kreisel criticized in [31] existing proof theory for "the lack of a clear and convincing analysis of the choice of methods of proof," and took as his ultimate aim "the discovery of objective criteria for such a choice". In this paper we have done exactly this for predicative set theory, using invariance of definitions and statements as our main criterion, following by this the ideas of Poincaré and Weyl. What is more, we have shown that the power of predicative reasoning goes well beyond the accepted Γ_0 limit given to it by Feferman and Schütte.

At this point it should be emphasized that we are *not* claiming that the predicative system $\mathcal{PW}$ which is developed in this paper is in any way complete for predicative set theory. Given Weyl's views about the open-ended nature of predicativity (which are adopted and followed in this paper), it is hard to believe that such a complete system exists — even from the point of view of a Platonist who tries to determine "from the outside" the extension and limit of predicative reasoning (as Feferman explicitly tried to do in [15]). Thus, in this paper we have deliberately confined ourselves only to methods that were accepted in one way or another by Feferman in some of his systems. However, there is no reason to continue to do so in future predicative extensions of $\mathcal{PW}$. One obvious direction here is to investigate what sorts of inductive definitions of operations *and predicates* are predicatively acceptable. (Note that $\mathcal{PW}$ allows to introduce new predicate symbols only via explicit definitions, but there is no reason to forbid predicative implicit definitions of predicates, as long as the invariance condition is observed.) According to the principles which guide us in this work, such a definition should be acceptable whenever it *uniquely* and invariantly determines in our framework the predicate or operation which it defines. This mean that an inductive definition which uniquely determines only some minimal predicate or operation is not acceptable. In contrast, an example of an implicit definition of a predicate that *should* be acceptable is the following inductive characterization of $\in^\star$, the transitive closure of $\in$: $y \in^\star x \leftrightarrow y \in x \vee \exists z \in x.y \in z$.

Another important goal for further research is to develop mathematics in $\mathcal{PW}$ (or in a predicative extension of it) in a way which is as natural as possible. Significant work in this direction has started in [7], and is extended and corrected in the Ph.D thesis of Nissan Levy [33]. (See also [32] for a part of his work.)

Finally, two interesting technical questions concerning $\mathcal{PW}$, which we have not tried to answer yet, are:

- What is the proof-theoretic ordinal of $\mathcal{PW}$?
- What is the minimal ordinal (from a platonistic point of view) that is not definable by a term of $\mathcal{PW}$? (Given a set theory S, we call such an ordinal *the set-theoretical ordinal* of S.)

Acknowledgment. I am grateful to Laura Crosilla, Jerhard Jäger and Nik Weaver for helpful comments on the first draft of this paper.

REFERENCES

[1] S. Abiteboul, R. Hull, and V. Vianu, ***Foundations of databases***, Addison-Wesley, 1995.

[2] R. Adams and Z. Luo, *Weyl's predicative classical mathematics as a logic-enriched type theory*, ***ACM Transactions on Computational Logic***, vol. 11 (2010), pp. 1–29.

[3] A. Avron, *Safety signatures for first-order languages and their applications*, ***First-order logic revisited*** (Hendricks et all, editor), Logos Verlag Berlin, 2004, pp. 37–58.

[4] ———, *Constructibility and decidability versus domain independence and absoluteness*, ***Theoretical Computer Science***, vol. 394 (2008), pp. 144–158.

[5] ———, *A new approach to predicative set theory*, ***Ways of proof theory*** (R. Schindler, editor), Onto Series in Mathematical Logic, vol. 2, onto verlag, 2010, pp. 31–63.

[6] ———, *Weyl reexamined: "Das Kontinuum" 100 years later*, ***The Bulletin of Symbolic Logic***, vol. 26 (2020), no. 1, pp. 26–79.

[7] A. Avron and L. Cohen, *Formalizing scientifically applicable mathematics in a definitional framework*, ***Journal of Formalized Reasoning***, vol. 9 (2016), no. 1, pp. 53–70.

[8] A. Avron, S. Lev, and N. Levi, *Safety, absoluteness, and computability*, ***27th eacsl annual conference on computer science logic (csl 2018)***, Schloss Dagstuhl-Leibniz-Zentrum fuer Informatik, 2018.

[9] J. Barwise, ***Admissible sets and structures: An approach to definability theory***, Perspective in Mathematical Logic, Springer-Verlag, Berlin, 1975.

[10] L. Cohen and A. Avron, *Applicable mathematics in a minimal computational theory of sets*, ***Logical Methods in Computer Science***, vol. 14 (2018).

[11] L. Crosilla, *Predicativity and Feferman*, ***Feferman on foundations***, Springer, 2017, pp. 423–447.

[12] ———, *Exploring predicativity*, ***Proof and computation: Digitization in mathematics, computer science, and philosophy***, World Scientific, 2018, pp. 83–108.

[13] ———, *Predicativity and constructive mathematics*, ***Objects, structures, and logics: Filmat studies in the philosophy of mathematics*** (G. Oliveri, C. Ternullo, and S. Boscolo, editors), Springer International Publishing, 2022, pp. 287–309.

[14] K. J. Devlin, ***Constructibility***, Perspectives in Mathematical Logic, Springer-Verlag, 1984.

[15] S. Feferman, *Systems of predicative analysis*, ***The Journal of Symbolic Logic***, vol. 29 (1964), no. 1, pp. 1–30.

[16] ———, *Predicative provability in set theory*, ***Bulletin of the American Mathematical Society***, vol. 72 (1966), no. 3, pp. 486–489.

[17] ———, *Persistent and invariant formulas for outer extensions*, ***Compositio Mathematica***, vol. 20 (1968), pp. 29–52.

[18] ———, *Predicatively reducible systems of set theory*, ***Axiomatic set theory: Proceedings of symposia in pure mathematics, part ii*** (D. S. Scott and T. Jech, editors), vol. 13, American mathematical society (providence), 1974, pp. 11–32.

[19] ———, *A more perspicuous formal system for predicativity*, ***Konstruktionen versus positionen*** (K. Lorenz, editor), vol. 2, Walter de Gruyters, Berlin, 1979, pp. 68–93.

[20] ———, *Weyl vindicated: Das Kontinuum 70 years later*, ***Remi e prospettive della logica e della scienza contemporanee*** (C. Cellucci and G. Sambin, editors), Cooperative Libraria Universitaria Editrice, Bologna, 1988, Reprinted with postcript in [23], 249-283.

[21] ———, *Reflecting on incompleteness*, ***The Journal of Symbolic Logic***, vol. 56 (1991), no. 1, pp. 1–49.

[22] ———, *Why a little bit goes a long way: Logical foundations of scientifically applicable mathematics*, ***Proceedings of the biennial meeting of the philosophy of science association***, vol. 2, 1992, Reprinted in [23], 284–298, pp. 442–455.

[23] ———, ***In the light of logic. logic and computation in philosophy***, Oxford University Press, 1998.

[24] ———, *Predicativity*, ***The oxford handbook of philosophy of mathematics and logic*** (S. Shapiro, editor), Oxford University Press, 2005, pp. 590–624.

[25] ———, *Why a little bit goes a long way: Logical foundations of scientifically applicable mathematics*, (2013), unpublished manuscript.

[26] S. Feferman and T. Strahm, *The unfolding of non-finitist arithmetic*, ***Annals of Pure and Applied Logic***, vol. 104 (2000), no. 1-3, pp. 75–96.

[27] M. Hallett, *Absoluteness and the skolem paradox*, ***Logic, mathematics, philosophy, vintage enthusiasms: Essays in honour of john l. bell*** (D. DeVidi, M. Hallett, and P. Clarke, editors), Springer Netherlands, Dordrecht, 2011, pp. 189–218.

[28] R. B. Jensen, *The fine structure of the constructible hierarchy*, ***Annals of mathematical logic***, vol. 4 (1972), no. 3, pp. 229–308.

[29] G. Kreisel, *Ordinal logics and the characterization of informal concepts of proof*, ***Proceedings of the international congress of mathematicians (edinburgh 1958)***, Cambridge University Press, 1958, pp. 288–299.

[30] ———, *The axiom of choice and the class of hyper- arithmetic functions*, ***Indagationes Mathematicae***, vol. 24 (1962), pp. 307–319.

[31] ———, *Principles of proof and ordinals implicit in given concepts*, ***Intuitionism and proof theory*** (A. Kino and et al., editors), North-Holland, 1970, pp. 489–516.

[32] N. Levi and A. Avron, *Analysis in a formal predicative set theory*, ***International workshop on logic, language, information, and computation***, Springer, 2021, pp. 167–183.

[33] Nissan Levi, *Predicativity and computability*, ***Ph.D. thesis***, Tel Aviv University, Israel, 2023.

[34] E. Nelson, ***Predicative arithmetic***, Mathematical notes, vol. 32, Princeton University Press, 1986.

[35] H. Poincaré, *Les mathématiques et la logique*, ***Revue de Métaphysique et de Morale***, vol. 13 (1905), pp. 813–835.

[36] ———, *Les mathématiques et la logique, ii*, ***Revue de Métaphysique et de Morale***, vol. 14 (1906), pp. 294–317.

[37] ———, *La logique de l'infini*, ***Revue de Métaphysique et de Morale***, vol. 17 (1909), no. 4, pp. 461–482.

[38] ———, *La logique de l'infini*, ***Scientia***, vol. 12 (1912), pp. 1–11.

[39] J. Richard, *Les principes des mathématiques et le problème des ensembles*, ***Revue générale des sciences pures et appliquées***, vol. 12 (1905), no. 16, p. 541.

[40] B. Russell, *Mathematical logic as based on the theory of types*, ***American journal of mathematics***, vol. 30 (1908), no. 3, pp. 222–262.

[41] K. Schütte, *Predicative well-orderings*, ***Formal systems and recursive functions*** (J. Crossley and M. Dummett, editors), North-Holland (Amsterdam), 1965, pp. 279–302.

[42] ———, ***Proof theory***, Grundlangen der mathematischen Wissenschften, vol. 225, Springer-Verlag, 1977.

[43] S. Shapiro, ***Foundations without foundationalism: A case for second-order logic***, vol. 17, Clarendon Press, 1991.

[44] J. R. Shoenfield, ***Mathematical logic***, AK Peters/CRC Press, 2018.

[45] S. G. Simpson, ***Subsystems of second order arithmetic***, perspectives in mathematical logic, Springer, 1999.

[46] J.D. Ullman, ***Principles of database and knowledge-base systems***, Computer Science Press, 1988.

[47] H. Wang, *The formalization of mathematics*, ***The Journal of Symbolic Logic***, vol. 19 (1954), pp. 241–266.

[48] N. Weaver, *Mathematical conceptualism*, ***arXiv preprint math/0509246***, (2005).

[49] ———, *Predicativity beyond gamma_0*, ***arXiv preprint math/0509244***, (2005).

[50] ———, *What is predicativism?*, unpublished manuscrit, 2013.

[51] H. Weyl, ***Das Kontinuum: Kritische untersuchungen uber die grundlagen der analysis***, Veit, Leipzig, 1918.

[52] ———, *Mathematics and logic: A brief survey serving as a preface to a review of the philosophy of bertrand russell*, ***American mathematical Monthly***, vol. 53 (1946), pp. 2–13.

[53] ———, ***The continuum: A critical examination of the foundation of analysis***, Thomas Jefferson University Press, 1987, Translated by Pollard, S. and Bole, T.

[54] AN Whitehead and B Russell, ***Principia mathematica***, vol. 1, Cambridge University Press, 1925, 2nd edition.

SCHOOL OF COMPUTER SCIENCE
TEL AVIV UNIVERSITY
TEL AVIV 6997801, ISRAEL
E-mail: aa@tauex.tau.ac.il