A MANY-SORTED APPROACH TO PREDICATIVE MATHEMATICS

Julian L. Hook

A MANY-SORTED APPROACH TO PREDICATIVE MATHEMATICS

Julian L. Hook

A DISSERTATION

PRESENTED TO THE

FACULTY OF PRINCETON UNIVERSITY

IN CANDIDACY FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE BY THE

DEPARTMENT OF

MATHEMATICS

June 1983

Contents

Part Three: Predicative Set Theory

A Many-Sorted Approach to Predicative Mathematics

Julian L. Hook

Abstract

We investigate calculus and set theory in a predicative theory whose consistency can be proved finitarily. Real numbers are introduced as objects of a second sort corresponding to equivalence classes of quotients of integers modulo infinitesimals; in particular, the theory of real closed fields is interpretable in our theory. Methods resembling nonstandard analysis are used in discussing calculus, with additional sorts representing sets of real numbers and continuous functions. A few second-order properties of the real numbers are discussed, and it is shown that similar techniques can be used to construct the p-adic numbers. We formulate a theory of infinite cardinals quite different from the Cantorian theory. The use of many-sorted theories, usually considered dispensable, is found to be essential in that our interpretability results would fail if the corresponding one-sorted theories were studied instead.

§0. *Introduction*

Classical mathematics, it has long been recognized, is impredicative. One of the primary sources of this impredicativity is the induction principle: each bound variable in each of the induction axioms in Peano arithmetic is understood to range over all "numbers" -- that is, over all objects satisfying all the axioms, including the one in question. From a classical point of view such circularity is entirely permissible; on the other hand, those who find it disturbing will be interested to know how much mathematics can be done predicatively. Our goal here is a predicative formulation of two branches of mathematics, calculus and set theory, using as a foundation the theory of predicative arithmetic developed by Nelson [1]. (References are listed at the end of this Introduction.)

Summary of Results

The present work is divided into three parts. Part One, consisting of §§1-3, is preparatory. The first section summarizes the main features of Nelson's theory $\underset{\sim}{Q}^0$, including the important concept of bounded induction and the nonclassical behavior of unbounded notions such as exponentiation. Also arithmetical in nature is §2, in which are presented refinements of $\underset{\sim}{Q}^0$ that will prove useful later. This section also illustrates a procedure to be followed continually: we strengthen a theory by adjoining new axioms, and show that the new theory is interpretable in the old. For technical reasons as well as for convenience, some of our theories will be many-sorted; an abstract logical introduction to many-sorted theories, including the appropriate definition of an interpretation, comprises §3.

The most important results presented here are probably the predicative reproductions of theorems of classical analysis in Part Two. After §4, in which "fractions" are introduced in the theory Q^0, the heart of the matter is reached in §5. Here the properties of exponentiation are used to divide numbers into two types, the "finite" and the "infinite". (There is an obvious precedent here in nonstandard analysis. On the other hand, it should be noted that the situation here corresponds to nothing in classical arithmetic, since the induction principle is explicitly violated: 0 is finite, and if n is finite then so is $n+1$, but not all numbers are finite.) This distinction induces notions of infinite fractions and of infinitesimals, and a relation $\sim$ of infinite closeness. Real numbers are introduced as objects of a second sort corresponding to equivalence classes of finite fractions modulo the relation $\sim$.
It follows fairly immediately that the real numbers satisfy the usual first-order axioms for real closed fields.

The theory is expanded further in §6. Additional sorts are adjoined for, among other things, continuous functions on the real numbers and closed sets of real numbers. Calculus is discussed in §7, the derivative being defined as the real number represented by a difference quotient $((f(b)-f(a))/(b-a)$ in which a and b are fractions such that $a \sim b$ but $a \neq b$.

Perhaps the flavor of the subject is best conveyed by citing a few specific theorems. By (6.28), if f is a continuous function and γ_1 and γ_2 are real numbers with $\gamma_1 \leq \gamma_2$, then

$$\mathrm{Dom}f = [\gamma_1,\gamma_2] \longrightarrow \exists\alpha(\alpha\in[\gamma_1,\gamma_2]\&\forall\beta(\beta\in[\gamma_1,\gamma_2] \longrightarrow f(\beta) \leq f(\alpha)))$$

-- that is, every continuous function on a closed interval attains a maximum. The statement of (7.2) is

$$\alpha<\beta\&[\alpha,\beta] \subseteq \mathrm{Dom}f\&\forall\gamma(\alpha<\gamma<\beta \longrightarrow f \text{ is differentiable at } \gamma) \longrightarrow$$

$$\exists\gamma(\alpha<\gamma<\beta\&\mathrm{Deriv}(f,\gamma) = (f(\beta)-f(\alpha))/(\beta-\alpha))$$

-- the mean value theorem. Finally, according to (7.27) and (7.28),

$$\alpha<\beta\&[\alpha,\beta] \subseteq \mathrm{Dom}f \longrightarrow \exists!g(\mathrm{Dom}g = [\alpha,\beta]\&\forall\gamma(\alpha<\gamma<\beta \longrightarrow$$

$$g \text{ is differentiable at } \gamma\&\mathrm{Deriv}(g,\gamma) = f(\gamma))\&g(\alpha) = 0)$$

-- every continuous function has a unique antiderivative up to an additive constant. The point of listing these theorems here is to emphasize that they are not poor approximations to classical theorems; rather, we are able to give predicative formulations of the concepts of analysis in such a way that the classical results are provable without so much as a change in statement.

The situation changes somewhat in §8, in which we discuss rational and algebraic numbers and decimal expansions; here the correspondence with traditional mathematics, though not hard to find, is less exact. In the final section of Part Two we show how the techniques used in constructing the real numbers can be modified to construct the p-adic numbers; we carry p-adic analysis as far as Hensel's lemma.

While predicative analysis duplicates many results of classical mathematics, the predicative set theory that is the subject of Part Three is decidedly nonclassical in appearance. The theory $\underset{\sim}{Q}^0$ includes a notion of set, but it is a bounded notion, and many collections of numbers -- for instance, the collection of all "finite" numbers -- do not form sets. The first task facing us, which we tackle in §10, is finding a way to refer to such collections within the theory. Using the arithmetical hierarchy of recursion theory as a guide, we define notions of Δ_2-collections, Δ_3-collections,... . Finally, in §11 is developed a theory of infinite cardinality. The cardinality of a collection is determined not by one-to-one corresondences but by approximating the collection from above and below by *sets* (in the sense of Q^0) and looking at *their* cardinalities. Infinite cardinals are partially ordered but in general not totally ordered; certain collections called "pseudosets", however, have cardinalities that are totally ordered. We prove several order properties of cardinals, and observe that pseudoset-cardinals admit very natural and interesting operations of addition and multiplication.

Why predicative mathematics?

The predicative point of view is particularly compatible with the philosophy of formalism, or nominalism -- namely, mathematics consists of a body of axioms, theorems, and formal proofs. Mathematical objects have no real existence, nor theorems physical significance; Lebesgue measure, uncountable sets, forcing, and the category of all categories are symbolic constructions and nothing more. Many mathematicians, even those who consider themselves formalists, seem to believe that the set ω of natural numbers really exists, and that the induction principle with all its impredicativity is unimpeachable because it is "correct". Following Nelson,' let us write

$x \wedge y$ for x^y and $2\uparrow n$ for $2\wedge(2\wedge(2\wedge\ldots\wedge 2))$ with n occurrences of 2; then, according to this belief, numbers like $2\uparrow 5$ or $2\uparrow(2\uparrow 5)$ are every bit as real as the number of women on the Supreme Court or the number of light bulbs in Maine. But numbers, too, are symbolic constructions, and "a construction", writes Nelson in [1,§1], "does not exist until it is made".

> The point is that to regard $2\uparrow 5$ as standing for a genetic number entails a philosophical commitment to some idealistic notion of existence. To a nominalist, $2\uparrow 4$ stands for a number, 65536, to which one can count; but $2\uparrow 5$ is a pair of numerals with a vertical arrow between them, and there is not a scintilla of evidence that it stands for a genetic number... . The infant counts on its fingers, the mathematician counts on ω - but the infant at least knows its fingers to exist. The mathematician's attitude towards ω has in practice been one of faith, and faith in a hypothetical entity of our own devising, to which are ascribed attributes of necessary existence and infinite magnitude, is idolatry. [1,§18]

As an indictment of impredicative methods, [1,§18] is clearly a hard act to follow; the reader is referred there for further discussion.

One of the attractive features of the theory $\mathcal{Q}^0$ is that it can be proved consistent with very little machinery: one can give a finitary proof using the "Hilbert Ansatz". (No claim is made for a *predicative* consistency proof; indeed, Gödel's incompleteness theorem applies to $\mathcal{Q}^0$, so one could not hope to prove the consistency of $\mathcal{Q}^0$ *within* $\mathcal{Q}^0$.) All the theories we shall use in our predicative investigation of mathematics will be proved consistent relative to $\mathcal{Q}^0$, usually via the construction of an interpretation, so these theories have finitary consistency proofs also. Irrespective of qualms about impredicativity, a mathematician -- especially a formalist -- should embrace any theory known to be both consistent and productive; we might summarize our work by saying that predicative mathematics satisfies these criteria.

Is predicative mathematics more trustworthy than classical mathematics? For some of us, at least, the answer is yes. Is it more productive, or even comparably productive? That remains to be seen, but it is at least productive enough to stand on its own. Is it more *natural*? That is the most subjective question of all. Occasionally someone speculates about what the mathematics of an alien civilization might look like; I, for one, find it rather less difficult to imagine little green men doing mathematics predicatively than to imagine them studying Lebesgue measure and forcing.

Acknowledgements

I wish to thank Simon Kochen for several helpful discussions, Hale Trotter for listening with interest and patience as I rambled about my work, and numerous graduate students for encouragement and friendship. It should be obvious already, though, that my biggest debt is to my advisor, Edward Nelson. It was he who kindled my interest in predicative mathematics in the first place; it was he who built the mathematical foundation on which my work is based and suggested the basic aims of my research; it was he who assured me that transcendental numbers were there if I would only look hard enough, pointed out that my construction of "the 7-adic numbers for all values of 7" could be modified to handle the p-adic numbers for all p at once, and gave me numerous other pushes in usually-appropriate directions. I am very grateful to have had an advisor so friendly, so encouraging,

and so able to tolerate the ups and downs of a student who somehow managed to take just enough time off from playing the piano to write a thesis in mathematics.

References

Our primary reference is the forthcoming book

[1] Edward Nelson, *Predicative Arithmetic*, to appear,

in which the groundwork for a predicative investigation of mathematics is carefully and completely laid. The best place to turn for logical preliminaries is

[2] Joseph R. Shoenfield, *Mathematical Logic*, Addison-Wesley (1967),

although the reader may find the discussion of eliminability of defined symbols in §74 of

[3] Stephen Cole Kleene, *Introduction to Metamathematics*, Van Nostrand (1952)

more straightforward, if less elegant, than Shoenfield's.

Many-sorted theories were investigated as early as 1938:

[4] A. Schmidt, Über deduktive Theorien mit mehreren Sorten von Grunddingen, *Math. Ann.* 115 (1938), 485-506.

A concise discussion in English can be found in Chapter 29 of

[5] J. Donald Monk, *Mathematical Logic*, Springer (1976),

and a more leisurely-paced one in Chapter XII of

[6] Hao Wang, *A Survey of Mathematical Logic*, North-Holland (1953);

the syntactic properties that are our primary concern in §3, however, seem not to have been widely studied. The later chapters of Wang's book present a version of predicative set theory based on a variant of Russell's theory of types.

We do not actually use nonstandard analysis here, but our handling of the real numbers in Part Two is strongly reminiscent of nonstandard analysis, particularly the syntax-oriented approach in

[7] Edward Nelson, Internal Set Theory: A New Approach to Nonstandard Analysis, *Bull. Amer. Math. Soc.* 83 (1977), 1165-98.

In addition to similar discussions of infinite and infinitesimal numbers and corresponding treatments of calculus, the reader will note parallels between the "internal" formulas of nonstandard analysis and the "bounded" formulas of predicative arithmetic. Most of the essential differences between our theory and nonstandard analysis stem from the essentially different behavior of exponentiation: whereas in nonstandard analysis a finite number raised to a finite power is finite, that is not true here at all. (In other words, exponentiation is internal but unbounded.) Some infinite numbers may differ from finite numbers by only one exponentiation, others by two or three or more; in this way the various degrees of exponentiability give rise to various levels of "infiniteness" and "infinitesimalness", a distinction that is lacking in ordinary nonstandard analysis. The notion that some infinitesimals may be more infinitesimal than others is investigated in a different context in

[8] Dawn Fisher, Extending Functions to Infinitesimals of Finite Order, *Am. Math. Monthly* 89 (1982), 443-9.

Finally, there are certain resemblances between our theory and the "Alternative Set Theory" described in

[9] Antonín Sochor, The Alternative Set Theory, in *Set Theory and Hierarchy Theory: A Memorial Tribute to Andrzej Mostowski*, Springer Lecture Notes in Math. 537 (1976), 259-71

and more completely in

[10] Peter Vopěnka, *Mathematics in the Alternative Set Theory*, B.G. Teubner (1979).

For instance: all sets in the strict sense are finite, but infinity arises because subclasses of sets need not be sets; real numbers are regarded as equivalence classes of quotients of integers. The methods used in Alternative Set Theory, though, are highly impredicative.

PART ONE

PRELIMINARIES

§1. *Fundamentals of Predicative Arithmetic*

This section is intended to serve both as a review of predicative arithmetic for those who have read Nelson's book [1] and as a summary of its relevant aspects for those who have not. One caveat is in order: the organization throughout is with an eye toward concise presentation of the essential points rather than toward systematic logical development. Much material that is vital for later parts of [1] but is not otherwise relevant here has been omitted, and much has been rearranged. The moral is that he who attempts to reconstruct all of Nelson's work using only this brief sketch as a guide faces a considerable challenge!

Robinson's theory

In the course of [1] Nelson builds up a powerful theory, here called Q^0, that will serve as the starting point for our investigation of predicative mathematics. At the heart of Q^0 is Raphael Robinson's theory Q. In a formulation that is particularly convenient in that all the nonlogical axioms are quantifier-free, this theory has as its nonlogical symbols the constant 0, unary function symbols S ("successor") and P ("predecessor"), and binary function symbols + and $\cdot$. The nonlogical axioms (hence the designation Ax) of Q are

1.1) $$\text{Ax } Sx \neq 0 \ ,$$

1.2) $$\text{Ax } Sx = Sy \longrightarrow x = y \ ,$$

1.3) $$\text{Ax } x+0 = x \ ,$$

1.4) $$\mathrm{A}x\ x + Sy = S(x+y)\ ,$$

1.5) $$\mathrm{A}x\ x \cdot 0 = 0\ ,$$

1.6) $$\mathrm{A}x\ x \cdot Sy = x \cdot y + x\ ,$$

and

1.7) $$\mathrm{A}x\ Px = y \longleftrightarrow Sy = x \vee (x = 0 \,\&\, y = 0)\ .$$

A finitary proof that Robinson's theory is consistent can be given using the Hilbert-Ackermann consistency theorem (the "Hilbert Ansatz").

The theory Q is of course sufficient for some arithmetical purposes. In it, for instance, we can prove the formulas

1.8) $$x \cdot S0 = x$$

and

1.9) $$x \neq 0 \longrightarrow \exists y(Sy = x)\ ,$$

and the binary predicate symbol $\leq$ can be immediately adjoined to Q via the defining axiom

1.10) $$\mathrm{Def}\ x \leq y \longleftrightarrow \exists z(x+z = y)\ .$$

Often of greater usefulness than provability or definability in Q, though, is the notation of *interpretability* in Q (see [2,§4.7]). We may regard as predicative any theory that is shown to be interpretable in Q -- that is, any theory for which an interpretation in some extension by definitions of Q can be constructed. Of course, the consistency of such a theory follows from that of Q. In [1,§6], Robinson's theory is quickly enlarged by the adjunction of the associative, distributive, and commutative laws, and the new theory is shown to be interpretable in the old. Each time we add new nonlogical axioms to our ever-growing theory (actually, each time with one exception to be noted in §5), we have a moral obligation to prove such an interpretability result.

Many more common symbols (in addition to $\leq$) can be defined in the theory being constructed. The following expressions all make sense in Nelson's theory Q^0 , and behave in ways that one might expect: $x < y$, $x-y$ (defined to be 0 if $x < y$), $Qt(x,y)$ and $Rm(x,y)$ (the quotient and remainder upon division of x into y), $Max(x,y)$ (the larger of the two numbers), $x|y$ (the divisibility relation), x is a prime. We can define the decimal digits 1,...,9, and may use ordinary decimal notation to refer to particular numbers. There are also unary function symbols $|\ |_2$ ($|x|_2$ is the largest power of 2 not exceeding x ; $|0|_2 = 1$) and Log (integer-valued logarithm to the base 2 ; $\text{Log } x = \text{Log } |x|_2$; $\text{Log } 0 = 0$). Conspicuously absent from this list is exponentiation, not so much because it is complicated as because it is ill-behaved. In fact, we shall see later in this section that a binary function symbol $\wedge$ for exponentiation *can* be defined in Q^0 ; the way in which it fails to be "well-behaved", though, should soon be apparent.

Smash

One of the most important features of the theory Q^0 is a binary function symbol # (pronounced "smash"). The basic property of this operation is $2^k \# 2^\ell = 2^{k \cdot \ell}$; that is, on powers of 2 , #

is to $\cdot$ as $\cdot$ is to $+$. For general numbers x and y , not necessarily powers of 2 , the value of $x\#y$ is the same as $|x|_2 \# |y|_2$. Therefore $x\#y$ is always a power of 2 . Smash is commutative, associative, and almost distributive over multipication (actually, $x\#(|y|_2 \cdot |z|_2) = (x\#|y|_2)\cdot(x\#|z|_2)$; it also satisfies $x\#1 = 1$, $x\#2 = |x|_2$, and $y \leq z \longrightarrow x\#y \leq x\#z$. Axioms describing some of these fundamental facts are a part of Q^0 , and the necessary interpretability result is proved in [1, §15].

The reader who wishes to think of $x\#y$ as $2\wedge(\text{Log}\, x \cdot \text{Log}\, y)$ may do so; indeed, that equality is a theorem of Q^0 . It cannot be emphasized too strongly, however, that exponentiation is not invariably so nice, and moreover that the "next symbol in line" after 0 , S , + , and $\cdot$ is not $\wedge$ but # .

Induction by relativization

The principle of mathematical induction, as discussed at some length in [1], is impredicative and has no place in the theory Q^0 . The objection that very little significant mathematics can be done without induction is well-founded, however; in fact, many induction proofs *can* be carried out in Q^0 . This point deserves elaboration.

To avoid depletion of valuable notational resources, let us continue to denote by Q the theory that has been described thus far. Let $\mathbb{C}[x]$ be a unary formula (that is, a formula with only the one free variable x) in the language of Q . Assume further that $\mathbb{C}$ is *inductive* (or, more precisely, inductive in the variable x in the theory Q) ; this means that $\mathbb{C}[0]\ \&\ \forall x(\mathbb{C}[x] \longrightarrow \mathbb{C}[Sx])$ is a theorem of Q. The induction principle would allow us to conclude that $\mathfrak{C}[x]$ holds for all x. Such an inference is for us unacceptable, though, unless we can show that the theory $Q[\mathfrak{C}]$ obtained from Q by adjoining $\mathfrak{C}$ as a new axiom is interpretable in Q.

As a first attempt toward constructing such an interpretation, we might try relativizing by the formula $\mathfrak{C}$ itself -- in effect, refining our concept of "number" so that only objects satisfying $\mathfrak{C}$ are considered. There are two difficulties with this approach. First, the formula $\mathfrak{C}$ might fail to respect the function symbols of Q. (To say that $\mathfrak{C}$ respects a function symbol $\underline{f}$ in Q means that $\mathfrak{C}[x_1]\&\mathfrak{C}[x_2]\&\ldots\&\mathfrak{C}[x_\lambda] \longrightarrow \mathfrak{C}[\underline{f}x_1x_2\ldots x_\lambda]$ is a theorem of Q.) Inductivity of $\mathfrak{C}$ ensures that $\mathfrak{C}$ respects 0 and S, but there is no such guarantee for $+$, $\cdot$, or $\#$. (We ignore defined symbols for the moment.) The second problem is that the relativization of $\mathfrak{C}$ by itself may fail to be a theorem of Q, in which case our interpretation fails to be an interpretation of the theory $Q[\mathfrak{C}]$. (By the relativization of $\mathfrak{C}$ by itself we mean the full relativization $\mathfrak{C}^{\mathfrak{C}}[x]$, which in this case is $\mathfrak{C}[x] \longrightarrow \mathfrak{C}_{\mathfrak{C}}[x]$, where $\mathfrak{C}_{\mathfrak{C}}$ is obtained from $\mathfrak{C}$ by relativizing each quantifier to $\mathfrak{C}$.)

The first of these difficulties is the more easily surmounted. A brief description of the method used in [1] follows. Write $\mathfrak{C}^1[x]$ for the unary formula $\forall y(y \leq x \longrightarrow \mathfrak{C}[y])$. Then $\mathfrak{C}^1$ is inductive in x, stronger than $\mathfrak{C}$ (that is, $\mathfrak{C}^1[x] \longrightarrow \mathfrak{C}[x]$ is a theorem of Q), and *hereditary* (which means that $\mathfrak{C}^1[x]\&w \leq x \longrightarrow \mathfrak{C}^1[w]$ is a theorem of Q). Now write $\mathfrak{C}^2[x]$ for the unary formula $\forall y(\mathfrak{C}^1[y] \longrightarrow \mathfrak{C}^1[y+x])$; the formula $\mathfrak{C}^2$ is inductive in x, stronger than $\mathfrak{C}^1$ (hence stronger than $\mathfrak{C}$), and hereditary, and in addition $\mathfrak{C}^2$ respects $+$. Next, write $\mathfrak{C}^3[x]$ for

$\forall y(\mathbb{C}^2[y] \longrightarrow \mathbb{C}^2[y\cdot x])$; the formula $\mathbb{C}^3$ is inductive in x, stronger than $\mathbb{C}$, and hereditary, and $\mathbb{C}^3$ respects both $+$ and $\cdot$. Finally, write $\mathbb{C}^4[x]$ for $\forall y(\mathbb{C}^3[y] \longrightarrow \mathbb{C}^3[y\#x])$. We summarize the properties of $\mathbb{C}^4$ (cf. [1,§15, Proposition III]) in

Metatheorem A . Let T be an extension of the theory Q , and let $\mathbb{C}$ be a unary formula in the language of T . Then the following is a theorem of T :

$$\mathbb{C}[0]\&\forall x(\mathbb{C}[x] \longrightarrow \mathbb{C}[Sx]) \longrightarrow$$

$$(\mathbb{C}^4[x] \longrightarrow \mathbb{C}[x])\&$$

$$(\mathbb{C}^4[x]\&w \leq x \longrightarrow \mathbb{C}^4[w])\&$$

$$\mathbb{C}^4[0]\&$$

$$(\mathbb{C}^4[x] \longrightarrow \mathbb{C}^4[Sx])\&$$

$$(\mathbb{C}^4[x_1]\&\mathbb{C}^4[x_2] \longrightarrow \mathbb{C}^4[x_1+x_2])\&$$

$$(\mathbb{C}^4[x_1]\&\mathbb{C}^4[x_2] \longrightarrow \mathbb{C}^4[x_1\cdot x_2])\&$$

$$(\mathbb{C}^4[x_1]\&\mathbb{C}^4[x_2] \longrightarrow \mathbb{C}^4[x_1\#x_2]) .$$

In light of Metatheorem A, it seems reasonable to construct an interpretation using $\mathbb{C}^4$ instead of $\mathbb{C}$. This method gives us at least an interpretation of the theory Q in itself, since $\mathbb{C}^4$ respects the function symbols of Q and since the axioms of Q are quantifier-free. We still face the problem, though, of whether the interpretation of the formula $\mathbb{C}$ -- namely, $\mathbb{C}^{\mathbb{C}^4}$ -- is a theorem of Q .

Nelson gives two examples that merit repetition here. If $\mathbb{C}[x]$ is $\exists y(SSO\cdot y = x\cdot(x+SO))$, then $\mathbb{C}$ is an inductive unary formula, and its relativization by $\mathbb{C}^4$ is $\mathbb{C}^4[x] \longrightarrow \exists y(\mathbb{C}^4[y]\&SSO\cdot y = x\cdot(x+SO))$. This formula is a theorem of Q , as the following argument shows. Suppose $\mathbb{C}^4[x]$. Then $\mathbb{C}[x]$ by Metatheorem A , so there exists y such that $SSO\cdot y = x\cdot(x+SO)$. Then $y \leq x\cdot(x+SO)$. But we have $\mathbb{C}^4[x]$ and $\mathbb{C}^4[SO]$ (by Metatheorem A), so we have $\mathbb{C}^4[x\cdot(x+SO)]$ (by Metatheorem A), and therefore $\mathbb{C}^4[y]$ (by Metatheorem A again). Thus $\exists y(\mathbb{C}^4[y]\&SSO\cdot y = x\cdot(x+SO))$, as desired. In this case, the interpretation associated with $\mathbb{C}^4$ is an interpretation of $Q[\mathbb{C}]$ in Q , and, as such, gives us the green light to work in the theory $Q[\mathbb{C}]$ if we wish.

Now let $\mathbb{C}[x]$ be $\exists y(y \neq 0\&\forall z(z \neq 0\&z \leq x \longrightarrow \exists w(z\cdot w = y)))$, which asserts that there is a number divisible by all numbers from 1 to x . Again, $\mathbb{C}$ is inductive in x in Q , but the method of proof employed in the preceding paragraph leads nowhere in this instance. There is here no reason to believe that $\mathbb{C}^{\mathbb{C}^4}$ is a theorem of Q , or even that $Q[\mathbb{C}]$ is interpretable in Q at all.

The crucial difference between these two examples is that in the first instance the induction is *bounded.* In other words, we can say in advance just how big, in terms of x and the function symbols of Q (that is, the function symbols appearing in Metatheorem A), the y such that $SSO\cdot y = x\cdot(x+SO)$ will have to be. (Answer: not bigger than $x\cdot(x+SO)$.) In the second instance, on the other hand, no such bound on y is apparent.

Let us make this general and precise.

Bounded induction

Let T be any theory whose language contains the binary predicate symbol $\leq$, and let $\mathbb{A}$ be a formula in this language. The initial occurrence of $\exists \underline{x}$ in a subformula $\exists \underline{x} \mathbb{B}$ of $\mathbb{A}$ is said to be *manifestly bounded* if $\mathbb{B}$ is of the form $\underline{x} \leq \underline{a} \& \mathbb{B}'$, where $\underline{a}$ is a term not containing the variable $\underline{x}$. The formula $\mathbb{A}$ is *manifestly bounded* if every occurrence of an existential quantifier is manifestly bounded; we regard $\forall$ as having been defined in terms of $\exists$, so the condition actually applies to *all* quantifiers in $\mathbb{A}$.

The important property of manifestly bounded formulas is essentially what was checked for the formula $\mathbb{C}$ in the first example above, namely the following "reflection principle". (Notation: $\mathbb{C}(\text{free } \mathbb{A})$ stands for $\mathbb{C}[\underline{x}_1] \& \ldots \& \mathbb{C}[\underline{x}_\lambda]$, where $\underline{x}_1, \ldots, \underline{x}_\lambda$ are the variables occurring free in $\mathbb{A}$.)

Metatheorem B. Let T' be a theory containing the binary predicate symbol $\leq$, let T be an extension of T' , and let $\mathbb{C}$ be a unary formula of T that respects all function symbols of T' and is hereditary. Let $\mathbb{A}$ be a manifestly bounded formula of T' . Then the formula $\mathbb{C}(\text{free } \mathbb{A}) \longrightarrow (\mathbb{A} \longleftrightarrow \mathbb{A}_{\mathbb{C}})$ is a theorem of T .

(This is basically [1,§7, Metatheorem 4].) It is a straightforward matter to show that Metatheorem B leads in short order to

Metatheorem C. Let $\mathbb{A}$ be a manifestly bounded formula of Q that is inductive in one of its free variables. Then $Q[\mathbb{A}]$ is interpretable in Q .

(The interpretation is determined by $\mathbb{C}^4$, where $\mathbb{C}$ is the unary formula obtained from $\mathbb{A}$ by appending to the front a universal quantifier on each free variable other than the one in which $\mathbb{A}$ is inductive.)

Let us now denote by Q' the extension of Q obtained by adjoining as new nonlogical axioms all "manifestly bounded induction" formulas of the form

MBI) $\mathrm{Ax}\ \mathbb{A}[0]\&\neg\exists y(y\leq x\&\mathbb{A}[y]\&\neg\mathbb{A}[Sy]) \longrightarrow (y\leq x \longrightarrow \mathbb{A}[y])$,

where $\mathbb{A}$ is a manifestly bounded formula in the language of Q . To legitimize Q' , we record

Metatheorem D. Let $\mathbb{B}_1,\ldots,\mathbb{B}_\lambda$ be theorems of Q' . Then $Q[\mathbb{B}_1,\ldots,\mathbb{B}_\lambda]$ is interpretable in Q .

(See [1,§7, Metatheorem 6]. The key observation is that each new axiom (MBI) is manifestly bounded and inductive in x in Q .)

It is an easy exercise to prove that the induction formula

Ind) $$(\mathbb{A}[0]\&\forall x(\mathbb{A}[x] \longrightarrow \mathbb{A}[Sx])) \longrightarrow \mathbb{A}[x]$$

is a theorem of Q' if $\mathbb{A}$ is manifestly bounded. The same is therefore true if $\mathbb{A}$ is simply *bounded* in Q' -- that is, if $\mathbb{A}$ is provably equivalent, in Q' , to a manifestly bounded formula. (This is the case with the formula $\exists y(SS0\cdot y = x\cdot(x+S0))$ considered earlier.) A mention of "bounded induction" in a proof is simply a reference to a theorem of the form (Ind) with $\mathbb{A}$ bounded.

Another useful form of induction available in Q' is the "bounded least number principle", which is

BLNP) $\exists x_1 \cdots \exists x_\lambda \mathbb{A}[x_1,\ldots,x_\lambda] \longrightarrow$

$$\exists x_1 \cdots \exists x_\lambda (\mathbb{A}[x_1,\ldots,x_\lambda] \& \neg \exists y_1 \cdots \exists y_\lambda (y_1 \leq x_1 \ \cdots \ y_\lambda \leq x_\lambda \& y_1 \neq x_1 \& \mathbb{A}[y_1,\ldots,y_\lambda]))$$

with $\mathbb{A}$ a bounded formula of $\mathcal{Q}'$. If λ is 1 , it is convenient to write $\min_x \mathbb{A}$ for the (bounded) formula $\mathbb{A}[x] \& \neg \exists y (y<x \& \mathbb{A}[y])$, so that (BLNP) becomes $\exists x \mathbb{A}[x] \longrightarrow \exists x \min_x \mathbb{A}$. As shown in [1,§8], every formula (BLNP) with $\mathbb{A}$ bounded is a theorem of $\mathcal{Q}'$.

The notion of boundedness can be extended to apply to defined symbols. A predicate symbol $\underline{p}$ adjoined to $\mathcal{Q}'$ via a defining axiom $\underline{p}x_1 \cdots x_\lambda \longleftrightarrow \mathbb{D}$ is said to be *bounded* if the formula $\mathbb{D}$ (in the language of $\mathcal{Q}'$) is bounded. Likewise, a function symbol $\underline{f}$ adjoined to $\mathcal{Q}'$ via a defining axiom $\underline{f}x_1 \cdots x_\lambda = y \longleftrightarrow \mathbb{D}$ (with appropriate existence and uniqueness conditions holding in $\mathcal{Q}'$) is *bounded* if the formula $\exists y \mathbb{D}$ (not just $\mathbb{D}$) is bounded. A standard fact about defined symbols is that they are eliminable, in the sense that every formula involving a defined symbol can be effectively replaced by an equivalent formula in which the symbol does not appear (see [3,§74]); a symbol is bounded precisely if its elimination from a bounded formula always yields a bounded formula. In the theory $\mathcal{Q}^0$ we may apply bounded induction to any bounded formula in which all defined symbols are bounded.

The symbols $\dot{-}$, Qt , Rm , Max , $|$, is a prime, $|\ |_2$, and Log mentioned previously are all bounded. (We can now see one of the several reasons why the operation $\#$ is so vital a part of $\mathcal{Q}^0$: numerous formulas and defined symbols, including Log , can be shown to be bounded in terms of $\#$ but not just in terms of the simpler operations S , + , and · .) The exponentiation symbol ∧ , more discussion of which will follow shortly, is *not* bounded.

With our discussion of bounded induction, we have completed the description of $\mathcal{Q}^0$ except for a large number of defined symbols -- some bounded, some unbounded. Let us use $\mathcal{Q}^0_b$ to denote the theory obtained when the unbounded symbols and their defining axioms are removed from $\mathcal{Q}^0$. We shall have occasion to use

Metatheorem E. Let $\mathbb{C}$ be a unary formula in an extension T of $\mathcal{Q}^0_b$. Assume that $\mathbb{C}$ is hereditary and respects 0 , S , + , · , and # . Then $\mathbb{C}$ respects every function symbol of $\mathcal{Q}^0_b$. Moreover, if $\mathbb{A}$ is a nonlogical axiom of $\mathcal{Q}^0_b$, then $\mathbb{A}^{\mathbb{C}}$ is a theorem of T .

(In particular, $\mathbb{A}$ could be the defining axiom of a bounded function symbol of $\mathcal{Q}^0$; it follows that such a symbol does not change its meaning if its defining axiom is relativized by $\mathbb{C}$. Metatheorem E is an easy extension of Metatheorem 7 in [1,§15].)

Sets, functions, and sequences

There is introduced in [1,§10] a procedure, the details of which need not concern us here, whereby a finite set of numbers can be encoded as a single number. This allows the definition of a unary predicate symbol indicating that a number is (i.e., encodes) a set. There is also a binary predicate symbol ϵ , and the extensionality property

1.11) a and b are sets & $\forall x(x \epsilon a \longleftrightarrow x \epsilon b) \longrightarrow a = b$

is a theorem of Q^0 . Both of these predicate symbols are bounded, as are several other defined symbols involving sets: $\cap$, $\cup$, $\subseteq$, { } (a unary function symbol: {x} is the set whose only member is x) , Card , and Bd (if a is a set, then Card a is the number of elements in a and Bd a is its largest element). Two useful theorems [1, (10.28) and (20.5)] are

1.12) $$x \in a \longrightarrow x < a$$

and

1.13) $$\text{Card } a \leq \text{Log } a.$$

Hence a bound on a automatically implies a bound on Bd a and a logarithmic bound on Card a . Conversely, if showing a formula to be bounded requires establishing a bound on a variable $\underline{x}$ that is always to designate a set, then it is sufficient to obtain a bound on Bd $\underline{x}$ and a logarithmic bound on Card $\underline{x}$. Not just any bound on Card $\underline{x}$ will do: a fact that will prove very useful later is that, in the absence of exponentiation, there may well be numbers larger than all logarithms!

The empty set, conveniently, is the number 0 . It is also sometimes convenient to know that the number 1 is not a set at all.

One should be cautioned against using sets that have not been shown to exist. There is no guarantee that an arbitrary set has a power set or that, given n , there is a set of all numbers from 0 to n . Above all, a "subclass" of a set may fail to be a set; that is, if a is a set and $\mathbb{A}[x]$ is a formula, those elements x of a such that $\mathbb{A}[x]$ holds may not form a set. There is, however, a principle of "bounded separation" [1,§11] that allows formation of the set $\{x \in a: \mathbb{A}[x]\}$ if $\mathbb{A}$ is a *bounded* formula. There is also a bounded unary function symbol Setlog that gives, for each n, the set of all numbers not exceeding Log n.

Ordered pairs are defined in [1] by the formula

1.14) $$\text{Def } \langle x,y\rangle = (x+y)\cdot(x+y) + y\,.$$

The definition satisfies the usual property

1.15) $$\langle x_1,y_1\rangle = \langle x_2,y_2\rangle \longrightarrow x_1 = x_2 \& y_1 = y_2$$

and also the convenient relations

1.16) $$x \le \langle x,y\rangle \& y \le \langle x,y\rangle$$

and

1.17) $$x_1 \le x_2 \& y_1 \le y_2 \longrightarrow \langle x_1,y_1\rangle \le \langle x_2,y_2\rangle\,.$$

The function symbol $\langle\ ,\ \rangle$ is bounded, as are the symbols Proj_1 $(\text{Proj}_1\langle x,y\rangle = x)$, Proj_2 $(\text{Proj}_2\langle x,y\rangle = y)$, and $\times$ (cartesian product).

There follows the usual definition of a function as a set of ordered pairs. A binary function symbol $\cdot(\cdot)$ is introduced, allowing the notation $f(x) = y$. For every function f there are sets Dom f and Ran f, satisfying a few obvious theorems. All of these defined symbols are bounded.

Let $\mathbb{D}[x,y]$ be a formula of Q^0 (possibly containing free variables besides x and y). If for every element x of a set a there exists some y such that $\mathbb{D}[x,y]$, then one might expect there to be a function f such that Dom f = a and $\forall x(x\in a \longrightarrow \mathbb{D}[x,f(x)])$. Again, the crucial property turns out to be boundedness -- this time boundedness of the formula $\exists y\mathbb{D}[x,y]$. The "bounded replacement principle" [1,§17] asserts that under this condition such a function exists; in fact, one such function is $\{\langle x,y\rangle: x\in a \,\&\, \min_y \mathbb{D}[x,y]\}$.

A *sequence* is defined as a function whose domain is the set of numbers from 1 to n for some n. (That such a set may not exist for every n is of no consequence as far as this definition is concerned.) Every sequence u has a length Ln u (possibly 0: the empty set 0 is a sequence!) satisfying the suggestive-looking inequality

1.18) $$\text{Ln } u \leq \text{Log } u ,$$

and also a largest term Sup u (the same as Bd Ran u). Corresponding to our earlier comment about sets is the fact [1,§19] that for the purpose of showing a formula to be bounded, establishing a bound on a sequence u is tantamount to establishing a bound on Sup u and a logarithmic bound on Ln u.

Let u and v be sequences. The relation sum(u,v) means that Ln u = Ln v and u(1) = v(1) and $\forall i(1\leq i<\text{Ln } u \longrightarrow v(i+1) = v(i)+u(i+1))$; in other words, v is the sequence of partial sums of u. For every sequence u there is a unique v such that sum(u,v) (this is a good exercise in the use of (BLNP)); this v is denoted $\sum u$. Note that $\sum u$ is a sequence, and that the sum of all the terms in u is the number $(\sum u)(\text{Ln } u)$. The notations prod (u,v) and Πu are defined similarly.

The juxtaposition u*v is the sequence whose length is Ln u+Ln v and whose terms are the terms of u followed by the terms of v . If s is a sequence all of whose terms are sequences, then s^* is the juxtaposition of all of those sequences; Ln s^* is the sum of the lengths of the sequences in s . If $1 \le i \le j \le \text{Ln } u$, then u[i,j] is the sequence that lists the terms of u from u(i) to u(j) ; its length is j-i+1 . If a is a set, then Enum a is the sequence that enumerates the elements of a in increasing order. The symbols · is a sequence, Ln , Sup, sum , $\sum$, prod , Π , ·*· , $\cdot^*$, ·[·,·] , and Enum, as well as a few others, are described fully in [1,§§19-20]; in particular, all are shown to be bounded.

Exponentiation

In [1,§13] appears the following definition of a bounded predicate symbol:

1.19) Def exp(x,k,f) <—> f is a function &∀i(i∈Dom f <—> i≤k)&

f(0) = 1 &∀i(i<k <—> f(i+1) = x·f(i)).

If exp(x,k,f) , then f(k) is the number we like to think of as x^k . The usual laws of exponents, albeit in rather unattractive forms, are easily established for exponential functions of this sort.

Now consider the definition

1.20) $$\text{Def } \varepsilon(k) \longleftrightarrow \exists f\ \exp(2,k,f)$$

of a presumably unbounded predicate symbol. Then $\varepsilon(k)$ essentially asserts the existence of a sequence whose terms are the first k powers of 2 . If $k = \text{Log } n$ for some n , then such a sequence can be formed; in fact, $\varepsilon(k)$ holds if and only if $\exists n(k = \text{Log } n)$ (or equivalently $\exists n(k \leq \text{Log } n))$. The unary formula $\varepsilon(k)$ is hereditary. Moreover, from the theorems $\text{Log } (2\cdot n) = S(\text{Log } n)$, $\text{Log } (m\cdot n) \geq \text{Log } m+\text{Log } n$, and $\text{Log } (m\#n) = \text{Log } m\cdot\text{Log } n$, it follows that ε respects S , $+$, and $\cdot$. (The last of these assertions is another example of the importance of $\#$. Note, though, that we cannot at this point prove that ε respects $\#$.) Writing $\varepsilon^3(k)$ for $\mathbb{C}^3[k]$, where $\mathbb{C}[k]$ is the unary formula $\varepsilon(k)$, we therefore have the theorem $\varepsilon^3(k) \longleftrightarrow \varepsilon(k)$.

If $x > 2$, then the existence of an f such that $\exp(x,k,f)$ is equivalent to the existence of a g such that $\exp\ (2,k\cdot(\text{Log } x+1),g)$, which just means $\varepsilon(k\cdot(\text{Log } x+1))$, or just $\varepsilon(k)$. Hence we have the alternative definition $\varepsilon(k) \longleftrightarrow \forall x\exists f\ \exp\ (x,k,f)$; it is this form of the definition that appears in [1,§14]. We may regard $\varepsilon(k)$ as the statement "k is exponentiable"; the base is irrelevant, but 2 is generally the most convenient choice.

The way to define the function symbol $\wedge$ for exponentiation should now be clear.

1.21) $$\text{Def } x\wedge k = z \longleftrightarrow \exists f(\exp(x,k,f)\&f(k) = z)\text{ , otherwise } z = 0 .$$

This is our first encounter with the "otherwise" notation, and some explanation is required. If there exists a z such that $\exists f(\exp(x,k,f) \& f(k) = z)$, then $x \wedge k$ is that z; if there is no such z, then $x \wedge k$ is 0. In definitions of this kind, the existence condition is automatic, but the uniqueness condition must still be verified. Note that in this instance the "otherwise" clause comes into play precisely if $\neg\varepsilon(k)$ (even if x is 1).

Often we will write x^k rather than $x \wedge k$; at other times the notation $x \wedge k$ will be clearer.

The symbol $\wedge$, like ε, is unbounded. The unboundedness lies in the fact that it is not clear, given k, how big an f such that $\exp(2,k,f)$ must be. If $k = \mathrm{Log}\ n$, then it can be shown [1,(16.32)] that $f \leq 18250\#(2\cdot n)\#(2\cdot n)$, but this is not sufficient to make ε or $\wedge$ bounded since the bound is in terms of n rather than in terms of k. If the symbol ε *were* bounded, then, since $\varepsilon(k)$ is inductive in k, we could conclude $\forall k\varepsilon(k)$ by bounded induction; Nelson gives an argument to show, however, that $\forall k\varepsilon(k)$ is *not* a theorem of Q^0. (More about this later: we shall eventually strengthen our theory by postulating the existence of a number N such that $\neg\varepsilon(N)$.) In any case, there is a bounded function symbol Explog with the property that $\mathrm{Explog}(x,k) = x \wedge \mathrm{Log}\ k$; if $x \neq 0$, then $\mathrm{Explog}(x,k)$ is never 0, since $\varepsilon(\mathrm{Log}\ k)$ always holds. (Note, incidentally, that by (1.13) and (1.18) we have $\varepsilon(\mathrm{Card}\ a)$ and $\varepsilon(\mathrm{Ln}\ u)$ also. It is a theorem of Q^0 that $\varepsilon(n)$ holds if and only if there is a set of all numbers not exceeding n.)

This completes our summary of the important features of the theory $\mathcal{Q}^0$. To be precise, we declare that $\mathcal{Q}^0$ is Nelson's theory $\mathcal{Q}_9$ -- the theory constructed in the first 20 sections of [1]. This is a provably consistent theory, and it is from here that we shall now embark on our voyage through predicative mathematics.

§2. *Hypersmashes and Higher Relativization Schemes*

It seems reasonable to ask whether the chain of function symbols 0 , S , $+$, $\cdot$, $\#$ can be extended in a natural way. This section is devoted to answering that question in the affirmative. The objects of our investigation will be binary function symbols $\#_1, \#_2, \ldots, \#_\mu, \ldots$ with the property that $2^k \#_1 2^\ell = 2^{k\#\ell}$, $2^k \#_2 2^\ell = 2^{k\#_1\ell}, \ldots, 2^k \#_\mu 2^\ell = 2^{k\#_{\mu-1}\ell}, \ldots$. (No claim is made about a ternary operation $x\#_n y$.) We shall see that, in a suitable extension Q^μ of Q^0 , the symbols $\#_1, \ldots, \#_\mu$ may be regarded as bounded.

Let us first note a reason why having these symbols at our disposal will be advantageous.

The problem

Consider the definitions

2.1)
$$\text{Def } \varepsilon_1(k) \longleftrightarrow \varepsilon(k) \& \varepsilon(2^k) ,$$
$$\text{Def } \varepsilon_2(k) \longleftrightarrow \varepsilon_1(k) \& \varepsilon_1(2^k) ,$$
$$\vdots$$
$$\text{Def } \varepsilon_\mu(k) \longleftrightarrow \varepsilon_{\mu-1}(k) \& \varepsilon_{\mu-1}(2^k) ,$$
$$\vdots$$

Then $\varepsilon_1(k)$ asserts that k is twice exponentiable, $\varepsilon_2(k)$ asserts that k is three times exponentiable,... . To nip confusion in the bud, we observe that the three dots do *not* conceal an induction. The subscripts $1, 2, \ldots, \mu, \ldots$ are (in Nelson's words) "genetic" rather than "formal"; we have not defined a binary relation $\varepsilon_n(k)$, nor do we even claim to have defined $\varepsilon_\mu(k)$ "for all μ" (whatever *that* would mean). We have simply shown the reader how to write down definitions of $\varepsilon_\mu(k)$ for as many μ as he likes. In practice, very small μ will suffice -- maybe $\mu = 25$ or even $\mu = 2$.

It is clear that every ε_μ is hereditary. We would like to know that every ε_μ respects 0 , S , $+$, and $\cdot$, as ε does. But the assertion that ε_μ respects $\cdot$ is equivalent to

$$\varepsilon_{\mu-1}(k)\&\varepsilon_{\mu-1}(\ell)\&\varepsilon_{\mu-1}(2^k)\&\varepsilon_{\mu-1}(2^\ell) \longrightarrow \varepsilon_{\mu-1}(k\cdot\ell)\&\varepsilon_{\mu-1}(2^k\#2^\ell) ,$$

and proving this requires showing that $\varepsilon_{\mu-1}$ respects $\#$ -- a definite problem, since we do not know even that ε itself respects $\#$. We can *make* ε respect $\#$, however, by introducing the symbol $\#_1$, thereby obtaining the theorem

$$\text{Log } x\#\text{Log } y = \text{Log } (x\#_1 y) ,$$

which suffices for the proof that ε_1 respects multiplication. The corresponding results for $\varepsilon_2,\varepsilon_3,\ldots,\varepsilon_\mu,\ldots$ require the symbols $\#_2,\#_3,\ldots,\#_\mu,\ldots$.

The reader who is willing to accept the fact that this program can be carried out may skip the remainder of §2.

Axioms for $\#_1$

We shall adjoin $\#_1$ to the theory Q^0 in a manner strongly reminiscent of the way in which Nelson adjoins $\#$ to the theory involving only 0 , S , $+$, and $\cdot$ [1,§§14-15].

Write $\varepsilon^4(k)$ for $\mathbb{E}^4[k]$, where $\mathbb{E}[k]$ is the unary formula $\varepsilon(k)$. Showing that ε respects # is tantamount to proving the theorem $\varepsilon^4(k) \longleftrightarrow \varepsilon(k)$. In any case, ε is inductive, so by Metatheorem A the following is a theorem of $\mathcal{Q}^0$:

2.2) $(\varepsilon^4(k) \longrightarrow \varepsilon(k))\&(\varepsilon^4(k)\&i\leq k \longrightarrow \varepsilon^4(i))\&\varepsilon^4(0)\&\ (\varepsilon^4(k) \longrightarrow \varepsilon^4(Sk))\&$

$(\varepsilon^4(k)\&\varepsilon^4(\ell) \longrightarrow \varepsilon^4(k+\ell))\&(\varepsilon^4(k)\&\varepsilon^4(\ell) \longrightarrow \varepsilon^4(k\cdot\ell))\&$

$(\varepsilon^4(k)\&\varepsilon^4(\ell) \longrightarrow \varepsilon^4(k\#\ell))$. ||

2.3) Def $x\wedge_1 k = z \longleftrightarrow \varepsilon^4(k)\&\exists f(\exp(x,k,f)\&f(k) = z)$, otherwise $z = 0$.

2.4) $$\varepsilon^4(k) \longrightarrow x\wedge_1 k = x\wedge k\ .\ ||$$

2.5) $\varepsilon^4(k)\&\varepsilon^4(\ell) \longrightarrow (x\cdot y)\wedge_1 k = (x\wedge_1 k)\cdot(y\wedge_1 k)\&x\wedge_1(k+\ell) = (x\wedge_1 k)\cdot(x\wedge_1 \ell)\&$

$x\wedge_1(k\cdot\ell) = (x\wedge_1 k)\wedge_1 \ell\&(2\wedge_1 k)\#(2\wedge_1 \ell) = 2\wedge_1(k\cdot\ell)$.

Proof. By (2.2) and (2.4), together with basic properties of $\wedge$ and # . ||

2.6) $$\text{Def } \lambda_1(x) \longleftrightarrow \exists k\ x \leq 2\wedge_1 k\ .$$

2.7) $(\lambda_1(x)\&w\leq x \longrightarrow \lambda_1(w))\&\lambda_1(0)\&(\lambda_1(x) \longrightarrow \lambda_1(Sx))\&(\lambda_1(x)\&$

$\lambda_1(y) \longrightarrow \lambda_1(x+y))\&(\lambda_1(x)\&\lambda_1(y) \longrightarrow \lambda_1(x\cdot y))\&(\lambda_1(x)\&\lambda_1(y) \longrightarrow$

$\lambda_1(x\#y))$.

Proof. The first two conjuncts are immediate. If $x \leq 2\wedge_1 k$ and $y \leq 2\wedge_1 \ell$, then $Sx \leq 2\wedge_1(k+1)$, $x+y \leq 2\wedge_1$ (Max (k,ℓ) +1) , $x\cdot y \leq 2\wedge_1(k+\ell)$, and $x\#y \leq 2\wedge_1(k\cdot\ell)$. ||

Of course $x \leq 2\wedge(\text{Log } x+1)$ always holds; since there is no guarantee that $\varepsilon^4(\text{Log } x)$ holds, however, $\lambda_1(x)$ may still be false. On the other hand, when we eventually prove (in a stronger theory than Q^0) that $\varepsilon^4(k) \longleftrightarrow \varepsilon(k)$, it will follow immediately that $\wedge_1$ is the same as $\wedge$ and that λ_1 holds universally. Here is one more soon-to-be-uninteresting definition:

2.8) Def $\text{Log}_1 x = k \longleftrightarrow |x|_2 = 2\wedge_1 k$, otherwise $k = 0$.

2.9) $$\varepsilon^4(\text{Log}_1 x) \text{ . } \|$$

2.10) $$\lambda_1(x) \longleftrightarrow \varepsilon^4(\text{Log } x) \text{ .}$$

Proof. The formula $\lambda_1(x)$ is true precisely if $x \leq 2\wedge k$ for some k such that $\varepsilon^4(k)$. If this is the case, then $k \geq \text{Log } x$, so $\varepsilon^4(\text{Log } x)$ by (2.2). Conversely, if $\varepsilon^4(\text{Log } x)$, then $\varepsilon^4(\text{Log } x+1)$ by (2.2), and certainly $x \leq 2\wedge(\text{Log } x+1)$. ||

2.11) $$\lambda_1(x) \longrightarrow \text{Log}_1 x = \text{Log } x \,\&\, |x|_2 = 2\wedge_1 \text{Log}_1 x \text{ .}$$

Proof. If $\lambda_1(x)$, then $\varepsilon^4(\text{Log } x)$ by (2.10), so (2.4) gives $|x|_2 = 2\wedge\text{Log } x = 2\wedge_1\text{Log } x$, as required by the definition (2.8). ||

2.12) $$\lambda_1(2\wedge_1 k) \text{ . } \|$$

Now for an "interesting" defining axiom:

2.13) $$\text{Def } x\#^1 y = 2\wedge_1(\mathrm{Log}_1 x\#\mathrm{Log}_1 y) .$$

2.14) $$x\#^1 y = 2\wedge(\mathrm{Log}_1 x\#\mathrm{Log}_1 y) .$$

Proof. By (2.13) and (2.4), it suffices to prove $\varepsilon^4(\mathrm{Log}_1 x\#\mathrm{Log}_1 y)$; this follows from (2.9) and (2.2). ||

2.15) $$\lambda_1(x\#^1 y) .$$

Proof. By (2.13) and (2.12). ||

2.16) $$x\#^1 y = |x\#^1 y|_2 .$$

Proof. By (2.14) and $\varepsilon(\mathrm{Log}_1 x\#\mathrm{Log}_1 y)$. ||

2.17) $$\mathrm{Log}_1(x\#^1 y) = \mathrm{Log}_1 x\#\mathrm{Log}_1 y .$$

Proof. By (2.16), (2.13), and (2.8). ||

2.18) $$\lambda_1(x)\&\lambda_1(y) \longrightarrow \mathrm{Log}(x\#^1 y) = \mathrm{Log}\ x\#\mathrm{Log}\ y .$$

Proof. By (2.15), (2.11), and (2.17). ||

Let $\overline{Q}^0$ be the extension by definitions of Q^0 obtained by adjoining $\wedge_1$, λ_1 , Log_1 , and $\#^1$ as above. Let $\hat{Q}^0$ be the theory obtained from Q^0 by adjoining a new binary function symbol $\#_1$ and the nonlogical axioms (2.19) and (2.20):

2.19) $$\mathrm{Ax}\ x\#_1 y = |x\#_1 y|_2\ ;$$

2.20) $$\mathrm{Ax\ Log}\ x\#_1 y = \mathrm{Log}\ x\#\mathrm{Log}\ y\ .$$

Let $\hat{Q}^0_b$ be obtained by removing from $\hat{Q}^0$ all unbounded defined symbols of Q^0 and their defining axioms (or equivalently by adjoining to Q^0_b the symbol $\#_1$ and the axioms (2.19) and (2.20)). Then $\hat{Q}^0$ is an extension by definitions of $\hat{Q}^0_b$, and we shall be free to work in $\hat{Q}^0$ as soon as we have proved

Metatheorem F_1. The theory $\hat{Q}^0_b$ is interpretable in Q^0.

Proof. We exhibit an interpretation I of $\hat{Q}^0_b$ in the extension by definitions $\overline{Q}^0$ of Q^0. Let the universe of I be λ_1. For each nonlogical (function or predicate) symbol $\underline{u}$ of Q^0_b (that is, for each $\underline{u}$ in $\hat{Q}^0_b$ other than $\#_1$), let $\underline{u}_I$ be $\underline{u}$, and let $(\#_1)_I$ be $\#^1$. By (2.7) and Metatheorem E, the interpreting formula λ_1 respects every function symbol of Q^0_b, and moreover $\mathbb{A}^I$ is a theorem of $\overline{Q}^0$ for every nonlogical axiom $\mathbb{A}$ of Q^0_b. By (2.15), λ_1 respects $(\#_1)_I$ in $\overline{Q}^0$, and the interpretations of axioms (2.19) and (2.20) hold by (2.16) and (2.18) (the function symbols $|\ |_2$, Log, and $\#$ being bounded). ||

It follows from (2.20) that ε respects $\#$. Hence it is a theorem of $\hat{Q}^0$ that $\varepsilon^4(k) \longleftrightarrow \varepsilon(k)$. If we extend $\hat{Q}^0$ by adjoining the defined symbols of $\overline{Q}^0$, it is then immediate that $x\wedge_1 k = x\wedge k$, that $\lambda_1(x)$ is equivalent to $\exists k x \leq 2\wedge k$ and therefore holds universally, and that Log_1 has the same meaning as Log . Furthermore, (2.19) and (2.20) together imply that $x\#_1 y = 2\wedge(\mathrm{Log}\ x\#\mathrm{Log}\ y)$; comparing (2.13), we see that $x\#_1 y = x\#^1 y$. Hereafter we shall always write $\#_1$.

Some hypersmash-arithmetic

Let us prove a few theorems in $\hat{Q}^0$. We are free, of course, to use $x\#_1 y = 2\wedge(\mathrm{Log}\ x\#\mathrm{Log}\ y)$.

2.21) $$x\#_1 0 = 2 .$$

Proof. This follows from $x\#0 = 1$: we have

$$x\#_1 0 = 2\wedge(\mathrm{Log}\ x\#0) = 2\wedge 1 = 2 . \quad \|$$

2.22) $$x\#_1 4 \leq x .$$

Proof. This follows from $x\#2 \leq x$: we have

$$x\#_1 4 = 2^k \text{ , where } k = \mathrm{Log}\ x\#2 \leq \mathrm{Log}\ x . \quad \|$$

2.23) $$x < (x\#_1 4)\#4 .$$

Proof. This follows from $x < (x\#2)\cdot 2$: we have

$$\mathrm{Log}\ ((x\#_1 4)\#4) = \mathrm{Log}\ (x\#_1 4)\cdot\mathrm{Log}\ 4 = (\mathrm{Log}\ x\#\mathrm{Log}\ 4)\cdot\mathrm{Log}\ 4$$

$$= (\mathrm{Log}\ x\#2)\cdot 2 > \mathrm{Log}\ x \text{ , whence (2.23).} \quad \|$$

The next three propositions follow from the corresponding facts about # .

2.24) $$x\#_1 y = y\#_1 x . \quad \|$$

2.25) $$x\#_1(y\#_1 z) = (x\#_1 y)\#_1 z \ . \ \|$$

2.26) $$y \leq z \longrightarrow x\#_1 y \leq x\#_1 z \ . \|$$

2.27) $$\text{Log Log } y = \text{Log Log } z \longrightarrow x\#_1 y = x\#_1 z \ .$$

Proof. This follows from $\text{Log } y = \text{Log } z \longrightarrow x\#y = x\#z$. Replacing x , y , and z by $\text{Log } x$, $\text{Log } y$, and $\text{Log } z$ in that theorem, we see that $\text{Log Log } y = \text{Log Log } z$ implies $\text{Log } x\#\text{Log } y = \text{Log } x\#\text{Log } z$; exponentiating then gives $x\#_1 y = x\#_1 z$. ||

2.28) $$\text{Log } (x\#_1 y) = |\text{Log}(x\#_1 y)|_2 \ .$$

Proof. This follows from $x\#y = |x\#y|_2$ and (2.20). ||

So a hypersmash is more than just a power of 2 : it is a power of 2 with an exponent that it also a power of 2 . If we start with $x\#y$, we can apply Log , then exponentiate, and get $x\#y$ back. If we start with $x\#_1 y$, we can apply Log *twice*, exponentiate twice, and get $x\#_1 y$ back.

2.29) $$x\#_1(y\#z) \leq (x\#_1 y)\#(x\#_1 z)\#x \ .$$

Proof. Let us first establish the lower-level analog $x\#(y\cdot z) \leq (x\#y)\cdot(x\#z)\cdot x$. We have

$$\text{Log } (x\#(y\cdot z)) = \text{Log } x\cdot\text{Log } (y\cdot z)$$

$$\leq \text{Log } x\cdot(\text{Log } y+\text{Log } z+1) = \text{Log } x\cdot\text{Log } y+\text{Log } x\cdot\text{Log } z+\text{Log } x \ .$$

Exponentiating the left side gives exactly $x\#(y\cdot z)$; exponentiating the right side gives at most $(x\#y)\cdot(x\#z)\cdot x$.

We now use this result to prove (2.29). We have

$$\mathrm{Log}\ (x\#_1(y\#z)) = \mathrm{Log}\ x\#\mathrm{Log}\ (y\#z) = \mathrm{Log}\ x\#(\mathrm{Log}\ y\cdot\mathrm{Log}\ z)$$

$$\leq (\mathrm{Log}\ x\#\mathrm{Log}\ y)\cdot(\mathrm{Log}\ x\#\mathrm{Log}\ z)\cdot\mathrm{Log}\ x\ .$$

Exponentiating gives (2.29). ||

Finally, a result about $\#$:

2.30) $$x \geq 8 \ \&\ y \geq 8 \longrightarrow x\cdot y < x\#y\ .$$

Proof. Note that $x \geq 3 \ \&\ y \geq 3 \longrightarrow x+y+1 < x\cdot y$. Therefore, if $x \geq 8$ and $y \geq 8$, then

$\mathrm{Log}\ (x\cdot y) \leq \mathrm{Log}\ x+\mathrm{Log}\ y+1 < \mathrm{Log}\ x\cdot\mathrm{Log}\ y = \mathrm{Log}\ (x\#y)$. ||

Induction on $\#_1$

Our work with $\#_1$ is almost complete. The one remaining order of business is the construction of a theory Q^1 in which $\#_1$ may be regarded as a bounded symbol in the most important way -- that is, in which we may apply induction on bounded formulas involving $\#_1$. To this end, we make the expected definition

$$\mathbb{C}^5[x] \ \text{ for } \ \forall y(\mathbb{C}^4[y] \longrightarrow \mathbb{C}^4[y\#_1 x])\ ;$$

here $\mathbb{C}[x]$ is a unary formula, and $\mathbb{C}^4[x]$ is as defined in §1 . The analog of Metatheorem A in this situation is

Metatheorem G_1 . Let T be an extension of the theory $\hat{Q}^0_b$, and let $\mathbb{C}$ be a unary formula in the language of T . Then the following is a theorem of T :

$$\mathbb{C}[0]\,\&\,\forall x(\mathbb{C}[x] \longrightarrow \mathbb{C}[Sx]) \longrightarrow$$

$$(\mathbb{C}^5[x] \longrightarrow \mathbb{C}[x]\,\&$$

$$(\mathbb{C}^5[x]\,\&\,w \leq x \longrightarrow \mathbb{C}^5[w])\,\&$$

$$\mathbb{C}^5[0]\,\&$$

$$(\mathbb{C}^5[x] \longrightarrow \mathbb{C}^5[Sx])\,\&$$

$$(\mathbb{C}^5[x_1]\,\&\,\mathbb{C}^5[x_2] \longrightarrow \mathbb{C}^5[x_1+x_2])\,\&$$

$$(\mathbb{C}^5[x_1]\,\&\,\mathbb{C}^5[x_2] \longrightarrow \mathbb{C}^5[x_1\cdot x_2])\,\&$$

$$(\mathbb{C}^5[x_1]\,\&\,\mathbb{C}^5[x_2] \longrightarrow \mathbb{C}^5[x_1\# x_2])\,\&$$

$$(\mathbb{C}^5[x_1]\,\&\,\mathbb{C}^5[x_2] \longrightarrow \mathbb{C}^5[x_1\#_1 x_2])\ .$$

Proof. We may work in an extension by definitions of T containing all the symbols of $\hat{Q}^0$, and we assume throughout that $\mathbb{C}[0]\&\forall x(\mathbb{C}[x] \longrightarrow \mathbb{C}[Sx])$. Our basic tools are (2.21)-(2.30) and Metatheorem A.

Suppose $\mathbb{C}^5[x]$. Since $\mathbb{C}^4$ is inductive, we have $\mathbb{C}^4[4]$, whence $\mathbb{C}^4[x\#_1 4]$ by the definition of $\mathbb{C}^5$ and (2.24). Since $\mathbb{C}^4$ respects $\#$, $\mathbb{C}^4[(x\#_1 4)\#4]$ holds; since $\mathbb{C}^4$ is hereditary, by (2.23) we have $\mathbb{C}^4[x]$ and therefore $\mathbb{C}[x]$. Thus $\mathbb{C}^5[x] \longrightarrow \mathbb{C}[x]$.

Suppose $\mathbb{C}^5[x]\&w \leq x\&\mathbb{C}^4[y]$. Then $\mathbb{C}^4[y\#_1x]$. By (2.26), $y\#_1w \leq y\#_1x$; hence $\mathbb{C}^4[y\#_1w]$. Thus $\mathbb{C}^5[x]\&w \leq x \longrightarrow \mathbb{C}^5[w]$.

$\mathbb{C}^5[0]$ is immediate from (2.21) and $\mathbb{C}^4[2]$.

Now suppose $\mathbb{C}^5[x_1]\&\mathbb{C}^5[x_2]$. We shall complete the proof of Metatheorem G_1 by showing $\mathbb{C}^5(x_1\#_1x_2)$, $\mathbb{C}^5(x_1\#x_2)$, $\mathbb{C}^5(x_1\cdot x_2)$, $\mathbb{C}^5(x_1+x_2)$, and $\mathbb{C}^5[Sx_1]$.

If $\mathbb{C}^4[y]$, then $\mathbb{C}^4[y\#_1x_1]$ (by $\mathbb{C}^5[x_1]$) , so $\mathbb{C}^4[(y\#_1x_1)\#_1x_2]$ (by $\mathbb{C}^5[x_2]$) , so $\mathbb{C}^4[y\#_1(x_1\#_1x_2)]$ (by (2.25)). This shows $\mathbb{C}^5[x_1\#_1x_2]$.

If $\mathbb{C}^4[y]$, then $\mathbb{C}^4[y\#_1x_1]$ and $\mathbb{C}^4[y\#_1x_2]$, so $\mathbb{C}^4[(y\#_1x_1)\#(y\#_1x_2)\#y]$. By (2.29), $y\#_1(x_1\#x_2) \leq (y\#_1x_1)\#(y\#_1x_2)\#y$, so $\mathbb{C}^4[y\#_1(x_1\#x_2)]$. This shows $\mathbb{C}^5[x_1\#x_2]$.

By (2.27) and (2.22), we have $y\#_18 = y\#_14 \leq y$. Hence $\mathbb{C}^4[y] \longrightarrow \mathbb{C}^4[y\#_18]$, or in other words $\mathbb{C}^5[8]$. Define $Fx = \text{Max}\ (x,8)$; then $\mathbb{C}^5[x] \longrightarrow \mathbb{C}^5[Fx]$. Still under the assumption that $\mathbb{C}^5[x_1]\&\mathbb{C}^5[x_2]$, we therefore have (by what we have already proved) $\mathbb{C}^5[Fx_1\#Fx_2]$. Using $Fx_1 \geq 8$, $Fx_2 \geq 8$, (2.30) , and the knowledge that $\mathbb{C}^5$ is hereditary, we get

$$x_1 \cdot x_2 \le Fx_1 \cdot Fx_2 \le Fx_1 \# Fx_2 \text{ , so } \mathbb{C}^5[x_1 \cdot x_2] \text{ ;}$$

$$x_1 + x_2 \le Fx_1 + Fx_2 \le Fx_1 \cdot Fx_2 \le Fx_1 \# Fx_2 \text{ , so } \mathbb{C}^5[x_1 + x_2] \text{ ;}$$

and $$Sx_1 \le SFx_1 \le Fx_1 + Fx_2 \le Fx_1 \cdot Fx_2 \le Fx_1 \# Fx_2 \text{ , so } \mathbb{C}^5[Sx_1] \text{ .}$$

Thus $\mathbb{C}^5[x_1] \& \mathbb{C}^5[x_2] \longrightarrow \mathbb{C}^5[Sx_1] \& \mathbb{C}^5[x_1+x_2] \& \mathbb{C}^5[x_1 \cdot x_2] \& \mathbb{C}^5[x_1 \# x_2] \& \mathbb{C}^5[x_1 \#_1 x_2]$. ||

Let Q^1 be the theory obtained from $\hat{Q}^0$ by adjoining as new nonlogical axioms all formulas of the form (MBI) (see §1) with $\mathbb{A}$ a manifestly bounded formula in the language of $\hat{Q}^0_b$. Let Q^1_b be obtained from Q^1 by removing all unbounded symbols of Q^0 together with their defining axioms; note that the new axioms (MBI) do not involve any of these symbols.

Metatheorem H_1 . Let T be a theory containing all the symbols of $\hat{Q}^0_b$. Let $\mathbb{C}$ be a unary formula of T that is hereditary and respects 0 , S , $+$, $\cdot$, $\#$, and $\#_1$. Then $\mathbb{C}$ respects every function symbol of $\hat{Q}^0_b$. Moreover, if T is an extension of $\hat{Q}^0_b$ and $\mathbb{A}$ is a nonlogical axiom of $\hat{Q}^0_b$, or if T is an extension of Q^1_b and $\mathbb{A}$ is a nonlogical axiom of Q^1_b , then $\mathbb{A}^{\mathbb{C}}$ is a theorem of T .

Proof. By Metatheorem E , $\mathbb{C}$ respects every function sumbol of Q^0_b ; by hypothesis, $\mathbb{C}$ respects $\#_1$ also. The last assertion as well follows from Metatheorem E if $\mathbb{A}$ is a nonlogical axiom of Q^0_b . If $\mathbb{A}$ is (2.19), (2.20), or one of the new axioms (MBI), then $\mathbb{A}$ is manifestly bounded; letting T' be $\hat{Q}^0_b$ in Metatheorem B , we see that $\mathbb{A} \longrightarrow \mathbb{A}^{\mathbb{C}}$ (that is, $\mathbb{A} \longrightarrow (\mathbb{C}(\text{free } \mathbb{A}) \longrightarrow \mathbb{A}_{\mathbb{C}})$) is a theorem of T

as long as $\mathcal{T}$ is an extension of $\hat{\mathcal{Q}}_b^0$. Therefore if $\mathbb{A}$ is a theorem of $\mathcal{T}$, so is $\mathbb{A}^{\mathbb{C}}$. ||

Metatheorem I_1. Let $\mathbb{A}$ be a manifestly bounded formula of $\hat{\mathcal{Q}}_b^0$ that is inductive in one of its free variables. Then $\hat{\mathcal{Q}}_b^0[\mathbb{A}]$ is interpretable in $\hat{\mathcal{Q}}_b^0$.

Proof. Let $\mathbb{A}$ be inductive in $\underline{x}$, and let $\mathbb{C}[\underline{x}]$ be the unary formula obtained from $\mathbb{A}$ by appending to the front a universal quantifier on each free variable other than $\underline{x}$. Then $\mathbb{C}[\underline{x}]$ is inductive in $\underline{x}$. By Metatheorem G_1, $\mathbb{C}^5$ is stronger than $\mathbb{C}$, is hereditary, and respects 0, S, $+$, $\cdot$, $\#$, and $\#_1$. By Metatheorem H_1, $\mathbb{C}^5$ respects every function symbol of $\hat{\mathcal{Q}}_b^0$, and moreover the relativization by $\mathbb{C}^5$ of every nonlogical axiom of $\hat{\mathcal{Q}}_b^0$ is a theorem of $\hat{\mathcal{Q}}_b^0$. Since $\mathbb{C}^5[\underline{x}] \longrightarrow \mathbb{A}$ is a theorem of $\hat{\mathcal{Q}}_b^0$, it follows from Metatheorem B that $\mathbb{A}^{\mathbb{C}^5}$ is a theorem of $\hat{\mathcal{Q}}_b^0$. Hence $\mathbb{C}^5$ determines an interpretation of $\hat{\mathcal{Q}}_b^0[\mathbb{A}]$ in $\hat{\mathcal{Q}}_b^0$. ||

Metatheorem J_1. Let $\mathbb{B}_1, \ldots, \mathbb{B}_\lambda$ be theorems of $\mathcal{Q}_b^1$. Then $\hat{\mathcal{Q}}_b^0[\mathbb{B}_1, \ldots, \mathbb{B}_\lambda]$ is interpretable in $\hat{\mathcal{Q}}_b^0$.

Proof. Each new axiom (MBI) is manifestly bounded and inductive in x in $\hat{\mathcal{Q}}_b^0$. Hence the conjunction $\mathbb{A}$ of all the new axioms (MBI) used in proving $\mathbb{B}_1, \ldots, \mathbb{B}_\lambda$ in $\mathcal{Q}_b^1$ is manifestly bounded and inductive in x in $\hat{\mathcal{Q}}_b^0$. By Metatheorem I_1, $\hat{\mathcal{Q}}_b^0[\mathbb{A}]$ is interpretable in $\hat{\mathcal{Q}}_b^0$. Now we need only observe that $\mathbb{B}_1, \ldots, \mathbb{B}_\lambda$ are theorems of $\hat{\mathcal{Q}}_b^0[\mathbb{A}]$ and apply the interpretation theorem [2,§4.7]. ||

Combining our various interpretability results, we see that if $\mathbb{B}_1,\ldots,\mathbb{B}_\lambda$ are theorems of an extension by definitions of Q^1 (itself an extension by definitions of Q^1_b), then the theory whose nonlogical axioms are $\mathbb{B}_1,\ldots,\mathbb{B}_\lambda$ is interpretable in Q^0, or even in Robinson's theory Q.

As in §1, it is a simple matter to check that every induction formula (Ind) in which $\mathbb{A}$ is a bounded formula of Q^1_b is a theorem of Q^1_b; in fact, the formula $\mathbb{A}$ may contain bounded symbols of some extension by definitions of Q^1_b. (Note that the bounds in a bounded formula may now involve $\#_1$.) We may also apply the bounded least number, bounded separation, and bounded replacement principles under corresponding conditions.

Higher smashes

From the foregoing discussion it should be clear how to proceed to further adjoin symbols $\#_2,\#_3,\ldots,\#_\mu,\ldots$. For $\#_2$, we begin working with ε^5 in Q^1; define $\wedge_2$, λ_2, Log_2, and $\#^2$ $(x\#^2y = 2\wedge_2(\mathrm{Log}_2 x\#_1\mathrm{Log}_2 y))$, forming the theory $\overline{Q}^1$; and establish the obvious analogs of (2.2)-(2.18). The theory $\hat{Q}^1$ consists of Q^1 together with the symbol $\#_2$ and the two axioms $x\#_2y = |x\#_2y|_2$ and $\mathrm{Log}\,(x\#_2y) = \mathrm{Log}\,x\#_1\mathrm{Log}\,y$. Using Metatheorem H_1 in place of Metatheorem E, we prove Metatheorem F_2, which asserts interpretability of $\hat{Q}^1_b$ ($\hat{Q}^1$ without the unbounded symbols of Q^0) in Q^1. We can prove in $\hat{Q}^1$ that ε respects $\#_1$, so that ε^5 is the same as ε, $\wedge_2$ the same as $\wedge$, λ_2 trivial, Log_2 the same as Log, and $\#_2$ the same as $\#^2$. We use (2.21)-(2.30) to establish their higher-level versions (for instance, (2.23) and (2.27) become

$x < (x\#_2 16)\#_1 16$ and $\text{Log Log Log } y = \text{Log Log Log } z \longrightarrow x\#_2 y = x\#_2 z)$, and use these results in proving Metatheorem G_2 . In the theory Q^2 we admit bounded induction on formulas containing $\#_2$; the remaining items on our level-2 agenda are Metatheorems H_2, I_2 , and J_2 , the last asserting that Q^2_b is finitely interpretable in $\hat{Q}^1_b$. At this point we are ready to go to work on $\#_3, \ldots$. The mimicry is sufficiently straightforward that only one point merits further attention: this concerns the numbers 2, 4, and 8 that appear in (2.21), (2.22)-(2.23), and (2.30) and in the proof of Metatheorem G_1. A little thought shows that these numbers will increase superexponentially with μ ; in particular, whereas for Metatheorem G_1 we used $\mathfrak{C}^4[4]$, in Metatheorem G_2 we shall need $\mathfrak{C}^5[16]$, and then $\mathfrak{C}^6[2^{16}]$, $\mathfrak{C}^7[2^{2^{16}}], \ldots$. Will this make our proofs unfeasible by the time we get to $\#_4$ or $\#_5$? Not at all: we have the theorems $16 = 4\#4$, $2^{16} = 16\#_1 16$, $2^{2^{16}} = 2^{16}\#_2 2^{16}, \ldots$, and we can use the knowledge that $\mathfrak{C}^5$ respects $\#$, $\mathfrak{C}^6$ respects $\#_1$, $\mathfrak{C}^7$ respects $\#_2, \ldots$.

The problem, revisited

Recall the definitions (2.1) of $\varepsilon_1, \varepsilon_2, \ldots$ and the discussion that motivated the introduction of $\#_1, \#_2, \ldots$. It is convenient to regard ε as ε_0 . Then for $\mu = 0,1,2,\ldots$, $\varepsilon_{\mu+1}(0)$ says that $\varepsilon_\mu(0)\&\varepsilon_\mu(1)$; hence if ε_μ respects 0 and S , then $\varepsilon_{\mu+1}$ repects 0 . Also, $\varepsilon_{\mu+1}(Sx)$ is $\varepsilon_\mu(Sx)\&\varepsilon_\mu(2^x+2^x)$; hence if ε_μ respects S and + , then $\varepsilon_{\mu+1}$ respects S . Next, $\varepsilon_{\mu+1}(x+y)$ is $\varepsilon_\mu(x+y)\&\varepsilon_\mu(2^x\cdot 2^y)$; hence if ε_μ respects + and $\cdot$, then $\varepsilon_{\mu+1}$ respects + . It is easy to see that if ε_μ respects $\cdot$ and $\#$

(respectively, $\#$ and $\#_1$, $\#_1$ and $\#_2,\ldots,\#_\lambda$ and $\#_{\lambda+1},\ldots)$, then $\varepsilon_{\mu+1}$ respects $\cdot$ (respectively $\#,\#_1,\ldots,\#_\lambda,\ldots)$. To summarize:

In Q^1, ε respects 0 , S , $+$, $\cdot$, $\#$;

ε_1 respects 0 , S , $+$, $\cdot$.

In Q^2, ε respects 0 , S , $+$, $\cdot$, $\#$, $\#_1$;

ε_1 respects 0 , S , $+$, $\cdot$, $\#$;

ε_2 respects 0 , S , $+$, $\cdot$.

$\vdots$

In Q^μ, ε respects 0 , S , $+$, $\cdot$, $\#$, $\#_1,\ldots,\#_{\mu-1}$;

ε_1 respects 0 , S , $+$, $\cdot$, $\#$, $\#_1,\ldots,\#_{\mu-2}$;

$\vdots$

ε_λ respects 0 , S , $+$, $\cdot$, $\#$, $\#_1,\ldots,\#_{\mu-\lambda-1}$;

$\vdots$

$\varepsilon_{\mu-1}$ respects 0 , S , $+$, $\cdot$, $\#$;

ε_μ respects 0 , S , $+$, $\cdot$.

The theory Q^μ in which we choose to work may depend on what is required of the various ε_λ . If we want ε_8 to respect $\cdot$, we shall not hesitate to work in Q^8 .

§3. *General Properties of Many-Sorted Theories*

In studying mathematics predicatively, we shall find it advantageous to work in a theory in which there is more than one kind of object. This section presents an introduction to such "many-sorted theories", with emphasis on the syntactic notions that will be of use to us; interpretations and extensions by definitions. (The generalization from the one-sorted to the many-sorted case is on the whole straightforward, even obvious. It is simply to familiarize the reader with the appropriate concepts and notation, and because I know of no good exposition to which he can be referred, that the main points are outlined here.) There follows an application of many-sorted theories to the general problem of constructing equivalence classes for a given equivalence relation.

In this section and this section only, numbered formulas need not be axioms, theorems, or definitions in a specific theory. Their function will be explained as necessary.

Many-sorted languages and theories

A ν-*sorted language* L $(\nu = 1,2,\ldots)$ may be described as follows. Associated with L are *sorts* $\sigma_1,\sigma_2,\ldots\sigma_\nu$. The symbols of L are the following:

(i) For each sort τ among $\sigma_1,\ldots,\sigma_\nu$, the *variables of sort* τ: $x_1^\tau, x_2^\tau, \ldots$.

(ii) Certain *predicate symbols*, each of which has a *type* $(\tau_1,\tau_2,\ldots,\tau_\lambda)$ for some (not necessarily distinct) sorts $\tau_1,\ldots,\tau_\lambda$ among $\sigma_1,\ldots,\sigma_\nu$. Such a function symbol is λ-*ary*, or of *degree* λ .

If $\tau_1,\ldots,\tau_\lambda$ are all the same sort τ, then the predicate symbol is said to be of sort τ. In particular, we require that for each sort τ there be a binary predicate symbol $=_\tau$ of sort τ (of type (τ,τ)).

(iii) Certain *function symbols*, each of which has a *type* $(\tau_1,\tau_2,\ldots,\tau_\lambda;\tau)$ for some (not necessarily distinct) sorts $\tau_1,\ldots,\tau_\lambda$, τ among $\sigma_1,\ldots,\sigma_\nu$. Such a function symbol is λ-*ary*, or of *degree* λ. If the sorts $\tau_1,\ldots,\tau_\lambda$, and τ are all the same, then the function symbol is said to be of sort τ. Of course, λ may be 0, in which case the function symbol is a *constant symbol of sort* τ.

(iv) The usual complement of connectives and quantifiers (as usual, $\neg$, $\vee$, and $\exists$ will suffice).

The *terms* of L are built up from the variables via the function symbols; each term has one of the sorts $\sigma_1,\ldots,\sigma_\nu$ associated with it. Every variable of sort τ is a term of sort τ. If $\underline{a}_1,\ldots,\underline{a}_\lambda$ are terms of sorts $\tau_1,\ldots,\tau_\lambda$, respectively, and if $\underline{f}$ is a λ-ary function symbol of type $(\tau_1,\ldots,\tau_\lambda;\tau)$, then $\underline{f}\ \underline{a}_1\cdots\underline{a}_\lambda$ is a term of sort τ.

If $\underline{a}_1,\ldots,\underline{a}_\lambda$ are terms of sort $\tau_1,\ldots,\tau_\lambda$, respectively, and if $\underline{p}$ is a λ-ary predicate symbol of type $(\tau_1,\ldots,\tau_\lambda)$, then $\underline{p}\ \underline{a}_1\cdots\underline{a}_\lambda$ is an *atomic formula* of L. In particular, $\underline{a} =_\tau \underline{b}$ is an atomic formula of L if $\underline{a}$ and $\underline{b}$ are terms of sort τ. *Formulas* of L are built up from atomic formulas in the usual way by using the connectives and quantifiers: if $\mathbb{A}$ and $\mathbb{B}$ are formulas and $\underline{x}$ is a variable of any sort whatsoever, then $\neg\mathbb{A}$, $\mathbb{A}\vee\mathbb{B}$, and $\exists\underline{x}\mathbb{A}$ are formulas.

Converting any standard system of logical axioms and rules of deduction for one-sorted languages into a corresponding system for a many-sorted language is a straightforward matter and need not be detailed here; the only modifications take the form of restrictions that all variables and function and predicate symbols be of sorts and types appropriate to one another. A *many-sorted theory* is specified by giving a many-sorted language and certain *nonlogical axioms* (formulas of the language). Again with only the most obvious of restrictions, all the usual syntactic results about theorems and proofs in first-order theories (for instance, the deduction theorem and the other results in Chapter 3 of [2]) carry over easily to the many-sorted case.

Extensions by definitions

Let $\mathbb{D}$ be a formula of a many-sorted theory T, and assume that $\underline{x}_1,\ldots,\underline{x}_\lambda$ are distinct variables of sorts $\tau_1,\ldots,\tau_\lambda$ respectively, with the property that no variable other than $\underline{x}_1,\ldots,\underline{x}_\lambda$ occurs free in $\mathbb{D}$. Form U from T by adjoining a new λ-ary predicate symbol $\underline{p}$ of type $(\tau_1,\ldots,\tau_\lambda)$ together with the (defining) axiom $\underline{p}\ \underline{x}_1\ldots\underline{x}_\lambda \longleftrightarrow \mathbb{D}$. Exactly as in [3,§74] or [2,§4.6], it can be shown that every formula $\mathbb{A}$ of U has a "translation" $\mathbb{A}'$ in T with the property that $\mathbb{A}'$ is a theorem of T if and only if $\mathbb{A}$ is a theorem of U. If the new symbol $\underline{p}$ does not occur in $\mathbb{A}$ (that is, if $\mathbb{A}$ is a formula of T), then $\mathbb{A}'$ is $\mathbb{A}$; hence U is a conservative extension of T.

In the analogous situation for function symbols, we have a theory $\mathcal{T}$, distinct variables $\underline{x}_1,\ldots,\underline{x}_\lambda,\underline{y},\underline{y}'$ of sorts $\tau_1,\ldots,\tau_\lambda,\tau,\tau$ respectively, and a formula $\mathbb{D}[\underline{x}_1,\ldots,\underline{x}_\lambda,\underline{y}]$ in which no variable other than those displayed occurs free. If we have proofs in $\mathcal{T}$ of the existence condition $\exists\underline{y}\,\mathbb{D}[\underline{x}_1,\ldots,\underline{x}_\lambda,\underline{y}]$ and the uniqueness condition $\mathbb{D}[\underline{x}_1,\ldots,\underline{x}_\lambda,\underline{y}] \& \mathbb{D}[\underline{x}_1,\ldots,\underline{x}_\lambda,\underline{y}'] \longrightarrow \underline{y} =_\tau \underline{y}'$, then we can adjoin to $\mathcal{T}$ a new function symbol $\underline{f}$ of type $(\tau_1,\ldots,\tau_\lambda;\tau)$ and the (defining) axiom $\underline{y} =_\tau \underline{f}\,\underline{x}_1\ldots\underline{x}_\lambda \longleftrightarrow \mathbb{D}[\underline{x}_1,\ldots,\underline{x}_\lambda,\underline{y}]$. Again, the translation procedure given in [3,§74] for one-sorted theories can be duplicated in the general case; it follows that the extension in question is conservative.

By iterating extensions of the two kinds just discussed, we obtain *extensions by definitions* of $\mathcal{T}$. Such a theory, being a conservative extension of $\mathcal{T}$, is consistent if and only if $\mathcal{T}$ is consistent.

Interpretations

We shall now define the notion of an *interpretation* of a many-sorted theory $\mathcal{U}$ in a many-sorted theory $\mathcal{T}$. Here $\mathcal{T}$ and $\mathcal{U}$ need not have the same sorts, or even the same number of sorts. First, an interpretation I associates with every sort τ of $\mathcal{U}$ a sort $I(\tau)$ of $\mathcal{T}$ and a unary predicate symbol U_τ of $\mathcal{T}$ of sort $I(\tau)$. With each λ-ary function symbol $\underline{f}$ of $\mathcal{U}$ of type $(\tau_1,\ldots,\tau_\lambda;\tau)$, I associates a λ-ary function symbol $\underline{f}_I$ of $\mathcal{T}$ of type $(I(\tau_1),\ldots,I(\tau_\lambda);I(\tau))$, and similarly for predicate symbols, including the equality symbols $=_\tau$. (It is not required that $(=_\tau)_I$ be $=_{I(\tau)}$. In this respect, the notion of interpretation used here generalizes that discussed in

[2,§4.7], even for one-sorted theories.) For all sorts $\tau_1,\ldots,\tau_\lambda,\tau$ of $\mathcal{U}$, all function symbols $\underline{f}$ of $\mathcal{U}$ of type $(\tau_1,\ldots,\tau_\lambda;\tau)$, and all predicate symbols $\underline{p}$ of $\mathcal{U}$ of type $(\tau_1,\ldots,\tau_\lambda)$, and with all variables understood to be of the proper sorts, formulas (3.1)-(3.5) are required to be theorems of $\mathcal{T}$:

3.1) $$\exists \underline{x}\, U_\tau \underline{x} \;;$$

3.2) $$U_{\tau_1}\underline{x}_1 \& U_{\tau_2}\underline{x}_2 \& \ldots \& U_{\tau_\lambda}\underline{x}_\lambda \longrightarrow U_\tau \underline{f}_I \underline{x}_1 \underline{x}_2 \ldots \underline{x}_\lambda \;;$$

3.3) $$U_\tau \underline{x} \longrightarrow \underline{x}(=_\tau)_I\, \underline{x};$$

3.4) $$U_{\tau_1}\underline{x}_1 \& U_{\tau_1}\underline{y}_1 \& \ldots \& U_{\tau_\lambda}\underline{x}_\lambda \& U_{\tau_\lambda}\underline{y}_\lambda \& \underline{x}_1 (=_{\tau_1})_I \underline{y}_1 \& \ldots \& \underline{x}_\lambda (=_{\tau_\lambda})_I \underline{y}_\lambda \longrightarrow$$
$$\underline{f}_I \underline{x}_1 \ldots \underline{x}_\lambda \,(=_\tau)_I\, \underline{f}_I \underline{y}_1 \ldots \underline{y}_\lambda \;;$$

3.5) $$U_{\tau_1}\underline{x}_1 \& U_{\tau_1}\underline{y}_1 \& \ldots \& U_{\tau_\lambda}\underline{x}_\lambda \& U_{\tau_\lambda}\underline{y}_\lambda \& \underline{x}_1 (=_{\tau_1})_I \underline{y}_1 \& \ldots \& \underline{x}_\lambda (=_{\tau_\lambda})_I \underline{y}_\lambda \longrightarrow$$
$$(\underline{p}_I \underline{x}_1 \ldots \underline{x}_\lambda \longrightarrow \underline{p}_I \underline{y}_1 \ldots \underline{y}_\lambda) \;.$$

Note that (3.1) and (3.2) are automatic if, as is often the case, U_τ holds universally for objects of sort $I(\tau)$; also, (3.3)-(3.5) are automatic whenever equality is interpreted by equality.

With each formula $\mathbb{B}$ of $\mathcal{U}$ there is associated a formula $\mathbb{B}^I$ of $\mathcal{T}$, called the *interpretation of* $\mathbb{B}$ by I . First, $\mathbb{B}_I$ is the formula of $\mathcal{T}$ obtained from $\mathbb{B}$ by replacing each function symbol $\underline{f}$ by $\underline{f}_I$, each predicate symbol $\underline{p}$ by $\underline{p}_I$, and each formula $\exists \underline{x}\, \mathbb{C}$ by $\exists \underline{x}(U_\tau \underline{x} \& \mathbb{C})$, where the variable $\underline{x}$ is of sort τ in $\mathcal{U}$.

(We are being a bit careless here in allowing the same letters to stand for variables of sort τ in $\mathcal{U}$ and variables of $I(\tau)$ in $\mathcal{T}$.) Then $\mathbb{B}^I$ is $U_{\tau_1}\underline{x}_1 \& \ldots \& U_{\tau_\lambda}\underline{x}_\lambda \longrightarrow \mathbb{B}_I$, where $\underline{x}_1,\ldots,\underline{x}_\lambda$ are the variables free in $\mathbb{B}$ (in some agreed-upon order) and are of sorts $\tau_1,\ldots,\tau_\lambda$ respectively (in $\mathcal{U}$). In order that I be an interpretation of $\mathcal{U}$ in $\mathcal{T}$, the final requirement is that $\mathbb{B}^I$ be a theorem of $\mathcal{T}$ for every nonlogical axiom $\mathbb{B}$ of $\mathcal{U}$.

The *interpretation theorem* for many-sorted theories asserts that under the above conditions, the interpretation $\mathbb{B}^I$ of every theorem $\mathbb{B}$ of $\mathcal{U}$ is a theorem of $\mathcal{T}$. The many-sortedness causes no problems at all in the proof. In fact, the only change necessary from the proof given in [2,§4.7] owes itself to our lenience in interpreting equality: the interpretations of identity and equality axioms need not *be* identity and equality axioms; rather, they are precisely (3.3)-(3.5), and are therefore provable in $\mathcal{T}$ by assumption.

The many-sorted theory $\mathcal{U}$ is *interpretable* in $\mathcal{T}$ if there is an interpretation of $\mathcal{U}$ in some extension by definitions of $\mathcal{T}$. By the above results, if $\mathcal{U}$ is interpretable in $\mathcal{T}$ and $\mathcal{T}$ is consistent, then $\mathcal{U}$ is consistent.

Equivalence classes

Let $\mathcal{T}$ be a (many-sorted) theory, let $\mathbb{A}[\underline{x}]$ be a formula of $\mathcal{T}$ with one free variable $\underline{x}$ of sort σ and no other free variables, and let $\sim$ be a binary predicate symbol of $\mathcal{T}$ of sort σ. Assume that (3.6)-(3.10) are theorems of $\mathcal{T}$:

3.6) $\exists \underline{x}\, \mathbf{A}[\underline{x}]$;

3.7) $\underline{x} \sim \underline{y} \longrightarrow \mathbf{A}[\underline{x}]$;

3.8) $\mathbf{A}[\underline{x}] \longrightarrow \underline{x} \sim \underline{x}$;

3.9) $\underline{x} \sim \underline{y} \longrightarrow \underline{y} \sim \underline{x}$;

3.10) $\underline{x} \sim \underline{y} \& \underline{y} \sim \underline{z} \longrightarrow \underline{x} \sim \underline{z}$.

The formula $\mathbf{A}[\underline{x}]$ may be regarded as defining the domain of the equivalence relation $\sim$.

If the theory T contains a reasonable amount of set theory, it may be possible to define in the traditional way the equivalence classes modulo the relation $\sim$, and perhaps even the set of all equivalence classes. This will not always be the case, however. It is our purpose now to show that it is always possible to study the equivalence classes with relative ease in a theory with one more sort than T .

Let $\tilde{T}$ be obtained from T by adjoining a new sort $\tilde{\sigma}$, whose objects will be regarded as the equivalence classes; a new binary predicate symbol ϵ of type $(\sigma, \tilde{\sigma})$; and three new nonlogical axioms (here Greek letters are variables of sort $\tilde{\sigma}$) :

3.11) $\exists \underline{x}(\mathbf{A}[\underline{x}] \& \forall \underline{y}(\underline{y} \epsilon \alpha \longleftrightarrow \underline{y} \sim \underline{x}))$;

3.12) $\mathbf{A}[\underline{x}] \longrightarrow \exists \alpha(\underline{x} \epsilon \alpha)$;

3.13) $\underline{x} \epsilon \alpha \& \underline{x} \epsilon \beta \longrightarrow \alpha =_{\tilde{\sigma}} \beta$.

We shall now show that $\tilde{T}$ is interpretable in T .

If τ is a sort in T, let $I(\tau)$ be τ and define $U_\tau \underline{x} \longleftrightarrow \underline{x} =_\tau \underline{x}$. Let $I(\tilde{\sigma})$ be σ, and define $U_{\tilde{\sigma}} \underline{x} \longleftrightarrow \mathbb{A}[\underline{x}]$. If $\underline{u}$ is a function or predicate symbol of T, let $\underline{u}_I$ be $\underline{u}$. Finally, let both $(=_{\tilde{\sigma}})_I$ and ϵ_I be $\sim$. We must verify conditions (3.1)-(3.5) and show that the interpretation by I of every nonlogical axiom of $\tilde{T}$ is a theorem of T.

Since U_τ holds universally if τ is not $\tilde{\sigma}$, and since there are no function symbols of $\tilde{T}$ whose types involve $\tilde{\sigma}$, we need not check (3.2); (3.1) is also automatic unless τ is $\tilde{\sigma}$, in which case (3.1) is just (3.6). If τ is not $\tilde{\sigma}$, then $(=_\tau)_I$ is $=_\tau$, so the only conditions among (3.3)-(3.5) that require checking are (3.3) when τ is $\tilde{\sigma}$ and (3.5) when $\underline{p}$ is $=_{\tilde{\sigma}}$ or ϵ. If τ is $\tilde{\sigma}$, then (3.3) is (3.8). If $\underline{p}$ is $=_{\tilde{\sigma}}$, then (3.5) is

$$\mathbb{A}[\underline{x}_1]\&\mathbb{A}[\underline{y}_1]\&\mathbb{A}[\underline{x}_2]\&\mathbb{A}[\underline{y}_2]\&\underline{x}_1 \sim \underline{y}_1 \& \underline{x}_2 \sim \underline{y}_2 \longrightarrow (\underline{x}_1 \sim \underline{x}_2 \longrightarrow \underline{y}_1 \sim \underline{y}_2),$$

which follows from (3.9) and (3.10). If $\underline{p}$ is ϵ, then (3.5) is practically the same:

$$\underline{x}_1 =_\sigma \underline{x}_1 \& y_1 =_\sigma \underline{y}_1 \& \mathbb{A}[\underline{x}_2]\&\mathbb{A}[\underline{y}_2]\&\underline{x}_1 =_\sigma \underline{y}_1 \& \underline{x}_2 \sim \underline{y}_2 \longrightarrow (\underline{x}_1 \sim \underline{x}_2 \longrightarrow \underline{y}_1 \sim \underline{y}_2).$$

If $\mathbb{B}$ is a formula of T, then $\mathbb{B}^I$ is just $\mathbb{B}$ save for a few embellishments of the form $\underline{x} =_\tau \underline{x}$; certainly $\mathbb{B}^I \longleftrightarrow \mathbb{B}$ is a theorem of T. In particular, if $\mathbb{B}$ is a nonlogical axiom of T, then $\mathbb{B}^I$ is a theorem of T. We must show that the same is true if $\mathbb{B}$ is one of the new axioms (3.11)-(3.13). First, $(3.11)^I$ is

$$\mathbb{A}[\alpha] \longrightarrow \exists \underline{x}(\underline{x} =_\sigma \underline{x} \& \mathbb{A}_I[\underline{x}] \& \forall \underline{y}(\underline{y} =_\sigma \underline{y} \longrightarrow (\underline{y} \sim \alpha \longleftrightarrow \underline{y} \sim \underline{x}))),$$

which is a theorem of T since $\mathbb{A}_I[\underline{x}] \longleftrightarrow \mathbb{A}[\underline{x}]$: just let $\underline{x}$ be α. Next, $(3.12)^I$ is $\underline{x} =_\sigma \underline{x} \longrightarrow (\mathbb{A}_I[\underline{x}] \longrightarrow \exists\alpha(\mathbb{A}[\alpha]\&\underline{x} \sim \alpha))$, which is even easier. Finally, $(3.13)^I$ is

$\underline{x} =_\sigma \underline{x} \& \mathbb{A}[\alpha] \& \mathbb{A}[\beta] \longrightarrow (\underline{x} \sim \alpha \& \underline{x} \sim \beta \longrightarrow \alpha \sim \beta)$, which is also a theorem of $\mathcal{T}$. The proof that I is an interpretation is complete.

If $\underline{f}$ is a λ-ary function symbol of sort σ in $\mathcal{T}$ such that $\underline{x}_1 \sim \underline{y}_1 \& \ldots \& \underline{x}_\lambda \sim \underline{y}_\lambda \longrightarrow \underline{f}\ \underline{x}_1 \ldots \underline{x}_\lambda \sim \underline{f}\ \underline{y}_1 \ldots \underline{y}_\lambda$ is a theorem of $\mathcal{T}$, one might reasonably expect that $\underline{f}$ "induces" a λ-ary function symbol $\tilde{\underline{f}}$ of sort $\tilde{\sigma}$ in $\tilde{\mathcal{T}}$. This is indeed the case, as will now be shown. In fact, a somewhat more general situation can be handled without much more effort.

Let $\tau_1, \ldots, \tau_\lambda, \tau_{\lambda+1}$ be (not necessarily distinct) sorts in $\mathcal{T}$. Let $\tau'_1, \ldots, \tau'_\lambda, \tau'_{\lambda+1}$ be sorts in $\tilde{\mathcal{T}}$ such that for $\kappa = 1, \ldots, \lambda+1$, if τ_κ is not σ , then τ'_κ is τ_κ , and if τ_κ is σ , then τ'_κ is either σ or $\tilde{\sigma}$. For $\kappa = 1, \ldots, \lambda+1$, let $\sim_\kappa$ be the binary predicate symbol of sort τ_κ in $\mathcal{T}$ defined as follows: if τ'_κ is τ_κ , then $\sim_\kappa$ is $=_{\tau_\kappa}$; if τ'_κ is $\tilde{\sigma}$, then $\sim_\kappa$ is $\sim$. For $\kappa = 1, \ldots, \lambda+1$, let ϵ_κ be the binary predicate symbol of type $(\tau_\kappa, \tau'_\kappa)$ in $\tilde{\mathcal{T}}$ defined as follows: if τ'_κ is τ_κ , then ϵ_κ is $=_{\tau_\kappa}$; if τ'_κ is $\tilde{\sigma}$, then ϵ_κ is ϵ .

Assume that $\underline{f}$ is a function symbol of type $(\tau_1, \ldots, \tau_\lambda; \tau_{\lambda+1})$ in $\mathcal{T}$ such that in $\mathcal{T}$ one can prove

3.14) $\underline{x}_1 \sim_1 \underline{y}_1 \& \ldots \& \underline{x}_\lambda \sim_\lambda \underline{y}_\lambda \longrightarrow \underline{f}\ \underline{x}_1 \ldots \underline{x}_\lambda \sim_{\lambda+1} \underline{f}\ \underline{y}_1 \ldots \underline{y}_\lambda$.

We claim that under this condition

3.15) $\tilde{\underline{f}}\ \underline{z}_1 \ldots \underline{z}_\lambda =_{\tau_{\lambda+1}} \underline{z}_{\lambda+1} \longleftrightarrow \exists \underline{x}_1 \ldots \exists \underline{x}_\lambda (\underline{x}_1\ \epsilon_1\ \underline{z}_1 \& \ldots \& \underline{x}_\lambda\ \epsilon_\lambda\ \underline{z}_\lambda\ \&$

$\underline{f}\ \underline{x}_1 \ldots \underline{x}_\lambda\ \epsilon_{\lambda+1}\ \underline{z}_{\lambda+1})$

is the defining axiom of a function symbol $\underline{\tilde{f}}$ of type $(\tau_1',\ldots,\tau_\lambda';\tau_{\lambda+1}')$ in $\tilde{T}$.

To establish the existence condition for $\underline{\tilde{f}}$, we argue in $\tilde{T}$ as follows. Given $\underline{z}_1,\ldots,\underline{z}_\lambda$, choose $\underline{x}_1,\ldots,\underline{x}_\lambda$ such that $\underline{x}_1 \epsilon_1 \underline{z}_1,\ldots,\underline{x}_\lambda \epsilon_\lambda \underline{z}_\lambda$. (For each $\kappa = 1,\ldots,\lambda$, if ϵ_κ is $=_{\tau_\kappa}$, just let $\underline{x}_\kappa$ be $\underline{z}_\kappa$; if ϵ_κ is ϵ , then use (3.11) to find such an $\underline{x}_\kappa$.) Let $\underline{x}_{\lambda+1}$ be $\underline{f}\,\underline{x}_1\ldots\underline{x}_\lambda$. We wish to show that there is some $\underline{z}_{\lambda+1}$ such that $\underline{x}_{\lambda+1} \epsilon_{\lambda+1} \underline{z}_{\lambda+1}$, for such a $\underline{z}_{\lambda+1}$ will necessarily satisfy the right side of (3.15). If $\epsilon_{\lambda+1}$ is $=_{\tau_{\lambda+1}}$, the assertion is trivial; assume therefore that $\epsilon_{\lambda+1}$ is ϵ . Then $\tau_{\lambda+1}$ is σ , $\tau_{\lambda+1}'$ is $\tilde{\sigma}$, and $\sim_{\lambda+1}$ is $\sim$. Certainly $\underline{x}_1 \sim_1 \underline{x}_1 \& \ldots \& \underline{x}_\lambda \sim_\lambda \underline{x}_\lambda$. (If $\sim_\kappa$ is $\sim$, then $\underline{x}_\kappa \sim \underline{x}_\kappa$ follows from $\underline{x}_\kappa \; \epsilon_\kappa \; \underline{z}_\kappa$, (3.11), (3.9), and (3.10).) Hence (3.14) implies that $\underline{x}_{\lambda+1} \sim_{\lambda+1} \underline{x}_{\lambda+1}$, which is to say $\underline{x}_{\lambda+1} \sim \underline{x}_{\lambda+1}$. By (3.7), $\mathbb{A}[\underline{x}_{\lambda+1}]$, so that by (3.12), $\exists \underline{z}_{\lambda+1}(\underline{x}_{\lambda+1} \in \underline{z}_{\lambda+1})$, as desired.

We now tackle the uniqueness condition, again arguing in $\tilde{T}$. Suppose $\underline{z}_{\lambda+1}$ and $\underline{w}_{\lambda+1}$ are such that, for some $\underline{x}_1,\ldots,\underline{x}_\lambda,\underline{y}_1,\ldots,\underline{y}_\lambda$ we have $\underline{x}_1 \epsilon_1 \underline{z}_1 \& \ldots \& \underline{x}_\lambda \epsilon_\lambda \underline{z}_\lambda \& \underline{f}\underline{x}_1\ldots\underline{x}_\lambda \epsilon_{\lambda+1} \underline{z}_{\lambda+1}$ and $\underline{y}_1 \epsilon_1 \underline{z}_1 \& \ldots \& \underline{y}_\lambda \epsilon_\lambda \underline{z}_\lambda \& \underline{f}\,\underline{y}_1\ldots\underline{y}_\lambda \epsilon_{\lambda+1} \underline{w}_{\lambda+1}$. For $\kappa = 1,\ldots,\lambda$, we have $\underline{x}_\kappa \epsilon_\kappa \underline{z}_\kappa \& \underline{y}_\kappa \epsilon_\kappa \underline{z}_\kappa$, whence $\underline{x}_\kappa \sim_\kappa \underline{y}_\kappa$ (using (3.11) if $\sim_\kappa$ is $\sim$). Let $\underline{x}_{\lambda+1}$ be $\underline{f}\,\underline{x}_1\ldots\underline{x}_\lambda$, and let $\underline{y}_{\lambda+1}$ be $\underline{f}\,\underline{y}_1\ldots\underline{y}_\lambda$; it follows from (3.14) that $\underline{x}_{\lambda+1} \sim_{\lambda+1} \underline{y}_{\lambda+1}$. But $\underline{x}_{\lambda+1} \epsilon_{\lambda+1} \underline{z}_{\lambda+1} \& \underline{y}_{\lambda+1} \epsilon_{\lambda+1} \underline{w}_{\lambda+1}$. If $\epsilon_{\lambda+1}$ is $=_{\tau_{\lambda+1}}$, then $\sim_{\lambda+1}$ is also $=_{\tau_{\lambda+1}}$, and so obviously $\underline{z}_{\lambda+1} =_{\tau_{\lambda+1}} \underline{w}_{\lambda+1}$. Otherwise, $\epsilon_{\lambda+1}$ is ϵ , $\sim_{\lambda+1}$ is $\sim$, and $\underline{z}_{\lambda+1} =_{\tau_{\lambda+1}} \underline{w}_{\lambda+1}$ follows from (3.11) and (3.13). Thus $\underline{z}_{\lambda+1} =_{\tau_{\lambda+1}} \underline{w}_{\lambda+1}$.

If every τ_κ is σ and every τ'_κ is $\tilde{\sigma}$, then $\tilde{\underline{f}}$ is the "induced" function symbol mentioned above; if every τ'_κ is τ_κ, $\tilde{\underline{f}}\ \underline{x}_1 \cdots \underline{x}_\lambda =_{\tau_{\lambda+1}} \underline{f}\ \underline{x}_1 \cdots \underline{x}_\lambda$ is a theorem of $\tilde{T}$. Observe that instead of using (3.15), we could equivalently have defined $\tilde{\underline{f}}$ by

$\tilde{\underline{f}}\ \underline{z}_1 \cdots \underline{z}_\lambda =_{\tau_{\lambda+1}} \underline{z}_{\lambda+1} \longleftrightarrow \forall \underline{x}_1 \cdots \forall \underline{x}_\lambda (\underline{x}_1 \in_1 \underline{z}_1 \& \cdots \& \underline{x}_\lambda \in_\lambda \underline{z}_\lambda \longrightarrow \underline{f}\ \underline{x}_1 \cdots \underline{x}_\lambda \in_{\lambda+1} \underline{z}_{\lambda+1})$. The reader may also check that, with notation as above, if $\underline{p}$ is a predicate symbol of type $(\tau_1,\ldots,\tau_\lambda)$ in T such that $\underline{x}_1 \sim_1 \underline{y}_1 \& \cdots \& \underline{x}_\lambda \sim_\lambda \underline{y}_\lambda \longrightarrow (\underline{p}\ \underline{x}_1 \cdots \underline{x}_\lambda \longrightarrow \underline{p}\ \underline{y}_1 \cdots \underline{y}_\lambda)$ is a theorem of T, then the definitions

$$\tilde{\underline{p}}\ \underline{z}_1 \cdots \underline{z}_\lambda \longleftrightarrow \exists \underline{x}_1 \cdots \exists \underline{x}_\lambda (\underline{x}_1 \in_1 \underline{z}_1 \& \cdots \& \underline{x}_\lambda \in_\lambda \underline{z}_\lambda \& \underline{p}\ \underline{x}_1 \cdots \underline{x}_\lambda)$$

and

$$\tilde{\underline{p}}\ \underline{z}_1 \cdots \underline{z}_\lambda \longleftrightarrow \forall \underline{x}_1 \cdots \forall \underline{x}_\lambda (\underline{x}_1 \in_1 \underline{z}_1 \& \cdots \& \underline{x}_\lambda \in \underline{z}_\lambda \longrightarrow \underline{p}\ \underline{x}_1 \cdots \underline{x}_\lambda)$$

of a predicate symbol $\tilde{\underline{p}}$ of type $(\tau'_1,\ldots,\tau'_\lambda)$ in $\tilde{T}$ are equivalent.

One significant example of a function symbol defined as above deserves special mention, namely the "quotient map". Assume that in an extension by definitions of T there is a constant symbol $\underline{e}$ of sort σ such that $\mathbb{A}[\underline{e}]$ is a theorem, and let $\underline{h}$ be the unary function symbol of sort σ defined by

$\underline{h}\ \underline{x} =_\sigma \underline{y} \longleftrightarrow (\mathbb{A}[\underline{x}] \& \underline{y} =_\sigma \underline{x}) \vee (\neg \mathbb{A}[\underline{x}] \& \underline{y} =_\sigma \underline{e})$. Replace T by an extension by definitions if necessary so that $\underline{e}$ and $\underline{h}$ are symbols of T, and form $\tilde{T}$. Regard $\underline{h}$ as a function symbol of type $(\tau_1;\tau_2)$, where τ_1 and τ_2 are both σ; let τ'_1 be σ, and let τ'_2 be $\tilde{\sigma}$. The appropriate form of (3.14) is

$\underline{x} =_\sigma \underline{y} \longrightarrow \underline{h}\ \underline{x} \sim \underline{h}\ \underline{y}$, which is a theorem of T since $\mathbb{A}[\underline{h}\ \underline{x}]$ always holds. By the general result above, 3.16) $$\underline{\tilde{h}}\ \underline{z}_1 =_{\tilde{\sigma}} \underline{z}_2 \longleftrightarrow \exists \underline{x}_1(\underline{x}_1 =_\sigma \underline{z}_1 \& \underline{h}\ \underline{x}_1 \in \underline{z}_2)$$

is a legitimate defining axiom for the quotient map symbol $\underline{\tilde{h}}$, a function symbol of type $(\sigma;\tilde{\sigma})$ in $\tilde{T}$ that we regard as assigning to each $\underline{x}$ such that $\mathbb{A}[\underline{x}]$ its equivalence class $\underline{\tilde{h}}\ \underline{x}$. Indeed, in this special case we can write $\underline{\tilde{x}}$ rather than $\underline{\tilde{h}}\ \underline{x}$. Observe that if $\underline{e}$ is as above, then by another appeal to the general result, the equivalence class $\underline{\tilde{e}}$ can be defined as a constant symbol of sort $\tilde{\sigma}$ in $\tilde{T}$. Cleaned up a bit in appearance, the definition (3.16) then becomes

3.17) $$\underline{\tilde{x}} =_{\tilde{\sigma}} \underline{z} \longleftrightarrow (\mathbb{A}[\underline{x}] \& \underline{x} \in \underline{z}) \vee (\neg \mathbb{A}[\underline{x}] \& \underline{z} =_{\tilde{\sigma}} \underline{\tilde{e}}) ,$$

or equivalently

3.18) $$\underline{\tilde{x}} =_{\tilde{\sigma}} \underline{z} \longleftrightarrow \underline{x} \in \underline{z} , \quad \text{otherwise} \quad \underline{z} =_{\tilde{\sigma}} \underline{\tilde{e}} .$$

Why many-sorted theories?

In elementary logic texts, many-sorted theories are generally regarded (if regarded at all) as rather unimportant objects. Monk [5] discusses them under the heading "inessential variations", and Shoenfield [2] neglects them entirely. At the heart of this point of view lies a theorem: every many-sorted theory can be effectively replaced by an equally powerful one-sorted theory. To see how this works, let us write T for the many-sorted theory in question and T^* for its one-sorted replacement. The nonlogical symbols of T^* are those of T (but now they are all of one sort) together with one new unary predicate symbol S_τ for each sort τ of T (the formula $S_\tau\underline{x}$ is to be thought of as saying "$\underline{x}$ is of sort τ") .

There is an obvious procedure for translating formulas of T into formulas of T^* ; the nonlogical axioms of T^* are the translations of the nonlogical axioms of T together with all formulas $\underline{\exists x} S_\tau \underline{x}$ and $S_{\tau_1} \underline{x}_1 \& \ldots \& S_{\tau_\lambda} \underline{x}_\lambda \longrightarrow S_\tau \underline{f}\, \underline{x}_1 \ldots \underline{x}_\lambda$, where $\underline{f}$ is a function symbol of type $(\tau_1, \ldots, \tau_\lambda ; \tau)$ in T . It is not hard to show that a formula of T is a theorem of T if and only if its translation into T^* is a theorem of T^* ; for details, see [6, Chapter XII].

In light of this replaceability of many-sorted theories by one-sorted theories, one might be led to believe that the use of many-sorted theories can accomplish nothing of importance. Such a conclusion is too hasty, however; a closer look at the above discussion of equivalence relations reveals as much. The many-sorted theory $\tilde{T}$ was shown to be interpretable in T , so that the consistency of T implies that of $\tilde{T}$. On the other hand, there is no such consistency proof for the one-sorted theory $\tilde{T}^*$, for $\tilde{T}^*$ may not be interpretable in T or in T^* at all! We shall see shortly than in certain simple cases this noninterpretability can actually be proved; the basic problem is that, even allowing equality to be interpreted by something other than equality, we would face insurmountable difficulties when it came time to write down *one* formula defining the interpretation of *both* relevant kinds of equality (equality for S_σ-objects and equality for $S_{\tilde{\sigma}}$-objects). For consistency proofs, then, many-sorted theories are fundamentally indispensable.

Of course, having defined $\tilde{T}$ and proved its interpretability in T, we could *now* pass to $\tilde{T}^*$ if we wished. It seems less confusing, however, to use a few Greek letters than to throw predicate symbols like S_σ and $S_{\tilde{\sigma}}$ into all our formulas. Once the technical preliminaries of this section are taken care of, manipulating objects of various sorts is not difficult at all; indeed, it could be argued that this is the way mathematicians really think in the first place. Surely we all think of a real number, a vector space, and a long exact homology sequence as animals of three different species rather than as three sets in Zermelo-Fraenkel set theory!

One more remark: although in keeping with the traditional conception of equivalence classes as collections, the symbol ϵ was used above for the relation between an object and its equivalence class, it should be noted that objects of sort $\tilde{\sigma}$ in $\tilde{T}$ are not really collections of objects of sort σ in any intrinsic way. Equivalence classes have usually been regarded as collections presumably because those collections could be formed, say in ZF, without introducing new sorts of objects. On the other hand, there is no reason why equivalence classes should have to be so huge and unwieldy -- and in the many-sorted approach, they aren't.

A noninterpretability proof

Let T be the theory with one sort σ, one binary predicate symbol $\sim$, axioms asserting that $\sim$ is an equivalence relation, and the additional axiom

3.19) $$\exists x \exists y (x \neq y \& x \sim y) .$$

Form $\tilde{T}$ as above; $\tilde{T}$ has two sorts, σ and $\tilde{\sigma}$. Consider the one-sorted theory $\tilde{T}^*$. In $\tilde{T}^*$ are two unary predicate symbols S_σ and $S_{\tilde{\sigma}}$, in addition to $=$, $\sim$, and ϵ ; the formulas

3.20) $$\exists x \, S_\sigma x$$

and

3.21) $$\exists x \, S_{\tilde{\sigma}} x$$

are axioms of $\tilde{T}^*$. There are theorems of $\tilde{T}^*$ asserting that $\sim$ is an equivalence relation on individuals satisfying S_σ and that $=$ is an equivalence relation on all individuals. Moreover, by (3.19), (3.11), and (3.13), the following are theorems of $\tilde{T}^*$:

3.22) $$\exists x \exists y (S_\sigma x \& S_\sigma y \& x \neq y \& x \sim y) ,$$

3.23) $$S_{\tilde{\sigma}} \alpha \longrightarrow \exists x (S_\sigma x \& \forall y (S_\sigma y \longrightarrow (y \in \alpha \longleftrightarrow y \sim x))) ,$$

3.24) $$S_\sigma x \& S_{\tilde{\sigma}} \alpha \; \& S_{\tilde{\sigma}} \beta \; \& x \in \alpha \& x \in \beta \longrightarrow \alpha = \beta .$$

The theory T has a model M consisting of exactly two elements a and b with $a \neq b$ but $a \sim^M b$; we shall use M to show that $\tilde{T}^*$ is not interpretable in T . (Since T is one-sorted, T is not significantly different from T^* , so essentially the same proof shows that $\tilde{T}^*$ is not interpretable in T^* . By adjoining to T an axiom asserting that there are exactly two individuals, one can convert this proof to a purely syntactic one that makes no mention of models.)

The important fact about M is that its two elements are not distinguished from each other by any formulas of T ; that is, a formula of T with one free variable is true about a in M if and only if it is true about b in M . In fact, one can readily prove, for a general formula $\mathbb{D}$ of T , that if a certain assignment of elements of M to the free variables in $\mathbb{D}$ makes $\mathbb{D}$ true in M , then the *opposite* assignment -- interchanging a and b -- also makes $\mathbb{D}$ true in M .

Now suppose I is an interpretation of $\tilde{T}^*$ in (an extension by definitions of) T . Then I provides us with a universe U_I and symbols $(S_\sigma)_I$, $(S_{\tilde{\sigma}})_I$, $=_I$, $\sim_I$, and ϵ_I , all defined by formulas of T . The interpretations of (3.20) and (3.21), namely

3.25) $$\exists x(U_I x \& (S_\sigma)_I x)$$

and

3.26) $$\exists x(U_I x \& (S_{\tilde{\sigma}})_I x) ,$$

are theorems of T . By the above remark about M , it follows that U_I , $(S_\sigma)_I$, and $(S_{\tilde{\sigma}})_I$ all hold universally in M (that is, they hold for both a and b in M). Combining this with the fact that the interpretation of every theorem of $\tilde{T}^*$ is valid in M shows that the following are valid in M : $\sim_I$ is an equivalence relation; $=_I$ is an equivalence relation;

3.27) $$\exists x \exists y (x \neq_I y \& x \sim_I y) ;$$

3.28) $$\exists x \forall y (y \in_I \alpha \longleftrightarrow y \sim_I x) ;$$

3.29) $$x \in_I \alpha \& x \in_I \beta \longrightarrow \alpha =_I \beta .$$

From (3.27) it follows that $=_I^M$ is the same relation as $=$ and $\sim_I^M$ the same as $\sim^M$: after all, there are only two possible equivalence relations on a set with two elements. Then (3.28) implies that $y \in_I \alpha$ holds universally in M -- that is, $a \in_I^M a$, $a \in_I^M b$, $b \in_I^M a$, and $b \in_I^M b$ are all true. But this contradicts (3.29), according to which $a \in_I^M a$ and $a \in_I^M b$ would imply $a = b$. Thus $\tilde{T}^*$ is not interpretable in T .

PART TWO

PREDICATIVE ANALYSIS

§4. *Arithmetic of Fractions*

Classically, when the real numbers are constructed from the natural numbers, the first step is invariably the construction of the rational numbers. Our predicative approach begins similarly; in this section we study the arithmetic of *fractions* -- quotients of natural numbers. For reasons that will become clear when we discuss the real numbers, however, it is a mistake to think of these fractions as rational numbers.

All the work in this section can be carried out in Nelson's theory Q^0 , *a fortiori* in any of the theories Q^μ described in §2. At the outset, we call the reader's attention to the fact that *all function and predicate symbols defined in this section are bounded.*

Fractions

Formally, a fraction is defined to be an ordered triple consisting of a "sign" (0 for negative, 1 for positive), a "numerator", and a "denominator"; all fractions are required to be in lowest terms. Strictly speaking, the first definition to be made is that of ordered triples:

4.1) $$\text{Def } \langle x_1,x_2,x_3\rangle = \langle x_1,\langle x_2,x_3\rangle\rangle \quad .$$

(See (1.14) for the definition of ordered pairs.) The existence condition for the definition

4.2) $$\text{Def } \mathrm{Gcd}(a,b) = d \longleftrightarrow ((a\neq 0 \vee b\neq 0) \,\&\, d|a \,\&\, d|b \,\&$$
$$\forall c(c|a \,\&\, c|b \longrightarrow c \leq d)) \vee (a=0 \,\&\, b=0 \,\&\, d=0)$$

follows easily from the bounded least number principle, and the uniqueness condition is obvious. This paves the way for

4.3) Def r is a fraction $\longleftrightarrow \exists z \exists a \exists b((z=0 \vee z=1)$ & $b \neq 0$ & $(a=0 \longrightarrow z=1$ & $b=1)$ & $\mathrm{Gcd}(a,b) = 1$ & $r = \langle z,a,b\rangle)$.

Note that $x_1 = \mathrm{Proj}_1\langle x_1,x_2,x_3\rangle$, $x_2 = \mathrm{Proj}_1\mathrm{Proj}_2\langle x_1,x_2,x_3\rangle$, and $x_3 = \mathrm{Proj}_2\mathrm{Proj}_2\langle x_1,x_2,x_3\rangle$; hence the definitions

4.4) Def Sign $r = \mathrm{Proj}_1\ r$,

4.5) Def Numer $r = \mathrm{Proj}_1\mathrm{Proj}_2\ r$,

4.6) Def Denom $r = \mathrm{Proj}_2\mathrm{Proj}_2\ r$.

To convert an arbitrary triple $\langle 0,a,b\rangle$ or $\langle 1,a,b\rangle$ into a fraction, we must prove that it has a unique expression in lowest terms:

4.7) $(z=0 \vee z=1)$ & $a \neq 0$ & $b \neq 0 \longrightarrow$

$\exists! r$(r is a fraction & Sign $r = z$ & $a \cdot$ Denom $r = b \cdot$ Numer r).

Proof. Existence of r is easy: let r be $\langle z, \mathrm{Qt}(\mathrm{Gcd}(a,b),a), \mathrm{Qt}(\mathrm{Gcd}(a,b),b)\rangle$. To prove **uniqueness,** suppose $\langle z,c,d\rangle$ and $\langle z,c',d'\rangle$ both satisfy the requirements for r . Then $a \cdot d = b \cdot c$ and $a \cdot d' = b \cdot c'$, so $a \cdot d \cdot b \cdot c' = b \cdot c \cdot a \cdot d'$ and hence $d \cdot c' = c \cdot d'$. It suffices therefore to prove $c = c'$; we show that both equal $\mathrm{Gcd}(c,c')$. If $c = \mathrm{Gcd}(c,c') \cdot c_1$ and $c' = \mathrm{Gcd}(c,c') \cdot c_1'$ with, say, $c_1 > 1$, then there is a prime p dividing c_1 . (For instance, let p be the smallest number larger than 1 that divides c_1 ; this exists by (BLNP).) Now $d \cdot c_1' = c_1 \cdot d'$, so $p | d \cdot c_1'$. Certainly $p \nmid c_1'$, for otherwise $\mathrm{Gcd}(c,c') \cdot p$ would be a common divisor of c and c' ; therefore $p|d$. (The skeptical reader may see [1, (9.37)] for a proof.) But $p|c$ & $p|d$ contradicts the fact that $\langle z,c,d \rangle$ is a fraction. Thus $c = \mathrm{Gcd}(c,c')$, and likewise $c' = \mathrm{Gcd}(c,c')$. ||

4.8) Def Reduc $s = r \longleftrightarrow r$ is a fraction & $\exists z \exists a \exists b$ $(s = \langle z,a,b \rangle$ & $(z=0 \vee z=1)$ & $b \neq 0$ & $((a=0$ & $r = \langle 1,0,1 \rangle) \vee$ $(a \neq 0$ & $z = \mathrm{Sign}\ r$ & $a \cdot \mathrm{Denom}\ r = b \cdot \mathrm{Numer}\ r)))$, otherwise $r = 0$.

The uniqueness condition for (4.8) follows from (4.7), which also guarantees that if s is $\langle 0,a,b \rangle$ or $\langle 1,a,b \rangle$ with $b \neq 0$, then Reduc s really is a fraction that represents s in lowest terms.

4.9) $$\text{Def } \hat{n} = \langle 1,n,1 \rangle \quad .$$

The fraction $\hat{n}$ is just the fraction representing the number n . The reader armed with (1.14) may be amused to calculate that $\hat{0} = 11$. The circumflex $\hat{}$ will consistently signal arithmetical operations and relations involving fractions.

4.10) Def $r \mathrel{\hat{<}} s \longleftrightarrow r$ and s are fractions &
$((\mathrm{Sign}\ r = \mathrm{Sign}\ s = 1$ & $\mathrm{Numer}\ r \cdot \mathrm{Denom}\ s < \mathrm{Numer}\ s \cdot \mathrm{Denom}\ r) \vee$
$(\mathrm{Sign}\ r = 0$ & $\mathrm{Sign}\ s = 1) \vee$
$(\mathrm{Sign}\ r = \mathrm{Sign}\ s = 0$ & $\mathrm{Numer}\ s \cdot \mathrm{Denom}\ r < \mathrm{Numer}\ r \cdot \mathrm{Denom}\ s))$.

4.11) Def $r \hat{\leq} s \longleftrightarrow r$ and s are fractions & $(r \hat{<} s \vee r = s)$.

The order axioms are now arithmetical trivialities; hardest is (4.15), which uses (4.7).

4.12) $$\neg\, r \hat{<} r . \quad \|$$

4.13) $$\neg (r \hat{<} s \;\&\; s \hat{<} r) . \; \|$$

4.14) $$r \hat{<} s \;\&\; s \hat{<} t \longrightarrow r \hat{<} t . \; \|$$

4.15) r and s are fractions $\longrightarrow r \hat{<} s \vee r = s \vee s \hat{<} r$. $\|$

We take the symbols $\hat{>}$ and $\hat{\geq}$ as abbreviations: $r \hat{>} s$ for $s \hat{<} r$, and $r \hat{\geq} s$ for $s \hat{\leq} r$.

Recall that there is a bounded function symbol Bd such that if a is a set, then Bd a is the largest number in a according to the ordering $<$. We next define a bounded function symbol $\hat{\text{Max}}$ such that if a is a set all of whose elements are fractions, then $\hat{\text{Max}}\, a$ is the largest fraction in a according to the ordering $\hat{<}$.

4.16) Def $\hat{\text{Max}}\, a = r_0 \longleftrightarrow a$ is a set of fractions & $r_0 \in a$ & $\forall r\, (r \in a \longrightarrow r \hat{\leq} r_0)$, otherwise $r_0 = 0$.

Uniqueness is clear from (4.13). That $\hat{\text{Max}}\, a$ is what it is supposed to be is the content of

4.17) a is a set of fractions & $a \neq 0 \longrightarrow \hat{\text{Max}}\, a \in a$ & $\forall r\, (r \in a \longrightarrow r \hat{\leq} \hat{\text{Max}}\, a)$.

Proof. What must be shown is that there is some r_0 satisfying $r_0 \in a \ \& \ \forall r(r \in a \longrightarrow r \hat{\leq} r_0)$. By the bounded replacement principle, there is a set b consisting of all denominators of elements of a . (Actually, bounded replacement gives a *function* mapping each element of a to its denominator; let b be the range of that function.) Let w be the product of all elements of b (that is, $w = (\Pi \cdot \mathrm{Enum}\ b)(\mathrm{Ln}\ \mathrm{Enum}\ b))$, and, using bounded separation and bounded replacement, form the sets $c_0 = \{\mathrm{Qt}\ (\mathrm{Denom}\ r, w \cdot \mathrm{Numer}\ r) : r \in a \ \& \ \mathrm{Sign}\ r = 0\}$ and $c_1 = \{\mathrm{Qt}\ (\mathrm{Denom}\ r, w \cdot \mathrm{Numer}\ r) : r \in a \ \& \ \mathrm{Sign}\ r = 1\}$. Since a is nonempty by assumption, either c_0 or c_1 is nonempty. If c_1 is nonempty, let x be the largest element of c_1 . Then the fraction r_0 such that $x = \mathrm{Qt}\ (\mathrm{Denom}\ r_0, w \cdot \mathrm{Numer}\ r_0)$ is precisely Reduc <1,x,w> , and this r_0 is easily seen to be the desired largest fraction in a . If c_1 is empty, then all the fractions in a are negative; let x be the *smallest* element of c_0 , and let r_0 =Reduc<0,x,w> . ||

4.18) Def $\hat{-}r = s \longleftrightarrow r$ is a fraction & $((r = \hat{0} \ \& \ s = \hat{0}) \lor$
$(r \neq \hat{0} \ \& \ s = \langle 1\text{-Sign } r, \text{Numer } r, \text{Denom } r \rangle))$, otherwise $s = 0$.

4.19) Def $r_1 \hat{+} r_2 = s \longleftrightarrow$

r_1 and r_2 are fractions &

$\exists z_1 \exists a_1 \exists b_1 \exists z_2 \exists a_2 \exists b_2 (r_1 = \langle z_1, a_1, b_1 \rangle \ \& \ r_2 = \langle z_2, a_2, b_2 \rangle \ \&$

$((z_1 = z_2 \ \& \ s = \mathrm{Reduc} \langle z_1, a_1 \cdot b_2 + a_2 \cdot b_1, b_1 \cdot b_2 \rangle) \lor$

$(z_1 = 0 \ \& \ z_2 = 1 \ \& \ \hat{-}r_1 \hat{<} r_2 \ \& \ s = \mathrm{Reduc} \langle 1, a_2 \cdot b_1 - a_1 \cdot b_2, b_1 \cdot b_2 \rangle) \lor$

$(z_1 = 0 \;\&\; z_2 = 1 \;\&\; r_2 \mathbin{\hat{<}} \hat{-}r_1 \;\&\; s = \text{Reduc} \langle 0, a_1 \cdot b_2 - a_2 \cdot b_1, b_1 \cdot b_2 \rangle) \vee$

$(z_1 = 1 \;\&\; z_2 = 0 \;\&\; r_1 \mathbin{\hat{<}} \hat{-}r_2 \;\&\; s = \text{Reduc} \langle 0, a_2 \cdot b_1 - a_1 \cdot b_2, b_1 \cdot b_2 \rangle) \vee$

$(z_1 = 1 \;\&\; z_2 = 0 \;\&\; \hat{-}r_2 \mathbin{\hat{<}} r_1 \;\&\; s = \text{Reduc} \langle 1, a_1 \cdot b_2 - a_2 \cdot b_1, b_1 \cdot b_2 \rangle) \vee$

$(r_1 = \hat{-}r_2 \;\&\; s = \hat{0})))$,

otherwise $s = 0$.

The following propositions are routine.

4.20) $$r_1 \mathbin{\hat{+}} r_2 = r_2 \mathbin{\hat{+}} r_1 \;. \quad \|$$

4.21) $$r_1 \mathbin{\hat{+}} (r_2 \mathbin{\hat{+}} r_3) = (r_1 \mathbin{\hat{+}} r_2) \mathbin{\hat{+}} r_3 \;. \quad \|$$

4.22) r is a fraction $\longrightarrow r \mathbin{\hat{+}} \hat{0} = r$. $\|$

4.23) r is a fraction $\longrightarrow r \mathbin{\hat{+}} (\hat{-}r) = \hat{0}$. $\|$

Let us agree to write $r \mathbin{\hat{-}} s$ as an abbreviation for $r \mathbin{\hat{+}} (\hat{-}s)$.

4.24) Def $r_1 \mathbin{\hat{\cdot}} r_2 = s \longleftrightarrow r_1$ and r_2 are fractions &

$((\text{Sign } r_1 = \text{Sign } r_2 \;\&\; s = \text{Reduc} \langle 1, \text{Numer } r_1 \cdot \text{Numer } r_2, \text{ Denom } r_1 \cdot \text{Denom } r_2 \rangle) \vee$

$(\text{Sign } r_1 \neq \text{Sign } r_2 \;\&\; s = \text{Reduc} \langle 0, \text{Numer } r_1 \cdot \text{Numer } r_2, \text{ Denom } r_1 \cdot \text{Denom } r_2 \rangle))$,

otherwise $s = 0$.

4.25) Def Recip $r = s \longleftrightarrow r$ is a fraction & $r \neq \hat{0}$ &

$s = \langle \text{Sign } r, \text{Denom } r, \text{Numer } r \rangle$, otherwise $s = 0$.

4.26) $$\text{Def } r_1 \mathbin{\hat{/}} r_2 = r_1 \mathbin{\hat{\cdot}} \text{Recip } r_2 \;.$$

4.27) $$r_1 \mathbin{\hat{\cdot}} r_2 = r_2 \mathbin{\hat{\cdot}} r_1 \;. \quad \|$$

4.28) $$r_1 \mathbin{\hat{\cdot}} (r_2 \mathbin{\hat{\cdot}} r_3) = (r_1 \mathbin{\hat{\cdot}} r_2) \mathbin{\hat{\cdot}} r_3 \ . \quad \|$$

4.29) $$r \text{ is a fraction} \longrightarrow r \mathbin{\hat{\cdot}} \hat{\hat{1}} = r \ . \quad \|$$

4.30) $$r \text{ is a fraction} \ \& \ r \neq \hat{0} \longrightarrow r \mathbin{\hat{\cdot}} \text{Recip } r = \hat{1} \ . \quad \|$$

4.31) $$r_1 \mathbin{\hat{\cdot}} (r_2 \mathbin{\hat{+}} r_3) = r_1 \mathbin{\hat{\cdot}} r_2 \mathbin{\hat{+}} r_1 \mathbin{\hat{\cdot}} r_3 \ . \quad \|$$

4.32) $$r_1 \text{ is a fraction} \ \& \ r_2 \mathrel{\hat{<}} r_3 \longrightarrow r_1 \mathbin{\hat{+}} r_2 \mathrel{\hat{<}} r_1 \mathbin{\hat{+}} r_3 \ . \quad \|$$

4.33) $$r_1 \mathrel{\hat{<}} r_2 \ \& \ \hat{0} \mathrel{\hat{<}} r_3 \longrightarrow r_1 \mathbin{\hat{\cdot}} r_3 \mathrel{\hat{<}} r_2 \mathbin{\hat{\cdot}} r_3 \ . \quad \|$$

Propositions (4.12)-(4.15), (4.20)-(4.23) and (4.27)-(4.33) are of course the standard ordered field axioms for fractions, so every elementary theorem about ordered fields is a theorem about fractions. (Example: $\neg \exists r \, (r \text{ is a fraction} \ \& \ r \mathbin{\hat{\cdot}} r = \hat{\hat{-1}})$.)

Two more useful functions to add to our supply are the absolute value and greatest integer functions.

4.34) Def $|r| = s \longleftrightarrow r$ is a fraction & $s = \langle 1, \text{Numer } r, \text{Denom } r \rangle$, otherwise $s = 0$.

4.35) $$r \text{ is a fraction} \longrightarrow |r| \mathrel{\hat{\geq}} \hat{0} \ \& \ (|r| \mathrel{\hat{=}} \hat{0} \longleftrightarrow r = \hat{0}) \ . \quad \|$$

4.36) $$|r_1 \mathbin{\hat{\cdot}} r_2| = |r_1| \mathbin{\hat{\cdot}} |r_2| \ . \quad \|$$

4.37) $$|r_1 \mathbin{\hat{+}} r_2| \mathrel{\hat{\leq}} |r_1| \mathbin{\hat{+}} |r_2| \ . \quad \|$$

4.38) $$r \mathrel{\hat{<}} \hat{0} \longrightarrow \exists! n \, (n \leq \text{Numer } r \ \& \ \hat{\hat{-n}} \mathrel{\hat{\leq}} r \ \& \ r \mathrel{\hat{<}} \hat{-}(n-1)\hat{\ }) \ .$$

Proof. All sufficiently large n satisfy $\hat{\hat{-n}} \mathrel{\hat{\leq}} r$. Use (BLNP) to find the smallest such n . ‖

4.39) $r \mathbin{\hat{\geq}} \hat{0} \longrightarrow \exists! n\ (n \leq \text{Numer } r \mathbin{\&} \hat{n} \mathbin{\hat{\leq}} r \mathbin{\&} r \mathbin{\hat{<}} (n+1)^{\wedge})$. ||

4.40) Def $[r] = s \longleftrightarrow r$ is a fraction &

$((r \mathbin{\hat{<}} \hat{0} \mathbin{\&} \exists n\ (n \leq \text{Numer } r \mathbin{\&} \hat{-}\hat{n} \mathbin{\hat{\leq}} r \mathbin{\&} r \mathbin{\hat{<}} \hat{-}(n-1)^{\wedge} \mathbin{\&} s = \hat{-}\hat{n})) \lor$

$(r \mathbin{\hat{\geq}} \hat{0} \mathbin{\&} \exists n\ (n \leq \text{Numer } r \mathbin{\&} \hat{n} \mathbin{\hat{\leq}} r \mathbin{\&} r \mathbin{\hat{<}} (n+1)^{\wedge} \mathbin{\&} s = \hat{n})))$,

otherwise $s = 0$.

Observe that if r is a fraction, then so is $[r]$. If r is positive and we want n rather than $\hat{n}$, we take Numer $[r]$.

As remarked earlier, all of the symbols we have defined are bounded. It will be helpful to record two of the bounds explicitly; the proofs are obvious but cumbersome.

4.41) $r_1 \mathbin{\hat{+}} r_2 \leq r_1 \cdot r_2$. ||

4.42) $r_1 \mathbin{\hat{\cdot}} r_2 \leq r_1 \cdot r_2$. ||

Polynomials

We wish to study polynomials whose coefficients are fractions. Such a polynomial can be identified with its sequence of coefficients; we require that the last (highest power) coefficient be different from $\hat{0}$.

4.43) Def f is a polynomial $\longleftrightarrow f$ is a sequence of fractions &

$f(\text{Ln } f) \neq \hat{0}$.

The fact that the domain of a sequence is $\{1,\dots,n\}$ rather than $\{0,\dots,n\}$ presents a minor inconvenience: we regard $f(1)$ as the constant term, $f(2)$ as the first-power coefficient, and so on. The zero polynomial is the empty sequence 0 (not $\hat{0}$) .

4.44) Def Deg f = n $\longleftrightarrow$ f is a polynomial & Ln f = n+1 , otherwise n = 0 .

If u is any sequence of fractions, we can obtain a polynomial by truncating u -- that is, by removing a string of $\hat{0}$'s from the end.

4.45) Def Truncate u = f $\longleftrightarrow$ u is a sequence of fractions & $\exists$n (1 $\le$ n $\le$ Ln u & u(n) $\neq \hat{0}$ & $\forall$i (n < i $\le$ Ln u $\longrightarrow$ u(i) = $\hat{0}$) & f = u[1,n]) , otherwise f = 0 .

4.46) u is a sequence of fractions $\longrightarrow$ Truncate u is a polynomial.

Proof. By (BLNP) there is an n such that $\min_n$ $\forall$i (n < i $\le$ Ln u $\longrightarrow$ u(i) = $\hat{0}$) ; then Truncate u is u[1,n] . This is either the empty sequence or a sequence whose last term u(n) is not $\hat{0}$. ||

Before showing how to evaluate a polynomial at a given fraction, we require some definitions.

4.47) Def u is a power sequence of r $\longleftrightarrow$ r is a fraction & u is a sequence & u(1) = $\hat{1}$ & $\forall$i (1 $\le$ i < Ln u $\longrightarrow$ u(i+1) = u(i) $\hat{\cdot}$ r) .

If u is a power sequence of r , then u(i) is intended to represent the fraction r^{i-1} . Of course this is not a $\cdot$-power but a $\hat{\cdot}$-power; the notation r $\hat{\wedge}$ i , however, will be sedulously avoided.

4.48) r is a fraction $\longrightarrow$ $\exists$u (u is a power sequence of r & Ln u = Log n & Sup u $\le$ Explog (r,n)) .

Proof. Bounded induction on n . The bound on Sup u is a consequence of (4.42). ||

Recall from §1 that a bound on a sequence u in terms of a number n requires a bound on Ln u that is logarithmic in n .

4.49) u and v are power sequences of r & Ln u = Ln v $\longrightarrow$ u = v .

Proof. Let k be Ln u , and use bounded induction on k . ||

4.50) Def Powerseq (r,n) = u $\longleftrightarrow$ u is a power sequence of r & Ln u = Log n , otherwise u = 1 .

(We choose u = 1 rather than u = 0 in the "otherwise" clause in order that Powerseq (r,n) not be a sequence at all if r is not a fraction.)

We shall have occasion to make several definitions much like (4.50) for which the appropriate conditions and bounds have similar proofs by bounded induction. In each such case, the preliminary definition corresponding to (4.47) and the theorems analogous to (4.48) and (4.49) will be omitted as long as they are straightforward. For example, here is the definition of the termwise product of two sequences of fractions:

4.51) Def Mult (u_1,u_2) = v $\longleftrightarrow$ u_1, u_2, and v are sequences of fractions & Ln u_1 = Ln u_2 = Ln v &
$\forall i\ (1 \leq i \leq \text{Ln}\ u_1 \longrightarrow v(i) = u_1(i)\ \hat{\cdot}\ u_2(i))$, otherwise v = 1.

(The bound on Sup v is Sup u_1·Sup u_2 , again by (4.42).)

It is sometimes useful to add sequences of fractions termwise even if they are of different lengths; we have in mind, of course, the addition of polynomials.

4.52) Def Add $(u_1,u_2) = v \longleftrightarrow u_1$, u_2 , and v are sequences of fractions & Ln v = Max (Ln u_1, Ln u_2) &
$\forall i\ (1 \leq i \leq \text{Ln } u_1 \ \&\ i \leq \text{Ln } u_2 \longrightarrow v(i) = u_1(i) \hat{+} u_2(i))$ &
$\forall i\ (\text{Ln } u_1 < i \leq \text{Ln } u_2 \longrightarrow v(i) = u_2(i))$ &
$\forall i\ (\text{Ln } u_2 < i \leq \text{Ln } u_1 \longrightarrow v(i) = u_1(i))$, otherwise $v = 1$.

4.53) Def $\hat{\sum} u = v \longleftrightarrow u$ and v are sequences of fractions & Ln u = Ln v & $v(1) = u(1)$ & $\forall i\ (1 \leq i < \text{Ln } u \longrightarrow v(i+1) = v(i) \hat{+} u(i+1))$, otherwise $v = 1$.

Just as $\sum u$ is the sequence of partial +-sums of the numbers in the sequence u , $\hat{\sum} u$ is the sequence of partial $\hat{+}$-sums of the fractions in u . The total sum is $(\hat{\sum} u)(\text{Ln } u)$. Likewise:

4.54) Def $\hat{\Pi} u = v \longleftrightarrow u$ and v are sequences of fractions & Ln u = Ln v & $v(1) = u(1)$ & $\forall i\ (1 \leq i < \text{Ln } u \longrightarrow v(i+1) = v(i) \hat{\cdot} u(i+1))$, otherwise $v = 1$.

The preceding definitions provide all the necessary tools for evaluating polynomials. Recall that Ln $f \leq$ Log f by (1.18); therefore (Powerseq (r,f)) $[1,\text{Ln } f]$ is the sequence whose terms are $\hat{1}, r, r^2, \ldots, r^{\text{Deg } f}$. Then Mult $(f,(\text{Powerseq } (r,f))[1,\text{Ln } f])$ is the sequence $f(1), f(2) \hat{\cdot} r, \ldots, f(\text{Ln } f) \hat{\cdot} r^{\text{Deg } f}$, and the desired value, to be called Polyvalue (f,r) , is the sum of all the terms in *this* sequence.

4.55) Def Polyvalseq $(f,r) = u \longleftrightarrow f$ is a sequence of fractions & r is a fraction & $u = \hat{\sum}$ (Mult(f,(Powerseq (r,f))[1,Ln f])), otherwise $u = 1$.

4.56) Def Polyvalue $(f,r) = s \longleftrightarrow (f = 0$ & r is a fraction & $s = \hat{0}) \vee$ (f is a sequence of fractions & $f \neq 0$ & r is a fraction & $s =$ (Polyvalseq (f,r))(Ln f)) , otherwise $s = 0$.

No matter how silly it may appear, the following requires proof.

4.57) Polyvalue (f,r) = Polyvalue (Truncate f,r) .

Proof. The idea is clear. We may assume that $f \neq 0$ is a sequence of fractions and that r is a fraction. As in (4.46), let $\min_n \forall i\ (n < i \leq \text{Ln } f \longrightarrow f(i) = \hat{0})$, so that Truncate f = f[1,n] .

If $1 \leq i \leq n$, then

$$(\text{Powerseq } (r,f))(i) = (\text{Powerseq } (r,\text{Truncate } f))(i) ,$$

and therefore

$$(\text{Mult } (f,(\text{Powerseq } (r,f))[1,\text{Ln } f]))(i)$$
$$= (\text{Mult}(\text{Truncate } f,(\text{Powerseq}(r,\text{Truncate } f))[1,\text{Ln}(\text{Truncate } f)]))(i) .$$

By bounded induction it follows that for such i ,

$$(\text{Polyvalseq } (f,r))(i) = (\text{Polyvalseq } (\text{Truncate } f,r))(i) .$$

But $n < i \leq \text{Ln } f \longrightarrow (\text{Mult}(f,(\text{Powerseq } (r,f))[1,\text{Ln } f]))(i) = \hat{0}$, so if $n < i \leq \text{Ln } f$, then (Polyvalseq (f,r))(i) = (Polyvalseq (f,r))(n) (another bounded induction). In particular,

$$\begin{aligned}\text{Polyvalue }(f,r) &= (\text{Polyvalseq }(f,r))(\text{Ln } f)\\ &= (\text{Polyvalseq }(f,r))(n)\\ &= (\text{Polyvalseq }(\text{Truncate } f,r))(n)\\ &= \text{Polyvalue }(\text{Truncate } f,r)\ ,\end{aligned}$$

since Ln(Truncate f) = n . ||

Hereafter, obvious and tedious arguments such as the above will be condensed drastically if not omitted.

Our next main object is a theorem (4.72) to the effect that if the coefficients of two polynomials f and g are close together, and if r is close to s , then Polyvalue (f,r) is close to Polyvalue (g,s) -- a kind of continuity result, in a sense to be made precise in §§5-6. The place to start, as usual, is with definitions.

4.58) Def Negseq u = v $\longleftrightarrow$ u and v are sequences of fractions & Ln u = Ln v & $\forall$i (1 $\leq$ i $\leq$ Ln u $\longrightarrow$ v(i) = $\hat{-}$u(i)), otherwise v = 1 .

4.59) Def Subt (u_1,u_2) = Add (u_1,Negseq u_2) .

4.60) f and g are sequences of fractions & r is a fraction $\longrightarrow$ Polyvalue (Add (f,g),r) = Polyvalue (f,r) $\hat{+}$ Polyvalue (g,r) .

Proof. By bounded induction on k = Max (Ln f,Ln g) . ||

4.61) f is a sequence of fractions & r is a fraction $\longrightarrow$ Polyvalue (Negseq f,r) = $\hat{-}$Polyvalue (f,r) . ||

4.62) f and g are sequences of fractions & r is a fraction $\longrightarrow$

Polyvalue (Subt (f,g),r) = Polyvalue (f,r) $\hat{-}$ Polyvalue (g,r). ||

The next proposition says that

$(a_1+a_2r+\ldots+a_{n+1}r^n) - (a_1+a_2s+\ldots+a_{n+1}s^n) = a_1(1-1) + a_2(r-s)+\ldots+$

$a_{n+1}(r^n-s^n)$. As expected, the proof is by bounded induction on Ln f.

4.63) f is a sequence of fractions and r and s are fractions $\longrightarrow$

Polyvalue (f,r) $\hat{-}$ Polyvalue (f,s)

= ($\hat{\sum}$ Mult(f,Subt((Powerseq(r,f))[1,Ln f],(Powerseq(s,f))[1,Lnf])))(Ln f).||

4.64) Def Reverse u = v $\longleftrightarrow$ u and v are sequences & Ln u = Ln v &

$\forall$i (1 $\leq$ i $\leq$ Ln u $\longrightarrow$ v(i) = u (Ln u+1-i)) , otherwise v = 1 .

Reverse u is the sequence with the same terms as u but in the reverse order. Another sequence that will prove useful is the one with terms $r^{n-1}, r^{n-2}\cdot s, r^{n-3}\cdot s^2, \ldots, r\cdot s^{n-2}, s^{n-1}$, for which we use the name Telseq in light of the "telescoping" property (4.67):
$r^n-s^n = (r-s)(r^{n-1}+r^{n-2}s+\ldots+s^{n-1})$.

4.65) Def Telseq (r,s,f) = u $\longleftrightarrow$ r and s are fractions &

f is a sequence & u = Mult(Reverse((Powerseq(r,f))[1,Ln f-1]),

(Powerseq(s,f))[1,Ln f-1]), otherwise u = 1 .

4.66) r and s are fractions & f is a sequence & 1 $\leq$ i $\leq$ Ln f-1 $\longrightarrow$

(Telseq (r,s,f))(i) = (Powerseq (r,f))(Ln f-i) $\hat{\cdot}$ (Powerseq (s,f))(i) .

Proof. By (4.65) and the definitions of Mult and Reverse. ||

4.67) r and s are fractions & f is a sequence & Ln f > 1 ⟶ $(\text{Powerseq}\ (r,f))(\text{Ln}\ f\text{-}1)\ \hat{-}\ (\text{Powerseq}\ (s,f))(\text{Ln}\ f\text{-}1)$ $= (r \hat{-} s)\ \hat{\cdot}\ (\hat{\sum}\ \text{Telseq}\ (r,s,f))(\text{Ln}\ f\text{-}1)$.

Proof. Prove by bounded induction on i that if $1 \le i \le \text{Ln}\ f\text{-}1$, then $(r \hat{-} s)\ \hat{\cdot}\ (\hat{\sum}\ \text{Telseq}\ (r,s,f))(i)$ $= (\text{Powerseq}\ (r,f))(\text{Ln}\ f)\ \hat{-}\ s\ \hat{\cdot}\ ((\text{Telseq}\ (r,s,f))(i))$. For $i = \text{Ln}\ f\text{-}1$, this is the desired result in light of (4.66). ||

Another definition, and a generalized triangle inequality:

4.68) Def Absset u = a ⟷ u is a sequence of fractions & a is a set & ∀r (r ∈ a ⟷ ∃s(s ∈ Ran u & r = |s|)), otherwise a = 1 .

4.69) u is a sequence of fractions ⟶ $|(\hat{\sum} u)(\text{Ln}\ u)|\ \hat{\le}\ (\text{Ln}\ u)\hat{}\ \hat{\cdot}\ \hat{\text{Max}}\ (\text{Absset}\ u)$.

Proof. By (4.37), (4.32), (4.31), and bounded induction, if $1 \le i \le \text{Ln}\ u$, then $|(\hat{\sum} u)(i)|\ \hat{\le}\ \hat{i}\ \hat{\cdot}\ \hat{\text{Max}}\ (\text{Absset}\ u)$. ||

We prove now that

$$|(a_1+a_2r+\ldots+a_{n+1}r^n) - (b_1+b_2r+\ldots+b_{n+1}r^n)|$$

$$\le (n+1) \cdot \text{Max}\ \{|a_i-b_i|\} \cdot (|r|^n+1) ,$$

and that

$$|(b_1+b_2r+\ldots+b_{n+1}r^n) - (b_1+b_2s+\ldots+b_{n+1}s^n)|$$
$$\leq (n+1)^2 \cdot \text{Max}\,\{|b_i|\} \cdot |r-s| \cdot (|r|^n+|s|^n+1) .$$

These two results are then easily combined via the triangle inequality to give (4.72).

4.70) f and g are sequences of fractions & Ln f = Ln g &

r is a fraction $\longrightarrow$ |Polyvalue (f,r) $\hat{-}$ Polyvalue (g,r)|

$\hat{\leq}$ (Ln f)$\hat{}$ $\hat{\cdot}$ $\hat{\text{Max}}$ (Absset(Subt(f,g)))$\hat{\cdot}$((Powerseq(|r|,f))(Ln f)$\hat{+}\hat{1}$) .

Proof. By (4.62), the left side is |Polyvalue (Subt(f,g),r)|. By (4.69) and the definition of Polyvalue, this is

$\hat{\leq}$ (Ln f)$\hat{}$ $\hat{\cdot}$ $\hat{\text{Max}}$ (Absset(Mult(Subt(f,g),(Powerseq(r,f))[1,Ln f]))) .

But every term in the sequence Mult(Subt(f,g),(Powerseq(r,f))[1,Ln f]) is, in absolute value,

$\hat{\leq}$ $\hat{\text{Max}}$(Absset(Subt(f,g))) $\hat{\cdot}$ $\hat{\text{Max}}$(Absset((Powerseq(r,f))[1,Ln f])) ,

and the latter factor is either (Powerseq(|r|,f))(Ln f) or $\hat{1}$, depending on whether |r| $\hat{\geq}$ $\hat{1}$ or |r| $\hat{<}$ $\hat{1}$. ||

4.71) g is a sequence of fractions & r and s are fractions $\longrightarrow$

|Polyvalue (g,r) $\hat{-}$ Polyvalue (g,s)|

$\hat{\leq}$ (Ln g)$\hat{}$ $\hat{\cdot}$ (Ln g)$\hat{}$ $\hat{\cdot}$ $\hat{\text{Max}}$ (Absset g) $\hat{\cdot}$ |r$\hat{-}$s|

$\hat{\cdot}$ ((Powerseq (|r|,g))(Ln g) $\hat{+}$ (Powerseq (|s|,g))(Ln g) $\hat{+}$ $\hat{1}$) .

Proof. By (4.63) and (4.69), the left side is

$\hat{\leq}$ (Ln g)$\hat{}$ $\hat{\cdot}$ $\hat{\text{Max}}$(Absset(Mult(g,

Subt((Powerseq(r,g))[1,Ln g],(Powerseq(s,g))[1,Ln g])))) .

This $\hat{\text{Max}}$ is certainly $\hat{\le}$ the product of $\hat{\text{Max}}$ (Absset g) with $\hat{\text{Max}}$ (Absset(Subt((Powerseq (r,g))[1,Ln g], (Powerseq(s,g))[1,Ln g]))) . Suppose $1 \le i \le$ Ln g . Then the i^{th} term of the sequence Subt(...,...) appearing above is (Powerseq (r,g))(i) $\hat{-}$ (Powerseq (s,g))(i) , which by the telescoping property equals $(r \hat{-} s) \,\hat{\cdot}\, (\hat{\sum}$ Telseq (r,s,g[1,i]))(i-1) . By (4.69) again and (4.66), this last quantity is, in absolute value, at most $|r \hat{-} s|$ times $(\text{Ln } g)^{\wedge}$ times $|(\text{Powerseq } (r,g))(i-j) \,\hat{\cdot}\, (\text{Powerseq } (s,g))(j)|$ for some j with $1 \le j \le i-1$, and this last factor can be no larger than (Powerseq (|r|,g))(Ln g) $\hat{+}$ (Powerseq (|s|,g))(Ln g) $\hat{+}\ \hat{1}$. Thus $\hat{\text{Max}}$ (Absset(Subt(...,...)))

$\hat{\le}$ $(\text{Ln } g)^{\wedge} \,\hat{\cdot}\, |r \hat{-} s| \,\hat{\cdot}\, ((\text{Powerseq}(|r|,g))(\text{Ln } g) \hat{+} (\text{Powerseq}(|s|,g))(\text{Ln } g) \hat{+} \hat{1})$,

as needed to complete the proof of (4.71). ||

4.72) f and g are sequences of fractions & Ln f = Ln g & r and s are fractions $\longrightarrow$ |Polyvalue (f,r) $\hat{-}$ Polyvalue (g,s)|

$\hat{\le}$ $((\text{Ln } f)^{\wedge} \,\hat{\cdot}\, \hat{\text{Max}}\ (\text{Absset}(\text{Subt}(f,g))) \,\hat{\cdot}\, ((\text{Powerseq}(|r|,f))(\text{Ln } f) \hat{+} \hat{1}))$

$\hat{+}$ $((\text{Ln } f)^{\wedge} \,\hat{\cdot}\, (\text{Ln } f)^{\wedge} \,\hat{\cdot}\, \hat{\text{Max}}\ (\text{Absset } g) \,\hat{\cdot}\, |r \hat{-} s|$

$\hat{\cdot}$ $((\text{Powerseq}(|r|,f))(\text{Ln } f) \hat{+} (\text{Powerseq}(|s|,f))(\text{Ln } f) \hat{+} \hat{1}))$. ||

The final result of this section is that beyond some point the values of a polynomial do not change sign. It is convenient to consider monic polynomials first.

4.73) Def f is monic $\longleftrightarrow$ f is a polynomial & f(Ln f) = $\hat{1}$.

4.74) Def Fixsign f = r $\longleftrightarrow$ f is monic & r = $\hat{\text{Max}}$ (Absset f) $\hat{\cdot}$ $(\text{Deg } f)^{\wedge}$ $\hat{+}\hat{1}$, otherwise r = 0 .

4.75) f is monic & $r \mathbin{\hat{\geq}}$ Fixsign f $\longrightarrow$ Polyvalue (f,r) $\hat{>} \hat{0}$.

Proof. Let n be Ln f , so Deg f = n-1 . If $r \mathbin{\hat{\geq}}$ Fixsign f, then $r \mathbin{\hat{\geq}} \hat{1}$ and also $r \mathbin{\hat{>}} |f(i)| \mathbin{\hat{\cdot}} \widehat{(n-1)}$ for every i with $1 \leq i \leq n$. The sequence of fractions (Powerseq (r,f))[1,n] is nondecreasing, so if $1 \leq i < n$, then

$$|(\mathrm{Powerseq}(r,f))(i) \mathbin{\hat{\cdot}} f(i)| \mathbin{\hat{\cdot}} \widehat{(n-1)} \mathbin{\hat{<}} (\mathrm{Powerseq}(r,f))(i) \mathbin{\hat{\cdot}} r \mathbin{\hat{\leq}} (\mathrm{Powerseq}(r,f))(n) .$$

Since f is monic, this inequality can be rewritten as follows:
if $1 \leq i < n$, then

$$|(\mathrm{Mult}\ (f,(\mathrm{Powerseq}(r,f))[1,n]))(i)| \mathbin{\hat{\cdot}} \widehat{(n-1)}$$
$$\mathbin{\hat{<}} (\mathrm{Mult}\ (f,(\mathrm{Powerseq}(r,f))[1,n]))(n) ;$$

that is, the n^{th} term of the sequence Mult(f,(Powerseq(r,f))[1,n]) is larger than $\widehat{(n-1)}$ times the absolute value of any one of the other n-1 terms. From (4.69) it follows that the n^{th} term is greater than the absolute value of the sum of all the other terms, and therefore that the sum of all n terms is positive. This sum is precisely Polyvalue (f,r) . ||

4.76) f is monic & $r \mathbin{\hat{\leq}} \hat{-}$Fixsign f $\longrightarrow$

(2|Deg f $\longrightarrow$ Polyvalue (f,r) $\hat{>} \hat{0}$) & (2∤Deg f $\longrightarrow$ Polyvalue (f,r) $\hat{<} \hat{0}$).

Proof. The inequalities in the proof of (4.75) remain true in absolute value, but now the sequence (Powerseq(r,f))[1,n] is alternating in sign. The sign of Polyvalue (f,r) is the same as the sign of the highest term, which depends on whether n is odd or even. ||

Generalizing (4.75) and (4.76) to the case of non-monic polynomials is easy once we have the following definition:

4.77) Def Normalize $f = g \longleftrightarrow f$ and g are polynomials & $\mathrm{Ln}\ f = \mathrm{Ln}\ g$ &
$\forall i\ (1 \leq i \leq \mathrm{Ln}\ f \longrightarrow g(i) = f(i) \hat{/} f(\mathrm{Ln}\ f))$, otherwise $g = 1$.

4.78) f is a polynomial & $f \neq 0 \longrightarrow$ Normalize f is monic &
$\forall r$ (r is a fraction $\longrightarrow$ Polyvalue (Normalize f,r) =
Polyvalue $(f,r) \hat{/} f(\mathrm{Ln}\ f))$. ||

In particular, (4.78) says that Polyvalue (Normalize f,r) has the same sign as Polyvalue (f,r) is $f(\mathrm{Ln}\ f)$ is positive and the opposite sign if $f(\mathrm{Ln}\ f)$ is negative. Hence there are four cases, depending on the sign of $f(\mathrm{Ln}\ f)$ and the parity of Deg f .

4.79) f is a polynomial & $f(\mathrm{Ln}\ f) \hat{>} \hat{0}$ & $2 \mid$ Deg $f \longrightarrow$
($r \hat{\geq}$ Fixsign (Normalize f) $\longrightarrow$ Polyvalue $(f,r) \hat{>} \hat{0}$) &
($r \hat{\leq} \hat{-}$Fixsign (Normalize f) $\longrightarrow$ Polyvalue $(f,r) \hat{>} \hat{0}$) . ||

4.80) f is a polynomial & $f(\mathrm{Ln}\ f) \hat{>} \hat{0}$ & $2 \nmid$ Deg $f \longrightarrow$
($r \hat{\geq}$ Fixsign (Normalize f) $\longrightarrow$ Polyvalue $(f,r) \hat{>} \hat{0}$) &
($r \hat{\leq} \hat{-}$Fixsign (Normalize f) $\longrightarrow$ Polyvalue $(f,r) \hat{<} \hat{0}$) . ||

4.81) f is a polynomial & $f(\mathrm{Ln}\ f) \hat{<} \hat{0}$ & $2 \mid$ Deg $f \longrightarrow$
($r \hat{\geq}$ Fixsign (Normalize f) $\longrightarrow$ Polyvalue $(f,r) \hat{<} \hat{0}$) &
($r \hat{\leq} \hat{-}$Fixsign (Normalize f) $\longrightarrow$ Polyvalue $(f,r) \hat{<} \hat{0}$) . ||

4.82) f is a polynomial & f(Ln f) $\hat{<}$ $\hat{0}$ & $2 \nmid$ Deg f $\longrightarrow$

(r $\hat{\geq}$ Fixsign (Normalize f) $\longrightarrow$ Polyvalue (f,r) $\hat{<}$ $\hat{0}$ &

(r $\hat{\leq}$ $\hat{-}$Fixsign (Normalize f) $\longrightarrow$ Polyvalue (f,r) $\hat{>}$ $\hat{0}$) . ||

§5. *A Theory with Real Numbers*

The primary objective of this section is the introduction of real numbers into our predicative theory. Using an observation of Nelson's about the unprovability of exponentiability, we first adjoin an axiom asserting the existence of "infinite" (nonexponentiable) numbers. Once we have infinite numbers, we have infinite fractions, infinitesimals, and a relation of infinite closeness; using the results of §3, we introduce a two-sorted theory in which the individuals of the second sort are the equivalence classes of finite fractions modulo this relation. These equivalence classes are our real numbers, and we show that they satisfy the axioms for real closed fields.

Nonexponentiable numbers

As Nelson points out in [1], one cannot prove in $\mathcal{Q}^0$ the formula $\forall k \varepsilon(k)$ asserting that every number is exponentiable; the same is true of the stronger theories $\mathcal{Q}^\mu$. One proof of this fact is sufficiently enlightening to merit a quick sketch here.

Let T be a consistent extension of $\mathcal{Q}^0$. Let ϕ be a new unary predicate symbol, and consider the formula

Fin) $$\phi(0) \;\&\; (\phi(x) \longrightarrow \phi(Sx)) \;.$$

In the theory $T[(\text{Fin})]$ in which (Fin) has been adjoined as a new axiom, we can certainly prove $\phi(0), \phi(1), \phi(2), \ldots,$ but we cannot immediately conclude $\forall x \phi(x)$ since no induction scheme is applicable to formulas involving ϕ . Let $\underline{a}$ be a variable-free term of T .

The intuition of many mathematicians is that $\underline{a}$ represents some number ρ and that there should be a proof of $\phi(\underline{a})$ in at most ρ steps. In [1,§18], however, Nelson cites an example of Simon Kochen to show that this intuition is wrong: there exist variable-free terms $\underline{a}$ such that $\phi(\underline{a})$ is not a theorem of $T[(\text{Fin})]$ at all. Nelson proceeds to prove a metatheorem [1,§18, Assertion 1] to the effect that if there *is* a proof in $T[(\text{Fin})]$ of $\phi(\underline{a})$, and if no formula in that proof contains more than τ quantifiers, then this proof must have *at least* a certain number (depending on $\underline{a}$ and τ) of formulas in it. As a consequence [1,§18, Assertion 2], no inductive formula of $T[(\text{Fin})]$ is stronger than $\phi(x)$ and respects exponentiation: if there were such a formula, one could use it to give short proofs of $\phi(\underline{a})$.

Now let T be Q^{μ}, and suppose $\forall k \varepsilon(k)$ is a theorem of T. Consider $T[(\text{Fin})]$. Write $\phi^1(x), \phi^2(x), \ldots$ for the formulas $\mathbb{C}^1[x], \mathbb{C}^2[x], \ldots$ (see §§1-2), where $\mathbb{C}[x]$ is $\phi(x)$. Since ϕ is inductive, it follows from Metatheorem G_{μ} that $\phi^{\mu+4}$ is stronger than ϕ, is hereditary, and respects 0, S, $+$, $\cdot$, $\#$, $\#_1, \ldots, \#_{\mu}$. By Metatheorem H_{μ}, $\phi^{\mu+4}$ respects every bounded function symbol of Q^{μ}; moreover, if $\mathbb{A}$ is a nonlogical axiom of Q^{μ}_{b} (that is, a nonlogical axiom of Q^{μ} other than the defining axiom of an unbounded symbol), then $\mathbb{A}^{\phi^{\mu+4}}$ is a theorem of $T[(\text{Fin})]$. Hence $\phi^{\mu+4}$ defines an interpretation of Q^{μ}_{b} in $T[(\text{Fin})]$.

The theorem $\forall k \varepsilon(k)$ of T implies that for every x and k there is a sequence u of length k such that $u(1) = x$ and $\forall i\ (1 \leq i < k \longrightarrow u(i+1) = u(i) \cdot x)$ -- a sequence whose terms are the first k powers of x . The existence of such a sequence is therefore a theorem of $\mathcal{Q}_b^\mu$, since $\mathcal{T}$ is just an extension of definitions of $\mathcal{Q}_b^\mu$. By the interpretation theorem, the relativization by $\phi^{\mu+4}$ of this theorem is a theorem of $\mathcal{T}[(\text{Fin})]$. Arguing in $\mathcal{T}[(\text{Fin})]$, then, if x and k satisfy $\phi^{\mu+4}$, then so does the sequence u , and because $\phi^{\mu+4}$ is hereditary, so does the last term of that sequence -- namely $x \wedge n$. Thus $\phi^{\mu+4}$ is an inductive formula of $\mathcal{T}[(\text{Fin})]$ that is stronger than ϕ and respects exponentiation, contrary to the aforementioned result of [1].

Arithmetic with infinitesimals

To $\mathcal{Q}^\mu$ adjoin a new constant symbol N and the axiom

5.1) $$\text{A}x \neg \varepsilon(\text{N}) ,$$

forming the theory $\tilde{\mathcal{Q}}^\mu$. This theory is consistent by the result just noted. Observe that since N is not exponentiable, Log N is exponentiable exactly once, Log Log N exactly twice, and so on; hence if we make the definitions

5.2)
$$\text{Def } \text{U}_1 = \text{Setlog N} ,$$
$$\text{Def } \text{U}_2 = \text{Setlog (Log N)} ,$$
$$\text{Def } \text{U}_3 = \text{Setlog (Log Log N)} ,$$
$$\vdots$$

it follows that for $\nu = 1,2,\ldots$ we have $\varepsilon_\nu(x) \longrightarrow x \in \text{U}_\nu$ and $x \in \text{U}_\nu \longrightarrow \varepsilon_{\nu-1}(x)$ (ε_0 means ε) .

Fix ν with $1 \leq \nu < \mu$, so that by §2, ε_ν respects 0, S, $+$, $\cdot$, and $\#$ in $\mathcal{Q}^\mu$ (hence in $\tilde{\mathcal{Q}}^\mu$).

5.3) Def r is limited $\longleftrightarrow$ r is a fraction & ε_ν(Numer [r]).

5.4) Def r is unlimited $\longleftrightarrow$ r is a fraction & $\neg$(r is limited).

The limited fractions are the ones we think of as "finite". A few obvious theorems:

5.5) $|r| \mathrel{\hat{\leq}} |s|$ & s is limited $\longrightarrow$ r is limited.

Proof. Since ε_ν is hereditary, it suffices to observe that if $|r| \mathrel{\hat{\leq}} |s|$, then Numer [r] $\leq$ Numer [s]. ‖

5.6) r is a fraction $\longrightarrow$ (r is limited $\longleftrightarrow \exists n(\varepsilon_\nu(n)$ & $|r| \mathrel{\hat{\leq}} \hat{n}$)).

Proof. If r is limited, let n be Numer [r]+1. Conversely, Numer $[\hat{n}] = n$, so if $\varepsilon_\nu(n)$ & $|r| \mathrel{\hat{\leq}} \hat{n}$, then r is limited by (5.5). ‖

5.7) r is a fraction $\longrightarrow$ (r is unlimited $\longleftrightarrow \exists n(\exists\varepsilon_\nu(n)$ & $\hat{n} \mathrel{\hat{\leq}} |r|$)).

Proof. If r is unlimited, let n be Numer [r]-1. The converse again follows from (5.5). ‖

5.8) r is limited $\longleftrightarrow$ $\hat{-}$r is limited. ‖

5.9) r is limited & s is limited $\longrightarrow$ r$\hat{+}$s is limited. ‖

5.10) r is limited & s is limited $\longrightarrow$ r$\hat{\cdot}$s is limited. ‖

The next definition is that of infinitesimal fractions.

5.11) Def r is infinitesimal $\longleftrightarrow$ r is a fraction &
($r \mathrel{\hat{=}} \hat{0} \vee$ Recip r is unlimited).

5.12) $|r| \mathrel{\hat{\leq}} |s|$ & s is infinitesimal $\longrightarrow$ r is infinitesimal.

Proof. If $|r| \mathrel{\hat{\leq}} |s|$, then $|\text{Recip } s| \mathrel{\hat{\leq}} |\text{Recip } r|$. Apply (5.5). ‖

5.13) r is infinitesimal $\longrightarrow$ $\hat{-}r$ is infinitesimal. ‖

5.14) r is infinitesimal & s is infinitesimal $\longrightarrow$ $r \mathbin{\hat{+}} s$ is infinitesimal.

Proof. If $r \mathbin{\hat{+}} s \neq \hat{0}$, then Recip $(r \mathbin{\hat{+}} s)$ is, in absolute value, at least half the smaller of Recip r and Recip s, and is therefore unlimited by (5.10) and the fact that $\hat{2}$ is limited. ‖

5.15) r is limited & s is infinitesimal $\longrightarrow$ $r \mathbin{\hat{\cdot}} s$ is infinitesimal.

Proof. Since r is limited and Recip s is unlimited, it follows from Recip $s \mathrel{\hat{=}} r \cdot$ Recip $(r \mathbin{\hat{\cdot}} s)$ that Recip $(r \mathbin{\hat{\cdot}} s)$ must be unlimited. ‖

5.16) r is a fraction & ¬(r is infinitesimal) & s is unlimited $\longrightarrow$
$r \mathbin{\hat{\cdot}} s$ is unlimited.

Proof. Take reciprocals and apply (5.15). ‖

Now the relation "infinitely close":

5.17) Def $r \sim s$ $\longrightarrow$ r and s are fractions & $r \mathbin{\hat{-}} s$ is infinitesimal.

5.18) r is infinitesimal $\longleftrightarrow$ $r \sim \hat{0}$. ‖

5.19) r is a fraction $\longrightarrow r \sim r$. ||

5.20) $r \sim s \longrightarrow s \sim r$. ||

5.21) $r \sim s \ \& \ s \sim t \longrightarrow r \sim t$.

Proof. By (5.14). ||

5.22) $r \sim s \longrightarrow \hat{-} r \sim \hat{-} s$. ||

5.23) $r_1 \sim s_1 \ \& \ r_2 \sim s_2 \longrightarrow r_1 \hat{+} r_2 \sim s_1 \hat{+} s_2$.

Proof. By (5.14). ||

5.24) r_1 and r_2 are limited & $r_1 \sim s_1 \ \& \ r_2 \sim s_2 \longrightarrow r_1 \hat{\cdot} r_2 \sim s_1 \hat{\cdot} s_2$.

Proof. The difference $r_1 \hat{\cdot} r_2 \hat{-} s_1 \hat{\cdot} s_2$ is equal to $r_1 \hat{\cdot} (r_2 \hat{-} s_2) \hat{+} s_2 \hat{\cdot} (r_1 \hat{-} s_1)$, which is infinitesimal by (5.15) and (5.14). ||

5.25) $r \sim s \ \& \neg(r$ is infinitesimal) $\longrightarrow$ Recip $r \sim$ Recip s .

Proof. The difference Recip $r \hat{-}$ Recip s can be written as $(s \hat{-} r) \hat{\cdot}$ Recip $r \hat{\cdot}$ Recip s . This is the product of an infinitesimal and two limited fractions, and is therefore infinitesimal. ||

5.26) $r_1 \sim s_1 \ \& \ r_2 \sim s_2 \ \& \ r_1$ is limited & $\neg(r_2$ is infinitesimal) $\longrightarrow$
$r_1 \hat{/} r_2 \sim s_1 \hat{/} s_2$.

Proof. By (5.24) and (5.25). ||

Our work in §4 is sufficient to show that polynomials behave nicely as regards the relation $\sim$. For starters, we note that "the sum of finitely many infinitesimals is infinitesimal":

5.27) u is a sequence of fractions & $\varepsilon_\nu(\text{Ln } u)$ & $\forall i(1 \leq i \leq \text{Ln } u \longrightarrow$ $u(i)$ is infinitesimal) $\longrightarrow$ $(\hat{\sum}u)(\text{Ln } u)$ is infinitesimal.

Proof. By (4.69), $|(\hat{\sum}u)(\text{Ln } u)| \hat{\leq} (\text{Ln } u)^\wedge \hat{\cdot} \hat{\text{Max}}$ (Absset u). The first factor on the right side is limited, and the second is infinitesimal. ||

5.28) u is a sequence of fractions & $\varepsilon_\nu(\text{Ln } u)$ & $\forall i$ $(1 \leq i \leq \text{Ln } u \longrightarrow$ $u(i)$ is limited) $\longrightarrow$ $(\hat{\sum}u)(\text{Ln } u)$ is limited. ||

5.29) r is limited & $\varepsilon_\nu(n)$ & u = Powerseq (r,n) $\longrightarrow$ u(Log n) is limited.

Proof. By (5.6), $|r| \hat{\leq} \hat{m}$ for some m with $\varepsilon_\nu(m)$. By bounded induction on n , $|(\text{Powerseq }(r,n))(\text{Log } n)| \hat{\leq} (\text{Explog }(m,n))^\wedge$. But $\varepsilon_\nu(m)$ & $\varepsilon_\nu(n)$ implies $\varepsilon_\nu(\text{Explog }(m,n))$. ||

In conjunction with the above proof, recall that Explog is a bounded function symbol of Q^0 ; since ε_ν respects 0 , S , + , · , and # , it respects every such function symbol by Metatheorem E. In fact, the bound on Explog [1,§19] does involve # ; hence the restriction that ν be strictly smaller than μ is really necessary. Note also that the notions limited, infinitesimal, and $\sim$ are *not* bounded, so we may never use induction directly on any formulas involving these symbols.

5.30) f and g are sequences of fractions & Ln f = Ln g & $\varepsilon_{\nu+1}$(Ln f) & $\forall$i (1 $\leq$ i $\leq$ Ln f $\longrightarrow$ f(i) is limited & f(i) ~ g(i)) & r is limited & r ~ s $\longrightarrow$ Polyvalue (f,r) ~ Polyvalue (g,s) .

Proof. By (4.72), |Polyvalue (f,r) $\hat{-}$ Polyvalue (g,s)|
$\hat{\leq}$ ((Ln f)$\hat{}$ $\hat{\cdot}$ M$\hat{a}$x (Absset(Subt(f,g))) $\hat{\cdot}$ ((Powerseq (|r|,f))(Ln f)$\hat{+}\hat{1}$))
+ ((Ln f)$\hat{}\hat{}$ $\hat{\cdot}$ (Ln f)$\hat{}$ $\hat{\cdot}$ M$\hat{a}$x(Absset g) $\hat{\cdot}$ |r$\hat{-}$s|
$\hat{\cdot}$ ((Powerseq(|r|,f))(Ln f) $\hat{+}$ (Powerseq(|s|,f))(Ln f)$\hat{+}\hat{1}$)) .

The right side is the sum of two terms, each of which is the product of several factors. One factor in each term (namely, M$\hat{a}$x(Absset(Subt(f,g))) in the first term, |r$\hat{-}$s| in the second) is infinitesimal by hypothesis, and the rest are limited (the powers of |r| and |s| are limited by (5.29) since $\varepsilon_{\nu+1}$(Ln f)) . ||

Is the condition $\varepsilon_{\nu+1}$(Ln f) really needed? Indeed it is: if ε_{ν}(n) but $\neg\varepsilon_{\nu+1}$(n) , then $2 \sim 2 + 1/2^n$ but $2^n \not\sim (2+1/2^n)^n$, so the polynomial x^n is not "continuous" in the sense described by (5.30).

It is convenient to know that every limited fraction is infinitely close to some fraction whose numerator and denominator are in U_ν . In fact, more can be said.

5.31) Def r is a U_ν-fraction $\longleftrightarrow$ r is a fraction & Numer r $\in U_\nu$ & Denom r $\in U_\nu$.

5.32) r is limited $\longrightarrow$ $\exists s_1 \exists s_2$ (s_1 and s_2 are U_ν-fractions & $s_1 \hat{<} r \hat{<} s_2$ & $s_1 \sim r \sim s_2$) .

Proof. We may assume $r \geq \hat{0}$. Let k be such that $\varepsilon_\nu(k)$ & $r \hat{<} \hat{k}$, and let n be such that $k \cdot n \in U_\nu$ but $\neg\, \varepsilon_\nu(n)$. Then $[r \hat{\cdot} \hat{n}]$ is $\hat{m}$ for some m , and in fact $m \in U_\nu$ because $m < k \cdot n$. Let s_1 be $(\hat{m} \hat{-} \hat{1}) \hat{/} \hat{n}$, and let s_2 be $(\hat{m} \hat{+} \hat{1}) \hat{/} \hat{n}$. The desired conclusions follow from $\hat{m} \hat{/} \hat{n} \hat{\leq} r \hat{<} (\hat{m} \hat{+} \hat{1}) \hat{/} \hat{n}$ and the fact that $\hat{1} \hat{/} \hat{n}$ is infinitesimal. ||

Note that the preceding proof required the existence of a nonzero infinitesimal; as such, this was the first time we actually used axiom (5.1). That axiom is essential in all that follows.

5.33) $\exists a(a$ is a set & $\forall r(r \in a \longleftrightarrow r$ is a U_ν-fraction)).

Proof. Let $\varepsilon_{\nu-1}(m)$ & $\neg\, m \in U_\nu$, and let $M = \langle 1,m,m \rangle$. Then $\varepsilon_{\nu-1}(M)$, so $\varepsilon(M)$ (because we specified $\nu \geq 1$) , so there is a set z consisting of all numbers from 1 to M . Now if r is a U_ν-fraction, then $r = \langle \text{Sign } r, \text{Numer } r , \text{Denom } r \rangle \leq \langle 1,m,m \rangle = M$; hence we may define a as the set $\{r \in z: r$ is a U_ν-fraction$\}$, which exists by bounded separation. ||

Sums and products of U_ν-fractions need not be U_ν-fractions, of course, and for this reason our principal objects of study are the limited fractions, not the U_ν-fractions. On the other hand, propositions (5.32) and (5.33) give some indication of why the bounded notion of a U_ν-fraction is a useful one: the U_ν-fractions form a set, and this set contains approximations to every limited fraction. These facts will be used frequently. For instance:

5.34) r is limited & $\hat{0} \mathrel{\hat{\leq}} r \longrightarrow \exists s(s \mathbin{\hat{\cdot}} s \sim r)$.

Proof. Assume $r \not\sim \hat{0}$. Let t be the larger of the fractions r and $\hat{1}$, so $r \mathrel{\hat{\leq}} t \mathbin{\hat{\cdot}} t$. By (5.33) and bounded separation, there is a set whose elements are the U_ν-fractions s such that $\hat{0} \mathrel{\hat{\leq}} s \mathrel{\hat{\leq}} t$ & $r \mathrel{\hat{\leq}} s \mathbin{\hat{\cdot}} s$, and this set, like every set of fractions, has a $\hat{<}$-smallest element, say s_0 . By (5.32) there is a U_ν-fraction s_1 such that $s_1 \mathrel{\hat{<}} s_0$ & $s_1 \sim s_0$; it follows that $s_1 \mathbin{\hat{\cdot}} s_1 \mathrel{\hat{<}} r \mathrel{\hat{\leq}} s_0 \mathbin{\hat{\cdot}} s_0$ and $s_1 \mathbin{\hat{\cdot}} s_1 \sim s_0 \mathbin{\hat{\cdot}} s_0$ (by (5.24)), so that $s_0 \mathbin{\hat{\cdot}} s_0 \sim r$, as desired. ||

5.35) f is a polynomial & $2 \nmid \text{Deg } f$ & $\varepsilon_{\nu+1}(\text{Ln } f)$ &
$\forall i\ (1 \leq i \leq \text{Ln } f \longrightarrow f(i)$ is limited) & $f(\text{Ln } f) \not\sim \hat{0} \longrightarrow$
$\exists r$ (r is limited & Polyvalue $(f,r) \sim \hat{0}$) .

Proof. This is basically like (5.34). Since all coefficients of f are limited and the highest-power coefficient is not infinitesimal, all coefficients of Normalize f are limited, and so is t = Fixsign (Normalize f). By (4.80) or (4.82), we have, depending on the sign of $f(\text{Ln } f)$, either Polyvalue $(f,t) \mathrel{\hat{>}} \hat{0}$ & Polyvalue $(f,\hat{-}t) \mathrel{\hat{<}} \hat{0}$ or Polyvalue $(f,t) \mathrel{\hat{<}} \hat{0}$ & Polyvalue $(f,\hat{-}t) \mathrel{\hat{>}} \hat{0}$. In the first case, let r be the smallest fraction in the set of all U_ν-fractions s such that $\hat{-}t \mathrel{\hat{\leq}} s \mathrel{\hat{\leq}} t$ & Polyvalue $(f,s) \mathrel{\hat{>}} \hat{0}$, and let s_1 be a U_ν-fraction such that $s_1 \mathrel{\hat{<}} r$ & $s_1 \sim r$. Then Polyvalue $(f,s_1) \mathrel{\hat{\leq}} \hat{0} \mathrel{\hat{<}}$ Polyvalue (f,r) (except possibly in case $s_1 \mathrel{\hat{<}} \hat{-}t$, in which case we can redefine s_1 to be $\hat{-}t$) , and by (5.30), Polyvalue $(f,s_1) \sim$ Polyvalue (f,r) . Hence Polyvalue $(f,r) \sim \hat{0}$. The proof in the other case is similar. ||

As particularly simple consequences of (5.35), we have the cases in which Deg f is 1,3,... :

5.36) a_0 and a_1 are limited & $a_1 \not\sim \hat{0} \longrightarrow$

$\exists r$ (r is limited & $a_0 \hat{+} a_1 \hat{\cdot} r \sim \hat{0}$) . ||

a_0, a_1, a_2, and a_3 are limited & $a_3 \not\sim \hat{0} \longrightarrow$

$\exists r$ (r is limited & $a_0 \hat{+} a_1 \hat{\cdot} r \hat{+} a_2 \hat{\cdot} r \hat{\cdot} r \hat{+} a_3 \hat{\cdot} r \hat{\cdot} r \hat{\cdot} r \sim \hat{0}$. ||

⋮

The two-sorted theory R_0

The preceding results show that the equivalence classes of limited fractions modulo the relation ~ behave very much like real numbers. The discussion of many-sorted theories in §3 provides the necessary tools for handling these equivalence classes and unifying our presentation.

To this point, we have been working in (an extension by definitions of) a theory $\tilde{Q}^{\mu}$ with only one sort, say n ("numbers"). Let us now adjoin to $\tilde{Q}^{\mu}$ a new sort r ("real numbers"); we shall use lower-case Greek letters for variables of sort r . Also adjoin a new binary predicate symbol ϵ of type (n, r) and three new nonlogical axioms:

5.37) Ax $\exists r$ (r is a fraction & r is limited & $\forall s$ ($s \in \alpha \longleftrightarrow s \sim r$)) ;

5.38) Ax r is a fraction & r is limited $\longrightarrow \exists \alpha (r \in \alpha)$;

5.39) Ax $r \in \alpha$ & $r \in \beta \longrightarrow \alpha =_r \beta$.

Call the resulting two-sorted theory $R_0^{\mu\nu}$. The superscripts, which we shall usually omit, remind us of the dependence on μ (the number of hypersmashes available in $\tilde{Q}^{\mu}$) and ν (the level of exponentiability used in defining "limited"); the subscript indicates that R_0 is the first in a chain of increasingly powerful theories to be developed in this section and the next.

Note that every symbol or axiom of $\tilde{Q}^{\mu}$ is a symbol or axiom of sort n in R_0 . In particular, there is in $\tilde{Q}^{\mu}$ a symbol ϵ , so there are *two* symbols ϵ in R_0 : the familiar ϵ from $\tilde{Q}^{\mu}$ of type (n,n) and the new symbol of type (n,r) . This is the first of several occasions on which we shall use one written symbol for function or predicate symbols of two or more different types. No ambiguity should arise as long as we always make sure we can recognize the sort of a term. This principle even allows dropping the sort-subscripts from the equality symbols $=_n$ and $=_r$.

Axioms (5.37)-(5.39) are exactly of the form (3.11)-(3.13), where $\mathbb{A}[\underline{x}]$ is the formula "$\underline{x}$ is a fraction & $\underline{x}$ is limited." Moreover, (3.6)-(3.10) are theorems of $\tilde{Q}^{\mu}$ for this $\mathbb{A}[\underline{x}]$. (Strictly speaking, to satisfy (3.7) we should first change the definition of $\sim$ so that it applies only to limited fractions; actually, though, this is irrelevant since such a change would not affect (5.37)-(5.39).) Therefore, by the general result of §3, R_0 is interpretable in $\tilde{Q}^{\mu}$. The interpretation I_0 is such that $I_0(n)$ and $I_0(r)$ are both n (of course); $U_n x \longleftrightarrow x = x$; $U_r x \longleftrightarrow x$ is a fraction & x is limited; $\underline{u}_{I_0}$ is $\underline{u}$ for every function or predicate symbol $\underline{u}$ of $\tilde{Q}^{\mu}$; and $(=_r)_{I_0}$ and $(\epsilon)_{I_0}$ are both $\sim$.

Mathematics in R_0

Having established the interpretability of R_0 in $\tilde{\mathfrak{Q}}^{\mu}$, we are free to work in R_0. The first order of business is transferring basic notions like addition and the order relation from fractions to real numbers via the axioms (5.37)-(5.39).

5.40) $$\text{Def } \tilde{0} =_{\mathcal{r}} \alpha \longleftrightarrow \hat{0} \in \alpha \ .$$

The existence condition follows from (5.38) and the uniqueness condition from (5.39). Likewise for the following definition:

5.41) Def $\tilde{r} =_{\mathcal{r}} \alpha \longleftrightarrow r$ is a fraction & r is limited & $r \in \alpha$, otherwise $\alpha =_{\mathcal{r}} \tilde{0}$.

We agree to abbreviate $\tilde{\hat{n}}$ to $\tilde{n}$. Though technically ambiguous, this notation is consistent with (5.40) and should not result in confusion.

5.42) $$\text{Def } \alpha_1 + \alpha_2 =_{\mathcal{r}} \beta \longleftrightarrow \exists r_1 \exists r_2 \, (r_1 \in \alpha_1 \ \& \ r_2 \in \alpha_2 \ \& \ r_1 \hat{+} r_2 \in \beta) \ .$$

This is a definition of the form (3.15); the preliminary result (3.14) is exactly (5.23). Therefore, as described in §3, (5.42) is a legitimate defining axiom for the "induced" function symbol + of type $(\mathcal{r}, \mathcal{r}; \mathcal{r})$. Similar remarks apply to the next several definitions.

5.43) $$\text{Def } -\alpha =_{\mathcal{r}} \beta \longleftrightarrow \exists r \, (r \in \alpha \ \& \ \hat{-} r \in \beta) \ .$$

We shorten $\alpha + (-\beta)$ to $\alpha - \beta$.

5.44) $$\text{Def } \alpha_1 \cdot \alpha_2 =_{\mathcal{r}} \beta \longleftrightarrow \exists r_1 \exists r_2 \, (r_1 \in \alpha_1 \ \& \ r_2 \in \beta_2 \ \& \ r_1 \hat{\cdot} r_2 \in \beta) \ .$$

5.45) Def $\alpha_1/\alpha_2 =_r \beta \longleftrightarrow (\alpha_2 \neq_r \tilde{0}$ & $\exists r_1 \exists r_2\ (r_1 \in \alpha_1$ & $r_2 \in \alpha_2$ & $r_1 \hat{/} r_2 \in \beta)) \vee (\alpha_2 =_r \tilde{0}$ & $\beta =_r \tilde{0})$.

For the order relation we must proceed carefully; it seems best to define $\leq$ first.

5.46) Def $\alpha \leq \beta \longleftrightarrow \exists r \exists s\ (r \in \alpha$ & $s \in \beta$ & $r \hat{\leq} s)$.

5.47) Def $\alpha < \beta \longleftrightarrow \alpha \leq \beta$ & $\alpha \neq_r \beta$.

5.48) Def $|\alpha| = \beta \longleftrightarrow \exists r\ (r \in \alpha$ & $|r| \in \beta)$.

Here, finally, is one way of defining the greatest integer function. The reader should have no trouble supplying the appropriate conditions.

5.49) Def $[\alpha] =_r \beta \longleftrightarrow \exists r\ (r \in \alpha$ & $[r] \in \beta)$ & $\neg \exists r\ (r \in \alpha$ & $[r] \hat{-} \hat{1} \in \beta)$.

The interpretation of RCF in R_0

The ordered field axioms (4.12)-(4.15), (4.20)-(4.23), and (4.27)-(4.33) for fractions have obvious counterparts for real numbers. It is surely not necessary to list them all; we record two sample proofs (for the theorems corresponding to (4.20) and (4.30)).

5.50) $$\alpha + \beta = \beta + \alpha \ .$$

Proof. There exist limited fractions r_1 , s_1 , r_2 , and s_2 such that $r_1 \in \alpha$, $s_1 \in \beta$, $r_1 \hat{+} s_1 \in \alpha+\beta$, $r_2 \in \alpha$, $s_2 \in \beta$, and $s_2 \hat{+} r_2 \in \beta+\alpha$. Now $r_1 \sim r_2$ and $s_1 \sim s_2$ by (5.37), and therefore $s_2 \hat{+} r_2 = r_2 \hat{+} s_2 \sim r_1 \hat{+} s_1$ by (4.20) and (5.23). It follows from (5.37) again that $s_2 \hat{+} r_2 \in \alpha+\beta$; then $\alpha+\beta = \beta+\alpha$ by (5.39). ||

5.51) $$\alpha \neq \tilde{0} \longrightarrow \alpha \cdot (\tilde{1}/\alpha) = \tilde{1} .$$

Proof. Take $r \in \alpha$; since $\alpha \neq \tilde{0}$, r is not infinitesimal and Recip $r = \hat{1}\hat{/}r$ is limited; in fact, $\hat{1}\hat{/}r \in \tilde{1}/\alpha$ (because $\hat{1} \in \tilde{1}$ and $r \in \alpha$) . But $r\hat{\cdot}(\hat{1}\hat{/}r) = \hat{1}$ by (4.30), so, by the definition (5.44), $\alpha \cdot (\tilde{1}/\alpha) = \tilde{1}$. ||

The absolute value properties (4.35)-(4.37) also hold for real numbers. More importantly, we have the following versions of theorems (5.34) and (5.36):

5.52) $$\alpha \geq \tilde{0} \longrightarrow \exists \beta (\beta \cdot \beta = \alpha) . \quad ||$$

5.53) $$\alpha_1 \neq \tilde{0} \longrightarrow \exists \beta (\alpha_0 + \alpha_1 \cdot \beta = \tilde{0}) . \quad ||$$

$$\alpha_3 \neq \tilde{0} \longrightarrow \exists \beta (\alpha_0 + \alpha_1 \cdot \beta + \alpha_2 \cdot \beta \cdot \beta + \alpha_3 \cdot \beta \cdot \beta \cdot \beta = \tilde{0}) . \quad ||$$

$$\vdots$$

In other words, all the real closed field axioms are theorems about real numbers in R_0 . This is the essence of

Metatheorem K. The theory *RCF* of real closed (ordered) fields is interpretable in R_0 .

Proof. The theory *RCF* has one sort, say σ , and nonlogical symbols 0 , 1 , $+$, $\cdot$, and $<$. To define the interpretation I , let $I(\sigma)$ be $\mathcal{r}$, define $U_\sigma \alpha \longleftrightarrow \alpha =_{\mathcal{r}} \alpha$, and let $(=_\sigma)_I$, 0_I , 1_I , $+_I$, $\cdot_I$, and $<_I$ be the symbols $=_{\mathcal{r}}$, $\tilde{0}$, $\tilde{1}$, $+$, $\cdot$, and $<$ of R_0 . Conditions (3.1)-(3.5) in the definition of an interpretation are all automatic since U_σ holds universally and $(=_\sigma)_I$ is $=_{\mathcal{r}}$;

that the interpretations of the nonlogical axioms of RCF are theorems of R_0 has already been noted. ||

6. *An Expanded Theory*

In the one-sorted theory Q^{μ} , objects of any kind whatsoever -- primes, ordered pairs, fractions, infinitesimals -- can serve as elements of a set, values of a function, or terms of a sequence. In the two-sorted theory R_0 , on the other hand, there is a new kind of object -- real numbers -- for which this is not the case. This section is concerned with successive refinements of R_0 in which one can study sets of real numbers, functions from the real numbers to the real numbers, sequences of real numbers, and even sequences of sets and sequences of functions. In principle, the methods used involve nothing more than further applications of the equivalence-class construction of §3; it seems advisable, though, to vary the approach slightly in order to make the notation more appealing. At the end of the section we check that the most complex of the theories constructed, a theory called R_4 , is interpretable in $\tilde{Q}^{\mu}$.

Sets of real numbers

We first discuss a theory $R_1^{\mu\nu}$, or just R_1 , designed to accommodate sets of real numbers. For these objects we have a new (third) sort s . We shall use lower-case Latin letters with the subscript s for variables of sort s ; corresponding conventions will be employed when we discuss other sorts later in this section. In R_1 there is a function symbol of type $(r;s)$ which we shall also denote by s ; if a is a set of fractions (in the sense of Q^{μ}) , then sa will be the corresponding set of real numbers. There is also a predicate symbol ϵ of type (r,s) with the obvious intended meaning.

The axioms of R_1 , in addition to those of R_0 , include three dealing with sets. The first describes exactly how the set of real numbers a corresponds to the set of fractions $\mathfrak{s}a$:

6.1) Ax a is a set of fractions $\longrightarrow \forall\alpha(\alpha \in \mathfrak{s}a \longleftrightarrow \exists r\ (r \in \alpha\ \&\ r \in a))$.

Observe that the first $\in$ here is the new one, of type (r, s) ; the second is of type (n, r) , and the third is our old friend from Q^μ . The remaining two axioms state that every set x_s corresponds to some set of fractions a and that a set is uniquely determined by its elements.

6.2) Ax $\exists a$ (a is a set of fractions & $\mathfrak{s}a =_s x_s$) .

6.3) Ax $\forall\alpha\ (\alpha \in x_s \longleftrightarrow \alpha \in y_s) \longrightarrow x_s =_s y_s$.

These axioms give us many simple sets. The simplest is undoubtedly $\mathfrak{s}0$.

6.4) $\forall\alpha(\alpha \notin x_s) \longleftrightarrow x_s = \mathfrak{s}0$.

Proof. Since 0 is the empty set in Q^μ , $\forall\alpha(\alpha \notin \mathfrak{s}0)$ follows from (6.1). Conversely, if $\forall\alpha(\alpha \notin x_s)$, then $\forall\alpha(\alpha \in x_s \longleftrightarrow \alpha \in \mathfrak{s}0)$, so $x_s = \mathfrak{s}0$ by (6.3). ||

Given a real number α , we can find a fraction r such that $r \in \alpha$ by (5.37). In Q^μ , we can form the singleton $\{r\}$; the set $\mathfrak{s}\{r\}$ then has α as its only member. If we had chosen a different r , $\mathfrak{s}\{r\}$ would still be the same by extensionality (6.3).

In other words, we have checked the existence and uniqueness conditions for the following definition.

6.5) Def $\{\alpha\} =_{\mathfrak{s}} x_{\mathfrak{s}} \longleftrightarrow \exists r(r \in \alpha \ \& \ \mathfrak{s}\{r\} = x_{\mathfrak{s}})$.

6.6) $$\beta \in \{\alpha\} \longleftrightarrow \beta = \alpha \ . \ \|$$

A similar method can be used to define closed intervals.

6.7) Def $[\alpha,\beta] =_{\mathfrak{s}} x_{\mathfrak{s}} \longleftrightarrow (\alpha = \beta \ \& \ x_{\mathfrak{s}} = \{\alpha\}) \vee$
$(\alpha \neq \beta \ \& \ \exists r \exists s \exists a(r \in \alpha \ \& \ s \in \beta \ \& \ a$ is a set &
$\forall t(t \in a \longleftrightarrow t$ is a U_ν-fraction & $r \ \hat{\leq} \ t \ \hat{\leq} \ s)$ &
$\mathfrak{s}a = x_{\mathfrak{s}}))$.

6.8) $\gamma \in [\alpha,\beta] \longleftrightarrow \alpha \leq \gamma \leq \beta$. $\|$

Of course, a set of fractions may contain fractions that are unlimited. Since such fractions do not represent real numbers, however, it is only the limited fractions in a that have any bearing on $\mathfrak{s}a$. In particular, if a is the set of all U_ν-fractions, then every real number is represented by some element of a , so $\mathfrak{s}a$ is the set of all real numbers.

6.9) Def $(-\infty,\infty) =_{\mathfrak{s}} x_{\mathfrak{s}} \longleftrightarrow \exists a(a$ is a set &
$\forall r(r \in a \longleftrightarrow r$ is a U_ν-fraction) & $\mathfrak{s}a = x_{\mathfrak{s}})$.

6.10) $\forall\alpha(\alpha \in (-\infty,\infty))$. $\|$

The reader should have no trouble defining $(-\infty,\beta]$ and $[\alpha,\infty)$.

6.11) Def $x_\delta \subseteq y_\delta \longleftrightarrow \forall\alpha(\alpha \in x_\delta \longrightarrow \alpha \in y_\delta)$.

6.12) Def $x_\delta \cup y_\delta =_\delta z_\delta \longleftrightarrow \exists a \exists b$ (a and b are sets of fractions & $\delta a = x_\delta$ & $\delta b = y_\delta$ & $\delta(a \cup b) = z_\delta$) .

The existence condition for (6.12) follows from (6.2) and the uniqueness condition from (6.3).

6.13) $\alpha \in x_\delta \cup y_\delta \longleftrightarrow \alpha \in x_\delta \vee \alpha \in y_\delta$. ||

Interestingly (or alarmingly, depending on one's point of view), the intersection of two sets of real numbers need not exist as a set of real numbers at all. Let a be the set of all positive U_ν-fractions with denominator 1 , so that δa is the set $\{1,2,3,\ldots\}$ (really $\{\tilde{1},\tilde{2},\tilde{3},\ldots\}$) of all positive integers (in the real numbers). It is easy to see that there is a set b consisting of all fractions of the form $n+1/2^n$ with $1 \leq n \in U_\nu$; then δb is the set of real numbers $\{1\frac{1}{2}, 2\frac{1}{4}, 3\frac{1}{8},\ldots\}$. A real integer m is in δb if and only if $1/2^m$ is infinitesimal -- that is, if and only if $\neg\varepsilon_{\nu+1}(m)$. The elements common to both δa and δb , therefore, are exactly these integers m . But they do not form a set, since very set of real numbers is δc for some set of fractions c and every set of fractions contains a smallest element.

The fact that the sets a and b in the above example contain unlimited elements is unimportant; indeed, by taking reciprocals we can convert the example to one in the unit interval. What is important is that sets of real numbers come from sets of fractions, and unbounded properties (like $\varepsilon_{\nu+1}$) cannot be used in defining sets of fractions. This idea is also at the heart of the following proposition, which asserts that "all sets are closed"; the argument is of the "overspill" variety common in nonstandard analysis.

6.13) $\forall\varepsilon(\varepsilon > \tilde{0} \longrightarrow \exists\beta(|\beta-\alpha| < \varepsilon \;\&\; \beta \in x_{\mathcal{s}})) \longrightarrow \alpha \in x_{\mathcal{s}}$.

Proof. Let a be a set of fractions such that $x_{\mathcal{s}} = \mathcal{s}a$; then $\forall\gamma(\gamma \in x_{\mathcal{s}} \longleftrightarrow \exists r(r \in \gamma \;\&\; r \in a))$. Take $s \in \alpha$, and form the set $\{|r \hat{-} s| : r \in \mathbf{a}\}$ of nonnegative fractions. This set has a smallest element e . The real number $\tilde{e}$ cannot be positive by hypothesis, so e is infinitesimal. That is, some r satisfies $r \in a$ and $r \sim s$. But $r \sim s \;\&\; s \in \alpha$ implies $r \in \alpha$, which together with $r \in a$ implies $\alpha \in x_{\mathcal{s}}$. ||

Functions

Just as sets of fractions give rise to sets of real numbers, certain functions whose domains and ranges are sets of fractions will give rise to functions from the real numbers to the real numbers. To clarify the meaning of "certain", we make the following definition in $\tilde{\mathcal{Q}}^{\mu}$:

6.14) Def f is a real function $\longleftrightarrow$ f is a function & Dom f and Ran f are sets of fractions & $\forall r \forall s(r \in \text{Dom } f \;\&\; s \in \text{Dom } f \;\&$ r is limited & $r \sim s \longrightarrow f(r)$ is limited & $f(r) \sim f(s))$.

Let $\mathcal{R}_2^{\mu\nu}$ be the theory obtained from $\mathcal{R}_1$ by adjoining a fourth sort $\mathcal{f}$ ("functions from the real numbers to the real numbers") ;

a function symbol $\mathfrak{f}$ of type $(\mathfrak{n};\mathfrak{f})$ (whose role will be similar to that of the function symbol $\mathfrak{s}$) ; a predicate symbol of degree 3, · maps · to · , of type $(\mathfrak{f},\mathfrak{r},\mathfrak{r})$; and three new nonlogical axioms:

6.15) Ax f is a real function $\longrightarrow$

$$\forall\alpha(\exists\beta(\mathfrak{f}f \text{ maps } \alpha \text{ to } \beta) \longleftrightarrow \exists r(r \in \alpha \;\&\; r \in \text{Dom } f)) \;\&$$
$$\forall\alpha\forall\beta\forall r(r \in \alpha \;\&\; r \in \text{Dom } f \;\&\; \mathfrak{f}f \text{ maps } \alpha \text{ to } \beta \longrightarrow f(r) \in \beta) .$$

6.16) Ax $\exists g(g$ is a real function & $\mathfrak{f}g =_{\mathfrak{f}} f_{\mathfrak{f}})$.

6.17) Ax $\forall\alpha\forall\beta(f_{\mathfrak{f}}$ maps α to $\beta \longleftrightarrow g_{\mathfrak{f}}$ maps α to $\beta) \longrightarrow$
$f_{\mathfrak{f}} =_{\mathfrak{f}} g_{\mathfrak{f}}$.

Our first theorem of R_2 allows the awkward predicate symbol · maps · to · , introduced here for simplicity of the interpretation in $\widetilde{\mathcal{Q}}^{\mu}$, to be replaced by more convenient notation.

6.18) $f_{\mathfrak{f}}$ maps α to β & $f_{\mathfrak{f}}$ maps α to $\gamma \longrightarrow \beta = \gamma$.

Proof. By (6.16), $f_{\mathfrak{f}}$ is $\mathfrak{f}g$ for some real function g . By (6.15) and hypothesis, some r satisfies $r \in \alpha$ & $r \in$ Dom g . Then, by (6.15) again, we have **g(r)** $\in \beta$ & **g(r)** $\in \gamma$, whence $\beta = \gamma$ by (5.39). ||

6.19) Def $f_{\mathfrak{f}}(\alpha) =_{\mathfrak{r}} \beta \longleftrightarrow f_{\mathfrak{f}}$ maps α to β , otherwise $\beta = \tilde{0}$.

6.20) $\forall\alpha\forall\beta\neg(f_{\mathfrak{f}}$ maps α to $\beta) \longleftrightarrow f_{\mathfrak{f}} = \mathfrak{f}0$.

Proof. Since 0 is the "empty function", so is $\mathfrak{f}0$ by (6.15). That $\mathfrak{f}0$ is the *only* empty function follows from (6.17). ||

6.21) f and g are real functions & $\mathfrak{f}f = \mathfrak{f}g \longrightarrow \mathfrak{s}(\text{Dom } f) = \mathfrak{s}(\text{Dom } g)$.

Proof. Suppose $\alpha \in \mathfrak{s}(\text{Dom } f)$. By (6.1), there exists r such that $r \in \alpha$ & $r \in \text{Dom } f$, so by (6.15), $\exists\beta(\mathfrak{f}f$ maps α to $\beta)$. If $\mathfrak{f}f = \mathfrak{f}g$, it follows that $\exists\beta(\mathfrak{f}g$ maps α to $\beta)$, that $\exists r(r \in \alpha$ & $r \in \text{Dom } g)$, and thus that $\alpha \in \mathfrak{s}(\text{Dom } g)$. Likewise $\alpha \in \mathfrak{s}(\text{Dom } g) \longrightarrow \alpha \in \mathfrak{s}(\text{Dom } f)$, so $\mathfrak{s}(\text{Dom } f) = \mathfrak{s}(\text{Dom } g)$ by (6.3). ||

6.23) Def Dom $f_{\mathfrak{f}} =_{\mathfrak{s}} x_{\mathfrak{s}} \longleftrightarrow \exists g(g$ is a real function & $\mathfrak{f}g = f_{\mathfrak{f}}$ & $x_{\mathfrak{s}} = \mathfrak{s}(\text{Dom } g))$.

6.24) $\alpha \in \text{Dom } f_{\mathfrak{f}} \longleftrightarrow \exists\beta(f_{\mathfrak{f}}$ maps α to $\beta)$. ||

Earlier we saw by example that the intersection of two sets of real numbers need not be a set of real numbers. There is a similar counterexample to the assertion that every function on the real numbers has a zero set. Define f on the positive integers by $f(n) = 1/2^n$, and extend f to the positive U_ν-fractions by piecewise linearity. Then f is a real function, and $\mathfrak{f}f(\alpha) = \tilde{0}$ if and only the greatest integer in α does not satisfy $\varepsilon_{\nu+1}$; hence there does not exist a set of real numbers $x_{\mathfrak{s}}$ such that $\forall\alpha(\alpha \in x_{\mathfrak{s}} \longleftrightarrow \mathfrak{f}f(\alpha) = \tilde{0})$.

The converse to this nontheorem, on the other hand, is true:

6.25) $\exists f_{\mathfrak{f}}(\text{Dom } f_{\mathfrak{f}} = (-\infty,\infty)$ & $\forall\alpha(f_{\mathfrak{f}}(\alpha) = \tilde{0} \longleftrightarrow \alpha \in x_{\mathfrak{s}}))$.

Proof. We may assume that $x_{\mathfrak{s}}$ is nonempty. Let a be a set of fractions such that $\mathfrak{s}a = x_{\mathfrak{s}}$. Define a real function g on U_ν-fractions by letting $g(r)$ be the smallest fraction in the set

$\{|s \hat{-} r| : s \in a\}$, and let $f_{\mathfrak{f}}$ be $\mathfrak{f}g$. Then $\mathrm{Dom}\ f_{\mathfrak{f}} = (-\infty,\infty)$, and $f_{\mathfrak{f}}(\alpha) = \tilde{0}$ if and only if some $r \in \alpha$ is infinitely close to some $s \in a$ -- that is, if and only if $\alpha \in x_{\mathfrak{s}}$. ||

Corresponding to the fact that all sets are closed is the following proposition, whose assertion is that "all functions are continuous".

6.26) $\alpha \in \mathrm{Dom}\ f_{\mathfrak{f}}\ \&\ \varepsilon > \tilde{0} \longrightarrow \exists\delta(\delta > \tilde{0}\ \&\ \forall\beta(\beta \in \mathrm{Dom}\ f_{\mathfrak{f}}\ \&\ |\beta - \alpha| < \delta \longrightarrow |f_{\mathfrak{f}}(\beta) - f_{\mathfrak{f}}(\alpha)| < \varepsilon))$.

Proof. Suppose $\alpha \in \mathrm{Dom}\ f_{\mathfrak{f}}\ \&\ \varepsilon > \tilde{0}$. Take a real function g such that $\mathfrak{f}g = f_{\mathfrak{f}}$; take $r \in \alpha$ such that $r \in \mathrm{Dom}\ g$; and take $e \in \varepsilon$. Then the set $\{|s \hat{-} r| : s \in \mathrm{Dom}\ g\ \&\ |g(s) \hat{-} g(r)|\ \hat{\geq}\ e \hat{/} \hat{2}\}$ has a smallest element t , which cannot be infinitesimal since g is a real function and $e \hat{/} \hat{2}$ is not infinitesimal. Let δ be $\tilde{t}/\tilde{2}$. If $\beta \in \mathrm{Dom}\ f_{\mathfrak{f}}$, then there is some s such that $s \in \beta\ \&\ s \in \mathrm{Dom}\ g$; if $|\beta - \alpha| < \delta$, then $|s \hat{-} r|\ \hat{<}\ t$, whence $|g(s) \hat{-} g(r)|\ \hat{<}\ e \hat{/} \hat{2}$, and thus $|f_{\mathfrak{f}}(\beta) - f_{\mathfrak{f}}(\alpha)| \leq \varepsilon/\tilde{2} < \varepsilon$. ||

Two more appealing facts are that every function on a bounded set is uniformly continuous (6.27) and that every such function attains a maximum (6.28).

6.27) $\mathrm{Dom}\ f_{\mathfrak{f}} \subseteq [\gamma_1,\gamma_2]\ \&\ \varepsilon > \tilde{0} \longrightarrow \exists\delta(\delta > \tilde{0}\ \&\ \forall\alpha\forall\beta(\alpha \in \mathrm{Dom}\ f_{\mathfrak{f}}\ \&\ \beta \in \mathrm{Dom}\ f_{\mathfrak{f}}\ \&\ |\beta - \alpha| < \delta \longrightarrow |f_{\mathfrak{f}}(\beta) - f_{\mathfrak{f}}(\alpha)| < \varepsilon))$.

Proof. Take a real function g such that $\mathfrak{f}g = f_{\mathfrak{f}}$ and such that $\mathrm{Dom}\ g$ contains no unlimited elements, and take $e \in \varepsilon$. For each fraction r in Dom g, let t be as in (6.26), so that $t \hat{\neq} \hat{0}$ and $\forall s(s \in \text{Dom } g \;\&\; |s \hat{-} r| \hat{<} t \longrightarrow |g(s) \hat{-} g(r)| \hat{<} e\hat{/}\hat{2})$. As r ranges through Dom g, these fractions t form a set. Let t_0 be the smallest fraction in this set, and let δ be $\tilde{t}_0/\tilde{2}$. ‖

6.28) $\text{Dom } f_{\mathscr{f}} \subseteq [\gamma_1, \gamma_2] \longrightarrow \exists \alpha(\alpha \in \text{Dom } f_{\mathscr{f}} \;\&\; \forall \beta(\beta \in \text{Dom } f_{\mathscr{f}} \longrightarrow f_{\mathscr{f}}(\beta) \leq f_{\mathscr{f}}(\alpha)))$.

Proof. Let g be a real function such that $\mathscr{f} g = f_{\mathscr{f}}$ and such that Dom g contains no unlimited elements. There is some r in Dom g such that $g(r)$ is largest; let α be $\tilde{r}$. ‖

It is important to understand how boundedness of Dom $f_{\mathscr{f}}$ is used in the above two proofs.

Defining the sum of two functions presents a mild difficulty, since if f and g are real functions with Dom $\mathscr{f} f$ = Dom $\mathscr{f} g$, it does not necessarily follow that Dom f = Dom g. The problem is by no means an insurmountable one, however. First consider the following situation. We have a real function f and a set a of fractions such that $\mathscr{s} a \subseteq \mathscr{s}(\text{Dom } f)$. How can we extend f to a real function f_1 such that Dom f_1 = Dom $f \cup a$ (and therefore necessarily $\mathscr{f} f_1 = \mathscr{f} f$) ? The simplest way is to define $f_1(r)$, for each r in a, to be equal to $f(s)$, where s is the element of Dom f closest to r. Of course, there may be *two* such elements, in which case we must specify which one to choose.

6.29) Def Extend $(f,a) = f_1 \longleftrightarrow$

f is a function & Dom f, Ran f, and a are sets of fractions &

f_1 is a function & Dom f_1 = Dom $f \cup a$ &

$$\forall r(r \in \text{Dom } f_1 \longrightarrow \exists s(s \in \text{Dom } f \ \&\ \forall t(t \in \text{Dom } f \longrightarrow$$
$$|t \hat{-} r| \hat{\geq} |s \hat{-} r|) \ \&\ \forall t(t \in \text{Dom } f \ \&\ |t \hat{-} r| = |s \hat{-} r| \longrightarrow$$
$$s \hat{\leq} t) \ \&\ f_1(r) = f(s))) \text{ , otherwise } f_1 = 0 \text{ .}$$

The uniqueness condition is at least as clear as the definition. Observe that $\text{Card } f_1 \leq \text{Card } f + \text{Card } a$ and $\text{Bd } f_1 \leq \langle f \cup a, f \rangle$, so that the function symbol Extend is bounded.

6.30) f is a real function & a is a set of fractions & $\mathscr{s}\, a \subseteq \mathscr{s}(\text{Dom } f) \longrightarrow \text{Extend }(f,a)$ is a real function & $\text{Dom Extend }(f,a) = \text{Dom } f \cup a$ & $\forall r(r \in \text{Dom } f \longrightarrow (\text{Extend }(f,a))(r) = f(r))$ & $\mathscr{f}\,(\text{Extend}(f,a)) = \mathscr{f} f$.

Proof. Let f_1 be Extend (f,a) . It suffices to show that if r is a limited fraction in a , then $f_1(r) \sim f(s)$ for some (hence every) s in Dom f with $r \sim s$. In fact, $f_1(r)$ is *equal* to $f(s)$, where s is the element of Dom f closest to r , and the assumption $\mathscr{s}\, a \subseteq \mathscr{s}(\text{Dom } f)$ implies that $r \sim s$ for this s . ||

The sum of two functions is now easy to define.

6.31) Def $f_{\mathscr{f}} + g_{\mathscr{f}} =_{\mathscr{f}} h_{\mathscr{f}} \longleftrightarrow \text{Dom } f_{\mathscr{f}} = \text{Dom } g_{\mathscr{f}}$ & $\exists f_0 \exists g_0 \exists h_0 (f_0, g_0,$ and h_0 are real functions & $\mathscr{f} f_0 = f_{\mathscr{f}}$ & $\mathscr{f} g_0 = g_{\mathscr{f}}$ & $\text{Dom } h_0 = \text{Dom } f_0 \cup \text{Dom } g_0$ & $\forall r(r \in \text{Dom } h_0 \longrightarrow h_0(r) = (\text{Extend }(f_0, \text{Dom } g_0))(r) \hat{+} (\text{Extend }(g_0, \text{Dom } f_0))(r)$ & $\mathscr{f} h_0 = h_{\mathscr{f}})$, otherwise $h_{\mathscr{f}} = \mathscr{f} 0$.

6.32) $\mathrm{Dom}\ f_\delta = \mathrm{Dom}\ g_\delta \longrightarrow \mathrm{Dom}\ (f_\delta + g_\delta) = \mathrm{Dom}\ f_\delta\ \&$
$\forall\alpha(\alpha \in \mathrm{Dom}\ f_\delta \longrightarrow (f_\delta + g_\delta)(\alpha) = f_\delta(\alpha) + g_\delta(\alpha))\ .\ \|$

6.33) $\mathrm{Def}\ -f_\delta =_\delta g_\delta \longleftrightarrow \exists f_0 \exists g_0 (f_0$ and g_0 are real functions $\&$
$\delta f_0 = f_\delta\ \&\ \mathrm{Dom}\ g_0 = \mathrm{Dom}\ f_0\ \&\ \forall r(r \in \mathrm{Dom}\ f_0 \longrightarrow g_0(r) =$
$\hat{-} f_0(r))\ \&\ \delta g_0 = g_\delta)\ .$

6.34) $\mathrm{Dom}\ (-f_\delta) = \mathrm{Dom}\ f_\delta\ \&\ \forall\alpha(\alpha \in \mathrm{Dom}\ f_\delta \longrightarrow (-f_\delta)(\alpha) = -(f_\delta(\alpha)))\ .\ \|$

6.35) $\mathrm{Def}\ f_\delta \cdot g_\delta =_\delta h_\delta \longleftrightarrow \mathrm{Dom}\ f_\delta = \mathrm{Dom}\ g_\delta\ \&$
$\exists f_0 \exists g_0 \exists h_0 (f_0$, g_0 , and h_0 are real functions $\&$
$\delta f_0 = f_\delta\ \&\ \delta g_0 = g_\delta\ \&\ \mathrm{Dom}\ h_0 = \mathrm{Dom}\ f_0 \cup \mathrm{Dom}\ g_0\ \&\ \forall r(r \in \mathrm{Dom}\ h_0 \longrightarrow$
$h_0(r) = (\mathrm{Extend}(f_0, \mathrm{Dom}\ g_0))(r)\ \hat{\cdot}\ (\mathrm{Extend}(g_0, \mathrm{Dom}\ f_0))(r))\ \&$
$\delta h_0 = h_\delta)$, otherwise $h_\delta = \delta 0$.

6.36) $\mathrm{Dom}\ f_\delta = \mathrm{Dom}\ g_\delta \longrightarrow \mathrm{Dom}\ (f_\delta \cdot g_\delta) = \mathrm{Dom}\ f_\delta\ \&$
$\forall\alpha(\alpha \in \mathrm{Dom}\ f_\delta \longrightarrow (f_\delta \cdot g_\delta)(\alpha) = f_\delta(\alpha) \cdot g_\delta(\alpha))\ .\ \|$

6.37) $\mathrm{Def}\ \alpha \cdot f_\delta =_\delta g_\delta \longleftrightarrow \exists f_0 \exists g_0 \exists s (f_0$ and g_0 are real functions $\&$
$\delta f_0 = f_\delta\ \&\ s \in \alpha\ \&\ \mathrm{Dom}\ f_0 = \mathrm{Dom}\ g_0\ \&\ \forall r(r \in \mathrm{Dom}\ f_0 \longrightarrow$
$g_0(r) = s \hat{\cdot} f_0(r))\ \&\ \delta g_0 = g_\delta)\ .$

6.38) $\mathrm{Dom}\ (\alpha \cdot f_\delta) = \mathrm{Dom}\ f_\delta\ \&\ \forall\beta(\beta \in \mathrm{Dom}\ f_\delta \longrightarrow (\alpha \cdot f_\delta)(\beta) = \alpha \cdot (f_\delta(\beta)))\ .\ \|$

6.39) $\mathrm{Def}\ \alpha / f_\delta =_\delta g_\delta \longleftrightarrow \forall\beta(\beta \in \mathrm{Dom}\ f_\delta \longrightarrow f_\delta(\beta) \neq \tilde{0})\ \&$
$\exists f_0 \exists g_0 \exists s (f_0$ and g_0 are real functions $\&\ \delta f_0 = f_\delta\ \&\ s \in \alpha\ \&$
$\mathrm{Dom}\ g_0 = \mathrm{Dom}\ f_0\ \&\ \forall r(r \in \mathrm{Dom}\ f_0 \longrightarrow g_0(r) = s \hat{/} f_0(r))\ \&$
$\delta g_0 = g_\delta)$, otherwise $g_\delta = \delta 0$.

6.40) $\forall\beta(\beta \in \text{Dom } f_{\mathfrak{f}} \longrightarrow f_{\mathfrak{f}}(\beta) \neq \tilde{0}) \longrightarrow \text{Dom } (\alpha/f_{\mathfrak{f}}) = \text{Dom } f_{\mathfrak{f}}$ &
$\forall\beta(\beta \in \text{Dom } f_{\mathfrak{f}} \longrightarrow (\alpha/f_{\mathfrak{f}})(\beta) = \alpha/(f_{\mathfrak{f}}(\beta)))$. ||

6.41) Def $f_{\mathfrak{f}}/g_{\mathfrak{f}} =_{\mathfrak{f}} f_{\mathfrak{f}}\cdot(\tilde{1}/g_{\mathfrak{f}})$.

6.42) $\text{Dom } f_{\mathfrak{f}} = \text{Dom } g_{\mathfrak{f}}$ & $\forall\alpha(\alpha \in \text{Dom } g_{\mathfrak{f}} \longrightarrow g_{\mathfrak{f}}(\alpha) \neq \tilde{0}) \longrightarrow$
$\text{Dom } (f_{\mathfrak{f}}/g_{\mathfrak{f}}) = \text{Dom } f_{\mathfrak{f}}$ & $\forall\alpha(\alpha \in \text{Dom } f_{\mathfrak{f}} \longrightarrow (f_{\mathfrak{f}}/g_{\mathfrak{f}})(\alpha) =$
$f_{\mathfrak{f}}(\alpha)/g_{\mathfrak{f}}(\alpha))$. ||

Two more useful operations are composition of functions and restriction of a function to a smaller domain.

6.43) Def $f_{\mathfrak{f}}\circ g_{\mathfrak{f}} =_{\mathfrak{f}} h_{\mathfrak{f}} \longleftrightarrow \forall\alpha(\alpha \in \text{Dom } g_{\mathfrak{f}} \longrightarrow g_{\mathfrak{f}}(\alpha) \in \text{Dom } f_{\mathfrak{f}})$ &
$\exists f_0 \exists g_0 \exists h_0(f_0$, g_0 , and h_0 are real functions & $\mathfrak{f}f_0 = f_{\mathfrak{f}}$ &
$\mathfrak{f}g_0 = g_{\mathfrak{f}}$ & $\text{Dom } h_0 = \text{Dom } g_0$ & $\forall r(r \in \text{Dom } h_0 \longrightarrow$
$h_0(r) = (\text{Extend}(f_0,\text{Ran } g_0))(g_0(r)))$ & $\mathfrak{f}h_0 = h_{\mathfrak{f}})$,
otherwise $h_{\mathfrak{f}} = \mathfrak{f}0$.

6.44) $\forall\alpha(\alpha \in \text{Dom } g_{\mathfrak{f}} \longrightarrow g_{\mathfrak{f}}(\alpha) \in \text{Dom } f_{\mathfrak{f}}) \longrightarrow \text{Dom } f_{\mathfrak{f}}\circ g_{\mathfrak{f}} = \text{Dom } g_{\mathfrak{f}}$ &
$\forall\alpha(\alpha \in \text{Dom } g_{\mathfrak{f}} \longrightarrow (f_{\mathfrak{f}}\circ g_{\mathfrak{f}})(\alpha) = f_{\mathfrak{f}}(g_{\mathfrak{f}}(\alpha)))$. ||

6.45) Def $f_{\mathfrak{f}}\lceil x_{\mathfrak{s}} =_{\mathfrak{f}} g_{\mathfrak{f}} \longleftrightarrow x_{\mathfrak{s}} \subseteq \text{Dom } f_{\mathfrak{f}}$ & $\exists f_0 \exists g_0 \exists a(f_0$ and g_0
are real functions & $\mathfrak{f}f_0 = f_{\mathfrak{f}}$ & a is a set of fractions &
$\mathfrak{s}a = x_{\mathfrak{s}}$ & $\text{Dom } g_0 = a$ & $\forall r(r \in a \longrightarrow g_0(r) = (\text{Extend}(f_0,a))(r))$ &
$\mathfrak{f}g_0 = g_{\mathfrak{f}})$, otherwise $g_{\mathfrak{f}} = \mathfrak{f}0$.

6.46) $x_{\mathfrak{s}} \subseteq \text{Dom } f_{\mathfrak{f}} \longrightarrow \text{Dom } (f_{\mathfrak{f}}\lceil x_{\mathfrak{s}}) = x_{\mathfrak{s}}$ & $\forall\alpha(\alpha \in x_{\mathfrak{s}} \longrightarrow (f_{\mathfrak{f}}\lceil x_{\mathfrak{s}})(\alpha) =$
$f_{\mathfrak{f}}(\alpha))$. ||

Sequences

The next of our refinements involves another new sort, $\mathcal{S}r$ ("sequences of real numbers"). It comes equipped with three function symbols: a symbol $\mathcal{S}r$ of type $(n;\mathcal{S}r)$, a symbol Ln of type $(\mathcal{S}r;n)$, and a symbol $\cdot(\cdot)$ of type $(\mathcal{S}r,n;r)$. The following axioms should by now be self-explanatory.

6.47) Ax u is a sequence of limited fractions $\longrightarrow$ Ln $(\mathcal{S}r\ u) =$ Ln u & $\forall i\ (1 \leq i \leq \text{Ln u} \longrightarrow u(i) \in (\mathcal{S}r\ u)(i))$.

6.48) Ax $\exists v$ (v is a sequence of limited fractions & $\mathcal{S}r\ v =_{\mathcal{S}r} u_{\mathcal{S}r}$) .

6.49) Ax Ln $u_{\mathcal{S}r}$ = Ln $v_{\mathcal{S}r}$ & $\forall i\ (1 \leq i \leq \text{Ln } u_{\mathcal{S}r} \longrightarrow u_{\mathcal{S}r}(i) =_r v_{\mathcal{S}r}(i)) \longrightarrow u_{\mathcal{S}r} =_{\mathcal{S}r} v_{\mathcal{S}r}$.

Call the resulting theory $R_3^{\mu\nu}$.

One advantage to the theory R_3 is that in it we can study polynomials with real coefficients in a more general setting than in §5 . The theorem (5.30), in light of (6.47) and (5.37), is just the uniqueness condition for the following definition.

6.50) Def Polyvalue $(u_{\mathcal{S}r},\alpha) =_r \beta \longleftrightarrow \varepsilon_{\nu+1}$ (Ln $u_{\mathcal{S}r}$) & $\exists v \exists r$ (v is a sequence of limited fractions & $\mathcal{S}r\ v = u_{\mathcal{S}r}$ & $r \in \alpha$ & Polyvalue $(v,r) \in \beta$) , otherwise $\beta = \tilde{0}$.

It is natural to regard polynomials as functions, of course, and the next proposition allows us to do just that.

6.51) $\varepsilon_{\nu+1}$ (Ln $u_{\mathcal{S}r}$) $\longrightarrow \exists f_{f}$(Dom $f_{f} = x_{s}$ & $\forall\alpha(\alpha \in x_{s} \longrightarrow f_{f}(\alpha) =$ Polyvalue $(u_{\mathcal{S}r},\alpha)))$.

Proof. Let a be a set of fractions such that $\mathfrak{s}a = x_{\mathfrak{s}}$, and let u_0 be a sequence of limited fractions such that $\mathcal{S}r\ u_0 = u_{\mathcal{S}r}$. The set $\{\langle r, \text{Polyvalue}\ (u_0, r)\rangle : r \in a\}$ is a real function, say f_0. Let $f_{\mathfrak{f}}$ be $\mathfrak{f} f_0$. ||

It is a simple matter to define elementary syntactic operations on sequences.

6.52) Def $u_{\mathcal{S}r}[i,j] =_{\mathcal{S}r} v_{\mathcal{S}r} \longleftrightarrow 1 \leq i \leq j \leq \text{Ln}\ u_{\mathcal{S}r}$ & $\exists u_0 \exists v_0$ (u_0 is a sequence of limited fractions & $\mathcal{S}r\ u_0 = u_{\mathcal{S}r}$ & $v_0 = u_0[i,j]$ & $\mathcal{S}r\ v_0 = v_{\mathcal{S}r}$), otherwise $v_{\mathcal{S}r} = \mathcal{S}r\ 0$.

6.53) $1 \leq i \leq j \leq \text{Ln}\ u_{\mathcal{S}r} \longrightarrow \text{Ln}\ u_{\mathcal{S}r}[i,j] = j-i+1$ & $\forall k\ (1 \leq k \leq j-i+1 \longrightarrow (u_{\mathcal{S}r}[i,j])(k) = u_{\mathcal{S}r}(k+i-1))$. ||

6.54) Def $u_{\mathcal{S}r} * v_{\mathcal{S}r} =_{\mathcal{S}r} w_{\mathcal{S}r} \longleftrightarrow \exists u_0 \exists v_0 \exists w_0$ (u_0 and v_0 are sequences of limited fractions & $\mathcal{S}r\ u_0 = u_{\mathcal{S}r}$ & $\mathcal{S}r\ v_0 = v_{\mathcal{S}r}$ & $w_0 = u_0 * v_0$ & $\mathcal{S}r\ w_0 = w_{\mathcal{S}r}$).

6.55) $\text{Ln}\ (u_{\mathcal{S}r} * v_{\mathcal{S}r}) = \text{Ln}\ u_{\mathcal{S}r} + \text{Ln}\ v_{\mathcal{S}r}$ & $\forall i\ (1 \leq i \leq \text{Ln}\ u_{\mathcal{S}r} \longrightarrow (u_{\mathcal{S}r} * v_{\mathcal{S}r})(i) = u_{\mathcal{S}r}(i))$ & $\forall i\ (\text{Ln}\ u_{\mathcal{S}r} < i \leq \text{Ln}\ u_{\mathcal{S}r} + \text{Ln}\ v_{\mathcal{S}r} \longrightarrow (u_{\mathcal{S}r} * v_{\mathcal{S}r})(i) = v_{\mathcal{S}r}(i - \text{Ln}\ u_{\mathcal{S}r}))$. ||

It is also possible to define the termwise sum and product of two sequences $u_{\mathcal{S}r}$ and $v_{\mathcal{S}r}$, the sequence of partial sums of $u_{\mathcal{S}r}$ as long as $\varepsilon_{\nu}(\text{Ln}\ u_{\mathcal{S}r})$, the sequence of partial products of $u_{\mathcal{S}r}$ as long as $\varepsilon_{\nu+1}(\text{Ln}\ u_{\mathcal{S}r})$, and a sequence $w_{\mathcal{S}r}$ of powers of α as long as $\varepsilon_{\nu+1}(\mathrm{Ln}\ w_{Sr})$. The restrictions are necessary: if $\neg\varepsilon_{\nu}(n)$, then the sum of n infinitesimals need not be infinitesimal, and if $\neg\varepsilon_{\nu+1}(m)$, then $\hat{2}$ raised to the m^{th} power is unlimited and so does not represent a real number. These operations provide an alternative way to evaluate a polynomial, which can be shown to be equivalent to (6.50). The details are straightforward.

More sequences

The last two sorts to be introduced in this section are Ss ("sequences of sets") and Sf ("sequences of functions"). The reader could probably provide the mechanical details himself, but for the record, $R_4^{\mu\nu}$ is the theory obtained by adjoining these two sorts to R_3 together with function symbols Ss of type $(n;Ss)$, Sf of type $(n;Sf)$, Ln of type $(Ss;n)$, Ln of type $(Sf;n)$, $\cdot(\cdot)$ of type $(Ss,n;s)$, and $\cdot(\cdot)$ of type $(Sf,n;f)$, and axioms (6.56)-(6.61).

6.56) Ax u is a sequence of sets of fractions $\longrightarrow$ $\mathrm{Ln}\ (Ss\ u) = \mathrm{Ln}\ u\ \&\ \forall i\ (1 \leq i \leq \mathrm{Ln}\ u \longrightarrow (Ss\ u)(i) =_s s(u(i)))$.

6.57) Ax $\exists v$ (v is a sequence of sets of fractions & $Ss\ v =_{Ss} u_{Ss}$).

6.58) Ax $\mathrm{Ln}\ u_{Ss} = \mathrm{Ln}\ v_{Ss}\ \&\ \forall i\ (1 \leq i \leq \mathrm{Ln}\ u_{Ss} \longrightarrow u_{Ss}(i) =_s v_{Ss}(i)) \longrightarrow u_{Ss} =_{Ss} v_{Ss}$.

6.59) Ax u is a sequence of real functions $\longrightarrow$ $\mathrm{Ln}\ (Sf\ u) = \mathrm{Ln}\ u\ \&$ $\forall i\ (1 \leq i \leq \mathrm{Ln}\ u \longrightarrow (Sf\ u)(i) =_f f(u(i)))$.

6.60) Ax $\exists v$ (v is a sequence of real functions & $Sf\ v =_{Sf} u_{Sf}$).

6.61) $\mathrm{Ax\ Ln}\ u_{S\mathit{f}} = \mathrm{Ln}\ v_{S\mathit{f}}\ \&\ \forall i\ (1 \leq i \leq \mathrm{Ln}\ u_{S\mathit{f}} \longrightarrow u_{S\mathit{f}}(i) =_{\mathit{f}} v_{S\mathit{f}}(i)) \longrightarrow u_{S\mathit{f}} =_{S\mathit{f}} v_{S\mathit{f}}$.

Clearly we can segment and juxtapose sequences of these two sorts. We can form the termwise sum, product, and composition of two sequences of functions, the sequence of partial sums of a sufficiently short (ε_ν) sequence of functions, the sequence of values of a given sequence of functions at a given real number, and, for that matter, the sequence of values of a given function at a given sequence of real numbers -- all under appropriate conditions on the domains of the functions. Given two sequences of sets, we can form the sequence which is their termwise union; given *one* such sequence, we can form the sequence of "partial unions" (whose last term is the union of all the sets in the original sequence). All of these assertions are easy to prove; we illustrate by showing that every sequence of functions has a sequence of domains.

6.62) $\exists v_{S\mathit{s}}(\mathrm{Ln}\ v_{S\mathit{s}} = \mathrm{Ln}\ u_{S\mathit{f}}\ \&\ \forall i\ (1 \leq i \leq \mathrm{Ln}\ u_{S\mathit{f}} \longrightarrow v_{S\mathit{s}}(i) = \mathrm{Dom}\ u_{S\mathit{f}}(i)))$.

Proof. Take a sequence u_0 of real functions such that $S\mathit{f}\ u_0 = u_{S\mathit{f}}$, let v_0 be the sequence of domains of the functions in the sequence u_0 (this is just bounded replacement), and let $v_{S\mathit{s}}$ be $S\mathit{s}\ v_0$. Then $\mathrm{Ln}\ v_{S\mathit{s}} = \mathrm{Ln}\ v_0 = \mathrm{Ln}\ u_0 = \mathrm{Ln}\ u_{S\mathit{f}}$, and if $1 \leq i \leq \mathrm{Ln}\ u_{S\mathit{f}}$, then $v_{S\mathit{s}}(i) = \mathit{s}(v_0(i)) = \mathit{s}(\mathrm{Dom}\ u_0(i)) = \mathrm{Dom}\ \mathit{f}(u_0(i)) = \mathrm{Dom}\ (S\mathit{f}\ u_0)(i) = \mathrm{Dom}\ u_{S\mathit{f}}(i)$. ||

The interpretation of R_4 *in* $\tilde{Q}^\mu$.

In §5 we pointed out that there is an interpretation I_0 of the two-sorted theory R_0 in (an extension by definitions of) the one-sorted theory $\tilde{Q}^\mu$. We show now that this interpretation can be extended to successive interpretations $I_1,\ldots,I_4$ of the theories $R_1,\ldots,R_4$ in $\tilde{Q}^\mu$. Actualy, nothing really new is involved. What we should do, of course, is define U_σ , $(=_\sigma)_I$, and $\underline{u}_I$ for every sort σ and nonlogical symbol $\underline{u}$ of R_4 , verify conditions (3.1)-(3.5) for an interpretation, and show that the interpretation of each of the new nonlogical axioms of this section is a theorem of $\tilde{Q}^\mu$. What we *shall* do is make the appropriate definitions and leave all but a few of the obvious verifications to the reader.

First recall how the interpretation I_0 works. We defined $U_n x \longleftrightarrow x = x$; $(=_n)_{I_0}$ is $=$; $\underline{u}_{I_0}$ is $\underline{u}$ for every nonlogical symbol $\underline{u}$ of $\tilde{Q}^\mu$; $U_r x \longleftrightarrow x$ is a limited fraction; $(=_r)_{I_0}$ and ϵ_{I_0} are both $\sim$. Now extend I_0 to an interpretation I_1 of R_1 in $\tilde{Q}^\mu$ by defining

6.63) Def $U_s a \longleftrightarrow a$ is a set of fractions,

6.64) Def $a(=_s)_{I_1} b \longleftrightarrow \forall r(r \epsilon a \;\&\; r$ is limited $\longrightarrow$

$\exists s(s \sim r \;\&\; s \epsilon b)) \;\&\; \forall s(s \epsilon b \;\&\; s$ is limited $\longrightarrow$

$\exists r(r \sim s \;\&\; r \epsilon a))$.

6.65) Def $s_{I_1} a = b \longleftrightarrow U_s a \;\&\; b = a$, otherwise $b = 0$.

6.66) Def $x(\epsilon_s)_{I_1} a \longleftrightarrow \exists r\ (r \sim x \ \&\ r \in a)$.

(In discussing interpretations, we must use care in distinguishing the various symbols ϵ ; here we write ϵ_s for the symbol of type (r, s) . Similar tricks will be employed with the symbols Ln and $\cdot(\cdot)$.)

Let us verify (3.1)-(3.5) this time. First (3.1) is $\exists a U_s a$, which is trivial. The form of (3.2) involving the function symbol s is $a = a \longrightarrow U_s\ (s_{I_1} a)$, which is clear from the definitions. Certainly $U_s a \longrightarrow a\ (=_s)_{I_1} a$, which is (3.3), is a theorem of $\tilde{Q}^\mu$, and (3.4) for the function symbol s is $x = x\ \&\ y = y\ \&\ x = y \longrightarrow s_{I_1}(x)\ (=_s)_{I_1} s_{I_1} y$, which is obvious. There are two forms of (3.5) to check: for $=_s$, (3.5) is $U_s x_1 \& U_s y_1 \& U_s x_2 \& U_s y_2 \& x_1 (=_s)_{I_1} y_1 \& x_2 (=_s)_{I_1} y_2 \longrightarrow (x_1 (=_s)_{I_1} x_2 \longrightarrow y_1 (=_s)_{I_1} y_2)$, and for ϵ_s , (3.5) is $U_r x\ \&\ U_r y\ \&\ U_s a\ \&\ U_s b\ \&\ x(=_r)_{I_1} y\ \&\ a(=_s)_{I_1} b \longrightarrow (x(\epsilon_s)_{I_1} a \longrightarrow y(\epsilon_s)_{I_1} b)$. Both of these are theorems of $\tilde{Q}^\mu$.

Let us also check the interpretations of axioms (6.1)-(6.3). First, $(6.1)^{I_1}$ is $a = a \longrightarrow$ (a is a set of fractions $\longrightarrow \forall\alpha(U_r \alpha \longrightarrow (\alpha(\epsilon_s)_{I_1} s_{I_1} a \longleftrightarrow \exists r(r = r\ \&\ r \in_{I_1} \alpha\ \&\ r \in a))))$ (this is a formula of $\tilde{Q}^\mu$, so of course all variables are of the same sort), which is clear when it is rewritten as a is a set of fractions $\longrightarrow$

$\forall\alpha(\alpha$ is a limited fraction $\longrightarrow (\alpha(\epsilon_{s})_{I_1} s_{I_1}$ a $\longleftrightarrow \exists r(r \sim \alpha$ & $r \in a)))$.

Nothing could be easier than $(6.2)^{I_1}$: $U_{s}x \longrightarrow \exists a(a = a$ & a is a set of fractions & $s_{I_1} a(=_{s})_{I_1} x)$ (just let a be x). Finally, $(6.3)^{I_1}$ is $U_{s}x$ & $U_{s}y \longrightarrow (\forall\alpha(U_{r}\alpha \longrightarrow (\alpha(\epsilon_{s})_{I_1} x \longleftrightarrow \alpha(\epsilon_{s})_{I_1} y)) \longrightarrow x(=_{s})_{I_1} y)$, which is exactly the way we defined $(=_{s})_{I_1}$.

We shall not be so meticulous in discussing the remaining sorts. All at once, here are the definitions for the interpretation I (really I_4) of R_4 in $\tilde{Q}^{\mu}$.

6.67) Def $U_{f}f \longleftrightarrow f$ is a real function.

6.68) Def $f(=_{f})_I g \longleftrightarrow \mathrm{Dom}\ f(=_{s})_I \mathrm{Dom}\ g$ & $\forall r\forall s(r \in \mathrm{Dom}\ f$ & $s \in \mathrm{Dom}\ g$ & r is limited & $r \sim s \longrightarrow f(r) \sim g(s))$.

6.69) Def $f_I f = g \longleftrightarrow U_{f}f$ & $g = f$, otherwise $g = 0$.

6.70) Def f maps_I x to y $\longleftrightarrow \exists r(r \sim x$ & $r \in \mathrm{Dom}\ f$ & $f(r) \sim y)$.

6.71) Def $U_{Sr}u \longleftrightarrow u$ is a sequence of limited fractions.

6.72) Def $u(=_{Sr})_I v \longleftrightarrow \mathrm{Ln}\ u = \mathrm{Ln}\ v$ & $\forall i\ (1 \leq i \leq \mathrm{Ln}\ u \longrightarrow u(i) \sim v(i))$.

6.73) Def $Sr_I u = v \longleftrightarrow U_{Sr}u$ & $v = u$, otherwise $v = 0$.

6.74) Def $(\mathrm{Ln}_{Sr})_I u = \mathrm{Ln}\ u$.

6.75) Def $(u(i)_{Sr})_I = r \longleftrightarrow (1 \leq i \leq \mathrm{Ln}\, u \;\&\; r = u(i))$, otherwise $r = \hat{0}$.

6.76) Def $U_{Ss} u \longleftrightarrow u$ is a sequence of sets of fractions .

6.77) Def $u(=_{Ss})_I v \longleftrightarrow \mathrm{Ln}\, u = \mathrm{Ln}\, v \;\&\; \forall i\, (1 \leq i \leq \mathrm{Ln}\, u \longrightarrow u(i)\ (=_s)_I v(i))$.

6.78) Def $Ss_I u = v \longleftrightarrow U_{Ss} u \;\&\; v = u$, otherwise $v = 0$.

6.79) Def $(\mathrm{Ln}_{Ss})_I u = \mathrm{Ln}\, u$.

6.80) Def $(u(i)_{Ss})_I = a \longleftrightarrow (1 \leq i \leq \mathrm{Ln}\, u \;\&\; a = u(i))$, otherwise $a = 0$.

6.81) Def $U_{Sf} u \longleftrightarrow u$ is a sequence of real functions.

6.82) Def $u(=_{Sf})_I v \longleftrightarrow \mathrm{Ln}\, u = \mathrm{Ln}\, v \;\&\; \forall i\, (1 \leq i \leq \mathrm{Ln}\, u \longrightarrow u(i)(=_f)_I v(i))$.

6.83) Def $Sf_I u = v \longleftrightarrow U_{Sf} u \;\&\; v = u$, otherwise $v = 0$.

6.84) Def $(\mathrm{Ln}_{Sf})_I u = \mathrm{Ln}\, u$.

6.85) Def $(u(i)_{Sf})_I = f \longleftrightarrow (1 \leq i \leq \mathrm{Ln}\, u \;\&\; f = u(i))$, otherwise $f = 0$.

Conditions (3.1)-(3.5) are no harder to verify than before. Let us spot-check the interpretations of a few random axioms, beginning with $(6.15)^I$, which is probably the most complex of all. This is

$f = f \longrightarrow$ (f is a real function $\longrightarrow$

$\forall\alpha(U_r \alpha \longrightarrow (\exists\beta(U_r \beta \;\&\; f_I f$ maps$_I$ α to β) $\longleftrightarrow$

$\exists r(r = r \ \& \ r \ \epsilon_I \ \alpha \ \& \ r \in \text{Dom } f))) \ \& \ \forall\alpha(U_r\alpha \longrightarrow \forall\beta(U_r\beta \longrightarrow$ $\forall r(r = r \longrightarrow (r \ \epsilon_I \ \alpha \ \& \ r \in \text{Dom } f \ \& \ \delta_I f \text{ maps}_I \ \alpha \text{ to } \beta \longrightarrow$ $f(r) \ \epsilon_I \ \beta)))))$. If we remember that $U_r\alpha \longleftrightarrow \alpha$ is a limited fraction, that $\delta_I f = f$ if f is a real function, and that ϵ_I is just $\sim$, this formula becomes little more than a restatement of the definition (6.70).

For $(6.48)^I$ we have

$U_{Sr}u \longrightarrow \exists v(v = v \ \& \ v$ is a sequence of limited fractions & $Sr_I v(=_{Sr})_I u)$, which is clear from (6.71)-(6.73). (Let v be u.)

Finally, $(6.58)^I$ is

$U_{Ss}u \ \& \ U_{Ss}v \longrightarrow ((Ln_{Ss})_I u = (Ln_{Ss})_I v \ \& \ \forall i \ (i = i \longrightarrow$ $(1 \le i \le (Ln_{Ss})_I u \longrightarrow (u(i)_{Ss})_I(=_s)_I(v(i)_{Ss})_I)) \longrightarrow u(=_{Ss})_I v)$,

which follows directly from (6.77).

A remark on induction

Conspicuously missing from the theories discussed in this section is the capability to use induction on any formulas that are not entirely of sort n . This deficiency appears unavoidable: all sorts other than n are intended to represent highly unbounded concepts, and in fact some very innocuous-looking forms of induction lead immediately to contradictions. For instance, if u_{Sr} is a sequence of real numbers, there need not exist a smallest i such that $u_{Sr}(i) = \tilde{0}$. It should be clear already, however, that all is not lost, or at least not all is lost; (6.62) is a good case in point. If we want to prove by induction a formula involving sorts other than n , we first use (5.37)-(5.39) and the axioms of this section to translate the hypotheses into a statement entirely of sort n ; if we are lucky, this formula will be (or can be weakened to be) sufficiently bounded that some induction scheme is applicable and yields a statement that, when translated back into the more convenient notation of the original sorts, implies the desired conclusion. Our study of calculus in §7 will provide many examples of this technique. A nice exercise at this point is to define the nth Fibonacci number and prove that if $\varepsilon_{\nu+1}(n)$, then it is equal to $(((1+\sqrt{5})/2)^n-((1-\sqrt{5})2)^n)/\sqrt{5}$.

§7. *A Survey of Calculus*

In this section we show that the theory R_4 is sufficient to reproduce the standard theorems of first-year calculus. The reader will recognize many techniques from nonstandard analysis.

The derivative

Our first objective is the definition of a derivative. Two preliminary definitions:

7.1) Def deriv $(f_{\mathfrak{f}}, \alpha, f_0, \xi) \longleftrightarrow \exists\gamma_1 \exists\gamma_2 (\gamma_1 < \gamma_2 \ \& \ \alpha \in [\gamma_1, \gamma_2] \subseteq \text{Dom } f_{\mathfrak{f}}) \ \&$

f_0 is a real function & $\mathfrak{f} f_0 = f_{\mathfrak{f}}$ &

$\forall r_1 \forall r_2 \, (r_1 \in \alpha \ \& \ r_2 \in \alpha \ \& \ r_1 \in \text{Dom } f_0 \ \& \ r_2 \in \text{Dom } f_0 \ \& \ r_1 \neq r_2 \longrightarrow$

$(f_0(r_1) \mathbin{\hat{-}} f_0(r_2)) \mathbin{\hat{/}} (r_1 \mathbin{\hat{-}} r_2) \in \xi)$.

7.2) Def $f_{\mathfrak{f}}$ is differentiable at $\alpha \longleftrightarrow \exists f_0 \exists \xi$ deriv $(f_{\mathfrak{f}}, \alpha, f_0, \xi)$.

A few remarks are in order. First, note that in (7.1) the requirement that α be contained in some interval that is a subset of Dom $f_{\mathfrak{f}}$ precludes the possibility that there might be only one r such that $r \in \alpha$ & $r \in$ Dom $f_{\mathfrak{f}}$. Also, observe that a function is allowed to be differentiable even at the endpoints of its domain. In our proofs in this section, we shall often treat only the case in which α is an interior point of Dom $f_{\mathfrak{f}}$, leaving any necessary modifications for the endpoints to the reader.

Now a uniqueness condition:

7.3) deriv $(f_{\mathfrak{f}}, \alpha, f_1, \xi_1)$ & deriv $(f_{\mathfrak{f}}, \alpha, f_2, \xi_2) \longrightarrow \xi_1 = \xi_2$.

Proof. Assume $[\alpha,\alpha+\varepsilon] \subseteq \mathrm{Dom}\, f_\delta$ for some $\varepsilon > \tilde{0}$. (If not, a similar argument applies to some set $[\alpha-\varepsilon,\alpha]$.) Suppose $\xi_1 < \xi_2$, and choose fractions s_1 and s_2 with $\xi_1 < \tilde{s}_1 < \tilde{s}_2 < \xi_2$. Let a_1 and a_2 be fractions such that $a_1 \in \alpha$, $a_1 \in \mathrm{Dom}\, f_1$, $a_2 \in \alpha$, and $a_2 \in \mathrm{Dom}\, f_2$. If $r \mathbin{\hat{>}} a_1$, $r \sim a_1$, and $r \in \mathrm{Dom}\, f_1$, then by the hypothesis $\mathrm{deriv}\,(f_\delta,\alpha,f_1,\xi_1)$ it follows that $f_1(r) = f_1(a_1) \mathbin{\hat{+}} (r \mathbin{\hat{-}} a_1) \mathbin{\hat{\cdot}} x$ for some $x \in \xi_1$; in particular; $f_1(r) \mathbin{\hat{<}} f_1(a_1) \mathbin{\hat{+}} (r \mathbin{\hat{-}} a_1) \mathbin{\hat{\cdot}} s_1$. Therefore the set $\{r \in \mathrm{Dom}\, f_1 \colon r \mathbin{\hat{>}} a_1 \;\&\; f_1(r) \mathbin{\hat{\geq}} f_1(a_1) \mathbin{\hat{+}} (r \mathbin{\hat{-}} a_1) \mathbin{\hat{\cdot}} s_1\}$, if nonempty, has a smallest element t_1, which is greater than and not infinitely close to a_1. Let τ_1 be $\tilde{t}_1$ if $t_1 \mathbin{\hat{<}} a_1 \mathbin{\hat{+}} \hat{1}$; if $t_1 \mathbin{\hat{\geq}} a_1 \mathbin{\hat{+}} \hat{1}$ or if the aforementioned set is empty, let τ_1 be $\alpha+\tilde{1}$. Likewise, if $r \mathbin{\hat{>}} a_2$, $r \sim a_2$, and $r \in \mathrm{Dom}\, f_2$, then $f_2(r) \mathbin{\hat{>}} f_2(a_2) \mathbin{\hat{+}} (r \mathbin{\hat{-}} a_2) \mathbin{\hat{\cdot}} s_2$; let τ_2 be either the real number represented by the smallest element of the set $\{r \in \mathrm{Dom}\, f_2 \colon r \mathbin{\hat{>}} a_2 \;\&\; f_2(r) \mathbin{\hat{\leq}} f_2(a_2) \mathbin{\hat{+}} (r \mathbin{\hat{-}} a_2) \mathbin{\hat{\cdot}} s_2\}$ or else $\alpha+\tilde{1}$ (if this set is empty or if its smallest element is $\mathbin{\hat{\geq}} a_2 \mathbin{\hat{+}} \hat{1}$).

Let γ be a real number greater than α and less than the smallest of τ_1, τ_2, and $\alpha+\varepsilon$. Some $r_1 \in \gamma$ is in $\mathrm{Dom}\, f_1$; since $\tilde{r}_1 = \gamma < \tau_1$ it follows that $f_1(r_1) \mathbin{\hat{<}} f_1(a_1) \mathbin{\hat{+}} (r_1 \mathbin{\hat{-}} a_1) \mathbin{\hat{\cdot}} s_1$. Likewise, some $r_2 \in \gamma$ is in $\mathrm{Dom}\, f_2$, and $f_2(r_2) \mathbin{\hat{>}} f_2(a_2) \mathbin{\hat{+}} (r_2 \mathbin{\hat{-}} a_2) \mathbin{\hat{\cdot}} s_2$. But this implies

$$\begin{aligned} f_\delta(\gamma) &= (f_1(r_1))^\sim \leq (f_1(a_1) \mathbin{\hat{+}} (r_1 \mathbin{\hat{-}} a_1) \mathbin{\hat{\cdot}} s_1)^\sim = f_\delta(\alpha) + (\gamma-\alpha)\cdot\tilde{s}_1 \\ &< f_\delta(\alpha) + (\gamma-\alpha)\cdot\tilde{s}_2 = (f_2(a_2) \mathbin{\hat{+}} (r_2 \mathbin{\hat{-}} a_2) \mathbin{\hat{\cdot}} s_2)^\sim \leq (f_2(r_2))^\sim = f_\delta(\gamma) , \end{aligned}$$

which is impossible. We conclude that $\xi_1 \not< \xi_2$, and similarly $\xi_2 \not< \xi_1$. Thus $\xi_1 = \xi_2$. ‖

7.4) Def Deriv $(f_{\mathfrak{f}},\alpha) =_{r} \xi \longleftrightarrow \exists f_0$ deriv $(f_{\mathfrak{f}},\alpha,f_0,\xi)$, otherwise $\xi = \tilde{0}$.

It is, of course, convenient to regard the derivative of a function as a function itself.

7.5) Def $f'_{\mathfrak{f}} =_{\mathfrak{f}} g_{\mathfrak{f}} \longleftrightarrow \forall\alpha(\alpha \in \text{Dom } f_{\mathfrak{f}} \longrightarrow f_{\mathfrak{f}}$ is differentiable at $\alpha)$ & Dom $g_{\mathfrak{f}}$ = Dom $f_{\mathfrak{f}}$ & $\forall\alpha(\alpha \in \text{Dom } g_{\mathfrak{f}} \longrightarrow g_{\mathfrak{f}}(\alpha) = \text{Deriv } (f_{\mathfrak{f}},\alpha))$, otherwise $g_{\mathfrak{f}} = \mathfrak{f}0$.

The uniqueness condition for (7.5) follows from (6.17). One defect of the notation $f'_{\mathfrak{f}}$ is that $f'_{\mathfrak{f}}$ will not exist (or, more properly, will be the empty function) if there is even one point in Dom $f_{\mathfrak{f}}$ at which $f_{\mathfrak{f}}$ is not differentiable. If $f_{\mathfrak{f}}$ is the absolute value function, for instance, then $f_{\mathfrak{f}}$ is differentiable at every real number except $\tilde{0}$, but its derivative does not define a function at all, since every function must have as its domain a (closed) set.

Basic properties of derivatives

If any functions at all are differentiable, polynomials had better be. First let us define the "derived polynomial" of a sequence of real numbers. Recall that the i^{th} term of a sequence corresponds to the term of degree i-1 in the **polynomial**.

7.6) Def Derivpoly $u_{Sr} =_{Sr} v_{Sr} \longleftrightarrow \text{Ln } v_{Sr} = \text{Ln } u_{Sr}-1$ & $\forall i \ (1 \leq i \leq \text{Ln } v_{Sr} \longrightarrow v_{Sr}(i) = \tilde{i}\cdot u_{Sr}(i+1))$.

7.7) $\forall\alpha(\alpha \in \mathrm{Dom}\ f_{\mathfrak{f}} \longrightarrow \exists\gamma_1\exists\gamma_2\ (\alpha \in [\gamma_1,\gamma_2] \subseteq \mathrm{Dom}\ f_{\mathfrak{f}})\ \&$
$f_{\mathfrak{f}}(\alpha) = \mathrm{Polyvalue}\ (u_{\mathcal{Sr}},\alpha)) \longrightarrow \forall\alpha(\alpha \in \mathrm{Dom}\ f_{\mathfrak{f}} \longrightarrow$
$f_{\mathfrak{f}}$ is differentiable at α & $f'_{\mathfrak{f}}(\alpha) =$
$\mathrm{Polyvalue}(\mathrm{Derivpoly}\ u_{\mathcal{Sr}},\alpha))$.

Proof. Assume $\varepsilon_{\nu+1}(\mathrm{Ln}\ u_{\mathcal{Sr}})$. (This is the interesting case. Otherwise, by the definition (6.50), both Polyvalue $(u_{\mathcal{Sr}},\alpha)$ and Polyvalue(Derivpoly $u_{\mathcal{Sr}},\alpha)$ are $\tilde{0}$ for every α .) Let Dom $f_{\mathfrak{f}} = \mathfrak{s}\,x_0$, let $u_{\mathcal{Sr}} = \mathcal{Sr}\ u_0$, and let f_0 be the real function defined on the set x_0 by $f_0(r) = \mathrm{Polyvalue}\ (u_0,r)$. For each α in Dom $f_{\mathfrak{f}}$, we are to show that deriv $(f_{\mathfrak{f}},\alpha,f_0,\mathrm{Polyvalue}(\mathrm{Derivpoly}\ u_{\mathcal{Sr}},\alpha))$. This entails nothing more than the obvious formalization in $\mathcal{R}_4$ of the following argument: if $r \in \alpha$ and $s \in \alpha$, then

$$\frac{f_0(r)-f_0(s)}{r-s} = \frac{1}{r-s}\Big(\sum_{i=1}^{\mathrm{Ln}\ u_0} u_0(i)\cdot r^{i-1} - \sum_{i=1}^{\mathrm{Ln}\ u_0} u_0(i)\cdot s^{i-1}\Big) = \sum_{i=2}^{\mathrm{Ln}\ u_0} u_0(i)\cdot \sum_{j=0}^{i-2} r^j s^{i-2-j}$$

$$= \sum_{i=1}^{\mathrm{Ln}\ u_0-1} u_0(i+1)\cdot \sum_{j=0}^{i-1} r^j s^{i-1-j} \sim \sum_{i=1}^{\mathrm{Ln}\ u_0-1} i\cdot u_0(i+1)\cdot r^{i-1}\ . \ \|$$

The next proposition asserts that differentiability is a "local" notion.

7.8) $\varepsilon > \tilde{0}\ \&\ [\alpha-\varepsilon,\alpha+\varepsilon] \subseteq \mathrm{Dom}\ f \longrightarrow$
$((f_{\mathfrak{f}}$ is differentiable at α & $\mathrm{Deriv}\ (f_{\mathfrak{f}},\alpha) = \xi) \longleftrightarrow$
$(f_{\mathfrak{f}}\restriction[\alpha-\varepsilon,\alpha+\varepsilon]$ is differentiable at α & $\mathrm{Deriv}\ (f_{\mathfrak{f}}\restriction[\alpha-\varepsilon,\alpha+\varepsilon],\alpha) = \xi))$.

Proof. Take $s_1 \in \alpha-\varepsilon$ and $s_2 \in \alpha+\varepsilon$. If deriv $(f_{\mathfrak{f}},\alpha,f_0,\xi)$, then the restriction g_0 of f_0 to the set $\{r \in \text{Dom } f_0 : s_1 \leq r \leq s_2\}$ satisfies deriv $(f_{\mathfrak{f}}\restriction[\alpha-\varepsilon,\alpha+\varepsilon],\alpha,g_0,\xi)$. Conversely, if deriv $(f_{\mathfrak{f}}\restriction[\alpha-\varepsilon,\alpha+\varepsilon],\alpha,g_1,\xi)$, then we may assume the smallest and largest elements of Dom g_1 are s_1 and s_2, respectively. Let f_1 be a real function with $\mathfrak{f}f_1 = f_{\mathfrak{f}}$, and let f_2 be the extension of g_1 to Dom $g_1 \cup \{r \in \text{Dom } f_1 : r \mathrel{\hat{<}} s_1 \vee r \mathrel{\hat{>}} s_2\}$ that agrees with f_1 on the latter set; then deriv $(f_{\mathfrak{f}},\alpha,f_2,\xi)$. ||

We next verify the usual sum and product rules for differentation.

7.9) $f_{\mathfrak{f}}$ is differentiable at $\alpha \longrightarrow -f_{\mathfrak{f}}$ is differentiable at α & Deriv $(-f_{\mathfrak{f}},\alpha) = -$Deriv $(f_{\mathfrak{f}},\alpha)$.

Proof. If deriv $(f_{\mathfrak{f}},\alpha,f_0,\xi)$ and h is the negative of f_0, then deriv $(-f_{\mathfrak{f}},\alpha,h,-\xi)$. ||

7.10) Dom $f_{\mathfrak{f}}$ = Dom $g_{\mathfrak{f}}$ & $f_{\mathfrak{f}}$ and $g_{\mathfrak{f}}$ are differentiable at α $\longrightarrow$ $f_{\mathfrak{f}} + g_{\mathfrak{f}}$ is differentiable at α & Deriv$(f_{\mathfrak{f}}+g_{\mathfrak{f}},\alpha)$ = Deriv $(f_{\mathfrak{f}},\alpha)$+Deriv $(g_{\mathfrak{f}},\alpha)$.

Proof. By (7.8) we may assume Dom $f_{\mathfrak{f}} = [\alpha-\varepsilon,\alpha+\varepsilon]$ for some $\varepsilon > \tilde{0}$. Let deriv $(f_{\mathfrak{f}},\alpha,f_0,\xi)$ and deriv $(g_{\mathfrak{f}},\alpha,g_0,\eta)$. Let f_1 be the extension of f_0 to Dom $f_0 \cup$ Dom g_0 defined by linear interpolation between successive values of f_0. That is, if $r \in$ Dom f_0, let $f_1(r)$ be $f_0(r)$; if $r \in$ Dom g_0 & $r \notin$ Dom f_0, s_1 is the greatest element of Dom f_0 smaller than r, and s_2 is the smallest element of Dom f_0 greater than r, then define

$f_1(r) = f_0(s_1) \hat{+} (f_0(s_2) \hat{-} f_0(s_1)) \hat{\cdot} (r \hat{-} s_1) \hat{/} (s_2 \hat{-} s_1)$. (There may be problems at the endpoints $\alpha-\varepsilon$ and $\alpha+\varepsilon$; if, say, r is smaller than every element of Dom f_0 , then define $f_1(r)$ by extrapolating from the first two values of f_0.) It is clear that f_1 is a real function and that $\mathfrak{f} f_1 = \mathfrak{f} f_0 = f_{\mathfrak{f}}$, and almost as clear that deriv $(f_{\mathfrak{f}}, \alpha, f_1, \xi)$. Let g_1 be the extension of g_0 to Dom $f_0 \,\hat{\cup}\,$ Dom g_0 obtained likewise, so that deriv $(g_{\mathfrak{f}}, \alpha, g_1, \eta)$. Let h be the pointwise sum of f_1 and g_1 on Dom $f_0 \cup$ Dom g_0 . Then $\mathfrak{f} h = f_{\mathfrak{f}} + g_{\mathfrak{f}}$, and if $r_1 \in \alpha$ & $r_2 \in \alpha$ & $r_1 \in$ Dom h & $r_2 \in$ Dom h & $r_1 \neq r_2$, then

$$\frac{h(r_1)-h(r_2)}{r_1-r_2} = \frac{f_1(r_1)-f_1(r_2)}{r_1-r_2} + \frac{g_1(r_1)-g_1(r_2)}{r_1-r_2} \in \xi+\eta \quad \text{, whence}$$

deriv $(f_{\mathfrak{f}}+g_{\mathfrak{f}}, \alpha, h, \xi+\eta)$, as desired. ||

7.11) Dom $f_{\mathfrak{f}}$ = Dom $g_{\mathfrak{f}}$ & $f_{\mathfrak{f}}$ and $g_{\mathfrak{f}}$ are differentiable at α $\longrightarrow$
$f_{\mathfrak{f}} \cdot g_{\mathfrak{f}}$ is differentiable at α &
Deriv $(f_{\mathfrak{f}} \cdot g_{\mathfrak{f}}, \alpha) = f_{\mathfrak{f}}(\alpha) \cdot$ Deriv $(g_{\mathfrak{f}}, \alpha) + g_{\mathfrak{f}}(\alpha) \cdot$ Deriv $(f_{\mathfrak{f}}, \alpha)$.

Proof. Proceed as in the proof of (7.10) through the construction of f_1 and g_1 . Let h be the pointwise product of f_1 and g_1 , so $\mathfrak{f} h = f_{\mathfrak{f}} \cdot g_{\mathfrak{f}}$. Suppose $r_1 \in \alpha$ & $r_2 \in \alpha$ & $r_1 \in$ Dom h & $r_2 \in$ Dom h & $r_1 \neq r_2$; let $d = r_1 \hat{-} r_2$, so $d \sim \hat{0}$ but $d \neq \hat{0}$. Then
$f_1(r_1) = f_1(r_2) \hat{+} s \hat{\cdot} d$ for some $s \in \xi =$ Deriv $(f_{\mathfrak{f}}, \alpha)$, and
$g_1(r_1) = g_1(r_2) \hat{+} t \hat{\cdot} d$ for some $t \in \eta =$ Deriv $(g_{\mathfrak{f}}, \alpha)$. It follows that
$h(r_1) = h(r_2) \hat{+} (f_1(r_2) \hat{\cdot} t \hat{+} g_1(r_2) \hat{\cdot} s) \hat{\cdot} d \hat{+} s \hat{\cdot} t \hat{\cdot} d \hat{\cdot} d$, and thus
$(h(r_1) \hat{-} h(r_2)) \hat{/} (r_1 \hat{-} r_2) \in f_{\mathfrak{f}}(\alpha) \cdot \eta + g_{\mathfrak{f}}(\alpha) \cdot \xi$. ||

7.12) $\forall\beta(\beta \in \text{Dom } f_{\mathfrak{f}} \longrightarrow f_{\mathfrak{f}}(\beta) \neq \tilde{0})$ & $f_{\mathfrak{f}}$ is differentiable at $\alpha \longrightarrow$ $\tilde{1}/f_{\mathfrak{f}}$ is differentiable at α & Deriv $(\tilde{1}/f_{\mathfrak{f}},\alpha) =$ $-\text{Deriv } (f_{\mathfrak{f}},\alpha)/(f_{\mathfrak{f}}(\alpha)\cdot f_{\mathfrak{f}}(\alpha))$.

Proof. Let deriv $(f_{\mathfrak{f}},\alpha,f_0,\xi)$, and let h be the pointwise reciprocal of f_0 . Then h is a real function and $\mathfrak{f}h = \tilde{1}/f_{\mathfrak{f}}$. If $r_1 \in \alpha$ & $r_2 \in \alpha$ & $r_1 \in \text{Dom } h$ & $r_2 \in \text{Dom } h$ & $r_1 \neq r_2$, then $(h(r_1)\hat{-}h(r_2))\hat{/}(r_1\hat{-}r_2) = (f_0(r_2)\hat{-}f_0(r_1))\hat{/}(f_0(r_1)\hat{\cdot}f_0(r_2)\hat{\cdot}(r_1\hat{-}r_2))$ $= \hat{-}s\hat{/}(f_0(r_1)\hat{\cdot}f_0(r_2))$ for some $s \in \xi$. Therefore deriv $(\tilde{1}/f_{\mathfrak{f}},\alpha,h,-\xi/(f_{\mathfrak{f}}(\alpha)\cdot f_{\mathfrak{f}}(\alpha)))$. ||

7.13) $\forall\beta(\beta \in \text{Dom } g_{\mathfrak{f}} \longrightarrow g_{\mathfrak{f}}(\beta) \neq \tilde{0})$ & Dom $f_{\mathfrak{f}}$ = Dom $g_{\mathfrak{f}}$ & $f_{\mathfrak{f}}$ and $g_{\mathfrak{f}}$ are differentiable at $\alpha \longrightarrow f_{\mathfrak{f}}/g_{\mathfrak{f}}$ is differentiable at α & Deriv $(f_{\mathfrak{f}}/g_{\mathfrak{f}},\alpha) = (g_{\mathfrak{f}}(\alpha)\cdot\text{Deriv } (f_{\mathfrak{f}},\alpha)-f_{\mathfrak{f}}(\alpha)\cdot\text{Deriv } (g_{\mathfrak{f}},\alpha))/(g_{\mathfrak{f}}(\alpha)\cdot g_{\mathfrak{f}}(\alpha))$.

Proof. By (7.11) and (7.12). ||

It does not take much effort to prove other simple theorems about derivatives, such as the chain rule and properties of local maxima and minima.

7.14) $\forall\beta(\beta \in \text{Dom } g_{\mathfrak{f}} \longrightarrow g_{\mathfrak{f}}(\beta) \in \text{Dom } f_{\mathfrak{f}})$ & $g_{\mathfrak{f}}$ is differentiable at α & $f_{\mathfrak{f}}$ is differentiable at $g_{\mathfrak{f}}(\alpha) \longrightarrow f_{\mathfrak{f}}\circ g_{\mathfrak{f}}$ is differentiable at α & Deriv $(f_{\mathfrak{f}}\circ g_{\mathfrak{f}},\alpha) = \text{Deriv } (f_{\mathfrak{f}},g_{\mathfrak{f}}(\alpha))\cdot\text{Deriv } (g_{\mathfrak{f}},\alpha)$.

Proof. Assume that Dom $g_{\mathfrak{f}}$ is some small interval containing α and that Dom $f_{\mathfrak{f}}$ is some small interval containing $g_{\mathfrak{f}}(\alpha)$. Let deriv $(g_{\mathfrak{f}},\alpha,g_0,\xi)$ and deriv $(f_{\mathfrak{f}},g_{\mathfrak{f}}(\alpha),f_0,\eta)$. As in the proof of (7.10), extend f_0 by linear interpolation to a real function f_1

defined on Dom $f_0 \cup$ Ran g_0 ; then deriv $(f_\delta, g_\delta(\alpha), f_1, \eta)$. If $r_1 \in \alpha$ & $r_2 \in \alpha$ & $r_1 \in$ Dom g_0 & $r_2 \in$ Dom g_0 & $r_1 \neq r_2$, then $g_0(r_1) = g_0(r_2) \hat{+} s \hat{\cdot} (r_1 \hat{-} r_2)$ for some $s \in \xi$; since $s \hat{\cdot} (r_1 \hat{-} r_2)$ is infinitesimal, it follows that $f_1(g_0(r_1)) = f_0(g_0(r_2)) \hat{+} t \hat{\cdot} s \hat{\cdot} (r_1 \hat{-} r_2)$ for some $t \in \eta$. Because $\delta(f_1 \circ g_0) = f_\delta \circ g_\delta$, the desired conclusion is then immediate. ||

7.15) $\varepsilon > \tilde{0}$ & $[\alpha, \alpha+\varepsilon] \subseteq$ Dom f_δ & f_δ is differentiable at α &
Deriv $(f_\delta, \alpha) > \tilde{0} \longrightarrow \exists\beta(\alpha < \beta < \alpha+\varepsilon$ & $f_\delta(\beta) > f_\delta(\alpha))$.

Proof. Let deriv $(f_\delta, \alpha, f_0, \xi)$, so $\xi > \tilde{0}$ by hypothesis. Take $a \in \alpha$ such that $a \in$ Dom f_0 , and take a fraction s with $\tilde{0} < \tilde{s} < \xi$. The set $\{r \in \text{Dom } f_0 : r \hat{>} a \text{ \& } f(r) \hat{>} f_0(a) \hat{+} s \hat{\cdot} (r \hat{-} a)\}$ contains all elements of Dom f_0 to the right of a and infinitely close to a ; it therefore also contains all sufficiently small elements of Dom f_0 to the right of a but *not* infinitely close to a . For such an r , $\alpha < \tilde{r} < \alpha+\varepsilon$ and $f_\delta(\tilde{r}) = (f_0(r))^\sim > (f_0(a))^\sim = f_\delta(\alpha)$. ||

7.16) $\varepsilon > \tilde{0}$ & $[\alpha-\varepsilon, \alpha] \subseteq$ Dom f_δ & f_δ is differentiable at α &
Deriv $(f_\delta, \alpha) < \tilde{0} \longrightarrow \exists\beta(\alpha-\varepsilon < \beta < \alpha$ & $f_\delta(\beta) > f_\delta(\alpha))$. ||

7.17) $\varepsilon > \tilde{0}$ & $[\alpha-\varepsilon, \alpha+\varepsilon] \subseteq$ Dom f_δ & f_δ is differentiable at α &
$\forall\beta(\alpha-\varepsilon < \beta < \alpha+\varepsilon \longrightarrow f_\delta(\beta) \leq f_\delta(\alpha)) \longrightarrow$ Deriv $(f_\delta, \alpha) = \tilde{0}$. ||

7.18) $\varepsilon > \tilde{0}$ & $[\alpha-\varepsilon, \alpha+\varepsilon] \subseteq$ Dom f_δ & f_δ is differentiable at α &
$\forall\beta(\alpha-\varepsilon < \beta < \alpha+\varepsilon \longrightarrow f_\delta(\beta) \geq f_\delta(\alpha)) \longrightarrow$ Deriv $(f_\delta, \alpha) = \tilde{0}$. ||

We now have all the tools for an easy proof of Rolle's theorem.

7.19) $\alpha < \beta$ & $[\alpha,\beta] \subseteq \text{Dom } f_{\mathfrak{f}}$ & $\forall\gamma(\alpha < \gamma < \beta \longrightarrow f_{\mathfrak{f}}$ is differentiable at $\gamma)$& $f_{\mathfrak{f}}(\alpha) = f_{\mathfrak{f}}(\beta) \longrightarrow \exists\gamma(\alpha < \gamma < \beta$ & Deriv $(f_{\mathfrak{f}},\gamma) = 0)$.

Proof. By (6.28) applied to the functions $f_{\mathfrak{f}}\upharpoonright[\alpha,\beta]$ and $-f_{\mathfrak{f}}\upharpoonright[\alpha,\beta]$, $f_{\mathfrak{f}}$ attains a maximum value and a minimum value on $[\alpha,\beta]$. If both of these occur at the endpoints α and β , then $f_{\mathfrak{f}}$ is constant on $[\alpha,\beta]$, so Deriv $(f_{\mathfrak{f}},\gamma) = \tilde{0}$ for *every* γ between α and β . Otherwise, either the maximum or the minimum of $f_{\mathfrak{f}}$ occurs at some γ with $\alpha < \gamma < \beta$, and either (7.17) or (7.18) ensures that Deriv $(f_{\mathfrak{f}},\gamma) = \tilde{0}$. ||

The usual adjustment by a linear function "to make ends meet" converts (7.19) into the mean value theorem.

7.20) $\alpha < \beta$&$[\alpha,\beta] \subseteq \text{Dom } f_{\mathfrak{f}}$&$\forall\gamma(\alpha < \gamma < \beta \longrightarrow f_{\mathfrak{f}}$ is differentiable at $\gamma) \longrightarrow$ $\exists\gamma(\alpha < \gamma < \beta$& Deriv $(f_{\mathfrak{f}},\gamma) = (f_{\mathfrak{f}}(\beta)-f_{\mathfrak{f}}(\alpha))/(\beta-\alpha))$.||

The following corollaries are immediate.

7.21) $\alpha < \beta$&$[\alpha,\beta] \subseteq \text{Dom } f_{\mathfrak{f}}$ & $\forall\gamma(\alpha < \gamma < \beta \longrightarrow f_{\mathfrak{f}}$ is differentiable at γ & Deriv $(f_{\mathfrak{f}},\gamma) \geq \tilde{0}) \longrightarrow f_{\mathfrak{f}}(\alpha) \leq f_{\mathfrak{f}}(\beta)$. ||

7.22) $\alpha < \beta$&$[\alpha,\beta] \subseteq \text{Dom } f_{\mathfrak{f}}$ & $\forall\gamma(\alpha < \gamma < \beta \longrightarrow f_{\mathfrak{f}}$ is differentiable at γ & Deriv $(f_{\mathfrak{f}},\gamma) > \tilde{0}) \longrightarrow f_{\mathfrak{f}}(\alpha) < f_{\mathfrak{f}}(\beta)$. ||

In connection with (7.22) it is natural to discuss inverse functions. We should not expect to attain the level of generality one might hope for: indeed, the function $\alpha \longmapsto \tilde{1}/\alpha$ on $[\tilde{1},\infty)$ has no inverse (its domain would not be a closed set). It seems necessary, therefore, to limit the discussion to functions defined on a bounded interval.

7.23) Def f_δ is one-to-one $\longleftrightarrow \forall\alpha\forall\beta(\alpha \in \text{Dom } f_\delta \ \& \ \beta \in \text{Dom } f_\delta \ \&$
$f_\delta(\alpha) = f_\delta(\beta) \longrightarrow \alpha = \beta)$.

7.24) $\text{Dom } f_\delta = [\gamma_1,\gamma_2] \ \& \ f_\delta$ is one-to-one $\longrightarrow$
$\exists g_\delta \forall\alpha\forall\beta(f_\delta$ maps α to $\beta \longleftrightarrow g_\delta$ maps β to $\alpha)$.

Proof. Let $f_\delta = \delta f_0$, and assume that $\text{Dom } f_0$ contains no unlimited elements. Nothing guarantees that f_0 is one-to-one; it may well happen that $f_0(r) = f_0(s)$ for some r and s with $r \sim s$ but $r \neq s$. This difficulty can be circumvented, however, by removing from f_0 all ordered pairs $\langle r,f_0(r)\rangle$ such that $f_0(r) = f_0(s)$ for some $s \mathrel{\hat{<}} r$. The resulting set f_1 (which exists by bounded separation) is a real function and is one-to-one; using injectivity of f_δ and boundedness of $\text{Dom } f_0$, it is easy to see that $\delta f_1 = \delta f_0 = f_\delta$. Let g_1 be the inverse of f_1 -- that is, the set $\{\langle s,r\rangle : \langle r,s\rangle \in f_1\}$. Again by injectivity of f_δ and boundedness of $\text{Dom } f_1$, g_1 is a real function. For all α and β we have

$$f_\delta \text{ maps } \alpha \text{ to } \beta \longleftrightarrow \exists r \exists s(r \in \alpha \ \& \ s \in \beta \ \& \ f_1(r) = s)$$

$$\longleftrightarrow \exists r \exists s(r \in \alpha \ \& \ s \in \beta \ \& \ g_1(s) = r)$$

$$\longleftrightarrow \delta g_1 \text{ maps } \beta \text{ to } \alpha \ . \quad \|$$

7.25) Def $f_\delta^{-1} = g_\delta \longleftrightarrow \forall\alpha\forall\beta(f_\delta$ maps α to $\beta \longleftrightarrow g_\delta$ maps β to $\alpha)$, otherwise $g_\delta = \delta 0$.

7.26) Dom $f_\mathfrak{f} = [\gamma_1,\gamma_2]$ & $f_\mathfrak{f}$ is one-to-one & $\gamma_1 \le \alpha \le \gamma_2$ & $f_\mathfrak{f}$ is differentiable at α & Deriv $(f_\mathfrak{f},\alpha) \neq \tilde{0} \longrightarrow f_\mathfrak{f}^{-1}$ is differentiable at $f_\mathfrak{f}(\alpha)$ & Deriv $(f_\mathfrak{f}^{-1}, f_\mathfrak{f}(\alpha)) = \tilde{1}/$Deriv $(f_\mathfrak{f},\alpha)$.

Proof. Let deriv $(f_\mathfrak{f},\alpha,f_0,\xi)$, so $\xi \neq \tilde{0}$; construct f_1 and g_1 as in the proof of (7.24), so $\mathfrak{f}g_1 = f_\mathfrak{f}^{-1}$. If $s_1 \in f_\mathfrak{f}(\alpha)$ & $s_2 \in f_\mathfrak{f}(\alpha)$ & $s_1 \in$ Dom g_1 & $s_2 \in$ Dom g_1 & $s_1 \neq s_2$, then $g_1(s_1) \in \alpha$ & $g_1(s_2) \in \alpha$ & $g_1(s_1) \in$ Dom f_1 & $g_1(s_2) \in$ Dom f_1 & $g_1(s_1) \neq g_1(s_2)$, so that $(f_1(g_1(s_1)) \hat{-} f_1(g_1(s_2))) \hat{/} (g_1(s_1) \hat{-} g_1(s_2)) \in \xi$. Since $f_1(g_1(s_1)) = s_1$ and $f_1(g_1(s_2)) = s_2$, it follows that $(g_1(s_1) \hat{-} g_1(s_2)) \hat{/} (s_1 \hat{-} s_2) \in \tilde{1}/\xi$. Therefore deriv $(f_\mathfrak{f}^{-1}, f_\mathfrak{f}(\alpha), g_1, \tilde{1}/\xi)$. ||

Integration

We turn our attention now to integration and the fundamental theorem of calculus. First a uniqueness condition:

7.27) Dom $g_\mathfrak{f}$ = Dom $h_\mathfrak{f} = [\alpha,\beta]$ & $\forall\gamma(\gamma \in [\alpha,\beta] \longrightarrow g_\mathfrak{f}$ and $h_\mathfrak{f}$ are differentiable at γ & Deriv $(g_\mathfrak{f},\gamma)$ = Deriv $(h_\mathfrak{f},\gamma))$ & $\exists\gamma(\gamma \in [\alpha,\beta]$ & $g_\mathfrak{f}(\gamma) = h_\mathfrak{f}(\gamma)) \longrightarrow g_\mathfrak{f} = h_\mathfrak{f}$.

Proof. By (7.9) and (7.10), the function $g_\mathfrak{f} - h_\mathfrak{f}$ has derivative $\tilde{0}$ at every $\gamma \in [\alpha,\beta]$. By the mean value theorem, $g_\mathfrak{f} - h_\mathfrak{f}$ is constant. Since $g_\mathfrak{f}$ and $h_\mathfrak{f}$ are equal at a point, they must therefore be equal everywhere. ||

The construction of the integral is contained in the proof of the following proposition.

7.28) $\alpha < \beta \,\&\, [\alpha,\beta] \subseteq \text{Dom } f_\delta \longrightarrow \exists g_\delta(\text{Dom } g_\delta = [\alpha,\beta] \,\&\, \forall\gamma(\gamma \in [\alpha,\beta] \longrightarrow$
g_δ is differentiable at γ & $\text{Deriv}(g_\delta,\gamma) = f_\delta(\gamma)) \,\&\, g_\delta(\alpha) = \tilde{0})$.

Proof. Let $f_\delta \upharpoonright [\alpha,\beta] = \delta f_0$. We may assume that the smallest element of $\text{Dom } f_0$ is $a \in \alpha$ and that the greatest is $b \in \beta$. The following argument is easily formalized in R_4 . Let $a = r_0, r_1, \ldots, r_k = b$ be a sequence enumerating the elements of $\text{Dom } f_0$ in increasing order, and define a function g_0 on $\text{Dom } f_0$ by setting $g(r_0) = \hat{0}$, $g(r_k) = \sum_{i=1}^{k} f(r_i) \hat{\cdot} (r_i \hat{-} r_{i-1})$ for $k = 1,\ldots,n$. If $0 \leq j < k \leq n$, then $(r_k \hat{-} r_j) \hat{\cdot} \text{Min } \{f_0(r_i) : j < i \leq k\} \leq g_0(r_k) - g_0(r_j)$ $\leq (r_k \hat{-} r_j) \hat{\cdot} \text{Max } \{f_0(r_i) : j < i \leq k\}$. It follows that g_0 is a real function and that if $r_j \sim r_k$, then $(g_0(r_k) \hat{-} g_0(r_j)) \hat{/} (r_k \hat{-} r_j) \in f_\delta(\tilde{r}_j)$. Let g_δ be δg_0 . ||

It follows from (7.27) that the function g_δ constructed in (7.28) is independent of the choice of f_0 .

7.29) Def $\int(f_\delta,\alpha,\beta) =_\delta g_\delta \longleftrightarrow \alpha < \beta \,\&\, [\alpha,\beta] \subseteq \text{Dom } f_\delta \,\&\, \text{Dom } g_\delta = [\alpha,\beta] \,\&$
$\forall\gamma(\gamma \in [\alpha,\beta] \longrightarrow g_\delta$ is differentiable at γ &
$\text{Deriv}(g_\delta,\gamma) = f_\delta(\gamma)) \,\&\, g_\delta(\alpha) = \tilde{0}$, otherwise $g_\delta = \delta 0$.

Observe that $\int(f_\delta,\alpha,\beta)$ is the *function* on $[\alpha,\beta]$ whose value at $\gamma \in [\alpha,\beta]$ is the real number we are accustomed to calling $\int_\alpha^\gamma f_\delta$. The reader should have no trouble proving that this value does not depend on β (as long as $\beta \geq \gamma$ and $[\alpha,\beta] \subseteq \text{Dom } f_\delta$) , as well as other basic properties of integrals. An attractive feature (not surprising in light of continuity) is that *every* function defined on $[\alpha,\beta]$ is integrable.

Higher derivatives and Taylor's theorem

The final topic of this section is Taylor's theorem. We start with a brief discussion of factorials.

7.30) Def $n! = k \longleftrightarrow (n=0 \ \&\ k=1) \vee \exists u\, (u$ is a sequence & $\mathrm{Ln}\ u = n \ \&\ u(1) = 1 \ \&\ \ \forall i (1 \leq i < n \longrightarrow u(i+1) = (i+1)\cdot u(i))\ \&$ $u(n) = k)$, otherwise $k = 0$.

The symbol ! is unbounded, of course; on the other hand, one can define a bounded function symbol Factlog such that Factlog $n = (\mathrm{Log}\ n)!$. (That Factlog is bounded follows from Factlog $n \leq$ Explog $(\mathrm{Log}\ n, n)$.) It follows that if $\varepsilon_{\nu+1}(n)$, then $\varepsilon_{\nu}(n!)$.

Now for a definition and a lemma.

7.31) Def $u_{S\mathfrak{s}}$ is a derivative sequence for $f_{\mathfrak{s}} \longleftrightarrow u_{S\mathfrak{s}}(1) = f_{\mathfrak{s}}$ &
$\forall i\ (1 \leq i \leq \mathrm{Ln}\ u_{S\mathfrak{s}} \longrightarrow \mathrm{Dom}\ u_{S\mathfrak{s}}(i) = \mathrm{Dom}\ f_{\mathfrak{s}})$ &
$\forall i\ (1 \leq i < \mathrm{Ln}\ u_{S\mathfrak{s}} \longrightarrow u_{S\mathfrak{s}}(i+1) = (u_{S\mathfrak{s}}(i))')$.

7.32) $\beta > \tilde{0}$ & $[\tilde{0},\beta] \subseteq \mathrm{Dom}\ f_{\mathfrak{s}}$ & $u_{S\mathfrak{s}}$ is a derivative sequence for $f_{\mathfrak{s}}$ &
$\varepsilon_{\nu+1}\ (\mathrm{Ln}\ u_{S\mathfrak{s}})$ & $\forall i\ (1 \leq i < \mathrm{Ln}\ u_{S\mathfrak{s}} \longrightarrow (u_{S\mathfrak{s}}(i))(\tilde{0}) = \tilde{0})$ &
$\forall \gamma(\tilde{0} < \gamma < \beta \longrightarrow (u_{S\mathfrak{s}}(\mathrm{Ln}\ u_{S\mathfrak{s}}))(\gamma) \geq \tilde{0}) \longrightarrow \forall\gamma(\tilde{0} < \gamma < \beta \longrightarrow$
$f\ (\gamma) \geq \tilde{0})$.

Proof. Let us write $f_{\mathfrak{s}}, f'_{\mathfrak{s}}, f^{(2)}_{\mathfrak{s}}, \ldots, f^{(n)}_{\mathfrak{s}}$ for the derivative sequence $u_{S\mathfrak{s}}$. By assumption, $\varepsilon_{\nu+1}(n)$, $f_{\mathfrak{s}}(\tilde{0}) = f'_{\mathfrak{s}}(\tilde{0}) = \ldots = f^{(n-1)}_{\mathfrak{s}}(\tilde{0}) = \tilde{0}$, and $f^{(n)}_{\mathfrak{s}}(\gamma) \geq \tilde{0}$ for all γ to the right of $\tilde{0}$ and sufficiently close to $\tilde{0}$; we are to show that $f_{\mathfrak{s}}$ has this last property as well.

The sequence $u_{S\delta}$ is represented by some sequence of real functions $f_0, f_1, \ldots, f_n$. It will suffice to show that if $e \hat{>} \hat{0}$ and $e \not\sim \tilde{0}$, then $f_0(r) \hat{+} e \hat{>} \hat{0}$ for all r in Dom f_0 such that $\hat{0} \hat{\leq} r \hat{\leq} b$, where b is some fraction representing β.

Choose a noninfinitesimal positive fraction d small enough that $d \hat{\cdot} (\hat{1} \hat{+} b \hat{+} \frac{1}{2!} b^2 \hat{+} \ldots \hat{+} \frac{1}{n!} b^n) \hat{\leq} e$; this is possible because $\varepsilon_{\nu+1}(n)$.

Define a new sequence of real functions $g_0, g_1, \ldots, g_n$ as follows: Dom g_i = Dom f_{n-i}, and for all r in the appropriate domains,

$$g_0(r) = f_n(r) \hat{+} d ,$$

$$g_1(r) = f_{n-1}(r) \hat{+} d \hat{+} dr ,$$

$$g_2(r) = f_{n-2}(r) \hat{+} d \hat{+} dr \hat{+} \frac{1}{2!} dr^2 ,$$

$$\vdots$$

$$g_n(r) = f_0(r) \hat{+} d \hat{+} dr \hat{+} \ldots \hat{+} \frac{1}{n!} dr^n .$$

Since $\delta f_0, \ldots, \delta f_n$ is a derivative sequence for $\delta f_0 = f_\delta$, it follows that $\delta g_n, \delta g_{n-1}, \ldots, \delta g_0$ is a derivative sequence for δg_n.

The formula $i \leq n \longrightarrow \forall r(r \in \text{Dom } g_i \ \& \ \hat{0} \hat{\leq} r \hat{\leq} b \longrightarrow g_i(r) \hat{>} \hat{0})$ is bounded; let us show that it is inductive in i. Since $f_\delta^{(n)} = \delta f_n \geq \tilde{0}$ between $\tilde{0}$ and β and since $d \hat{>} \hat{0}$ but $d \not\sim \hat{0}$, it follows that $g_0 \hat{>} \hat{0}$ between $\hat{0}$ and b. Now suppose the same is true of g_i, where $0 \leq i < n$, and consider g_{i+1}. Since $[\tilde{0}, \beta] \subseteq \text{Dom } \delta g_{i+1}$, the smallest r_0 such that $\hat{0} \hat{\leq} r_0$ and $r_0 \in \text{Dom } g_{i+1}$ is necessarily infinitesimal; since $\delta f_{n-i-1}(\tilde{0}) = \tilde{0}$ by hypothesis, it follows that $g_{i+1}(r_0) \sim d$, so in particular $g_{i+1}(r_0)$ is positive and noninfinitesimal.

Now $\delta g_i = (\delta g_{i+1})' \geq \tilde{0}$ between $\tilde{0}$ and β , so by (7.21) δg_{i+1} is nondecreasing on that interval. Thus $g_{i+1} \hat{>} \hat{0}$ between $\hat{0}$ and b .

By bounded induction, $g_n \hat{>} \hat{0}$ between $\hat{0}$ and b . On this interval, $f_0 \hat{+} e \hat{>} g_n$; hence the proof is complete.||

With (7.32) in hand, Taylor's theorem becomes straightforward. All we lack is the definition of the Taylor polynomial.

7.33) Def $P(f_\delta, u_{S\delta}) =_\delta g_\delta \longleftrightarrow \exists\alpha\exists\beta(\alpha < \tilde{0} < \beta$ & Dom $f_\delta = [\alpha,\beta])$ &
$u_{S\delta}$ is a derivative sequence for f_δ & $\varepsilon_{\nu+1}(\mathrm{Ln}\ u_{S\delta})$ &
Dom g_δ = Dom f_δ & $\exists v_{S\kappa}(\mathrm{Ln}\ v_{S\kappa} = \mathrm{Ln}\ u_{S\delta}$ &
$\forall i(1 \leq i \leq \mathrm{Ln}\ v_{S\kappa} \longrightarrow v_{S\kappa}(i) = (u_{S\delta}(i))(\tilde{0})/((i-1)!)^{\sim})$ &
$\forall\alpha(\alpha \in$ Dom $g_\delta \longrightarrow g_\delta(\alpha) =$ Polyvalue $(v_{S\kappa},\alpha)))$, otherwise $g_\delta = \delta 0$.

Note that if $\varepsilon_{\nu+1}(i)$, then $((i-1)!)^{\wedge}$ is limited, so the existence of $v_{S\kappa}$ in (7.33) presents no problem.

We are now ready to state Taylor's theorem. For simplicity, we are considering only Taylor polynomials centered at $\alpha = \tilde{0}$; the results extend easily to other values of α .

7.34) $\alpha < \tilde{0} < \beta$ & $[\alpha,\beta] \subseteq$ Dom f_δ & $u_{S\delta}$ is a derivative sequence for f_δ &
$\varepsilon_{\nu+1}(\mathrm{Ln}\ u_{S\delta})$ & $\forall\gamma(\gamma \in [\alpha,\beta] \longrightarrow |(u_{S\delta}(\mathrm{Ln}\ u_{S\delta}))(\gamma)| \leq \xi)$ &
$\mathrm{Ln}\ w_{S\kappa} = \mathrm{Ln}\ u_{S\delta}$ & $\forall i(1 \leq i < \mathrm{Ln}\ u_{S\delta} \longrightarrow w_{S\kappa}(i) = \tilde{0})$ &
$w_{S\kappa}(\mathrm{Ln}\ u_{S\delta}) = \tilde{1} \longrightarrow \forall\gamma(\gamma \in [\alpha,\beta] \longrightarrow$
$|f_\delta(\gamma) - (P(f_\delta, u_{S\delta}[1,\mathrm{Ln}\ u_{S\delta}-1]))(\gamma)|/|$Polyvalue $(w_{S\kappa},\gamma)|$
$\leq \xi/((\mathrm{Ln}\ u_{S\delta}-1)!)^{\sim})$.

Proof. In something more nearly resembling English, the assertion is that if $f_{\mathit{f}}, f'_{\mathit{f}}, \ldots, f_{\mathit{f}}^{(n+1)}$ is a derivative sequence for a function f_{f} defined in some interval around $\tilde{0}$, and if g_{f} is defined in that interval by $g_{\mathit{f}}(\gamma) = f_{\mathit{f}}(\tilde{0}) + f'_{\mathit{f}}(\tilde{0})\cdot\gamma + f_{\mathit{f}}^{(2)}(\tilde{0})\cdot\gamma^2/\tilde{2} + \ldots + f_{\mathit{f}}^{(n)}(\tilde{0})\cdot\gamma^n/(n!)^{\sim}$, then $|f_{\mathit{f}}(\gamma) - g_{\mathit{f}}(\gamma)|/|\gamma^{n+1}| \leq \xi/((n+1)!)^{\sim}$, where ξ is the maximum value attained by $|f_{\mathit{f}}^{(n+1)}|$.

Define a function h_{f} on $[\alpha, \beta]$ by $h_{\mathit{f}}(\gamma) = g_{\mathit{f}}(\gamma) + \xi\cdot\gamma^{n+1}/((n+1)!)^{\sim} - f_{\mathit{f}}(\gamma)$. We know that f_{f} has $n+1$ derivatives, and the polynomial $g_{\mathit{f}}(\gamma) + \xi\cdot\gamma^{n+1}/((n+1)!)^{\sim}$ certainly does, so there exists a derivative sequence $h_{\mathit{f}}, h'_{\mathit{f}}, \ldots, h_{\mathit{f}}^{(n+1)}$. It is easy to check that $h_{\mathit{f}}(\tilde{0}) = h'_{\mathit{f}}(\tilde{0}) = \ldots = h_{\mathit{f}}^{(n)}(\tilde{0}) = \tilde{0}$ and that $h_{\mathit{f}}^{(n+1)}(\gamma) = \xi - f_{\mathit{f}}^{(n+1)}(\gamma) \geq \tilde{0}$. By (7.32), $h_{\mathit{f}} \geq \tilde{0}$ to the right of $\tilde{0}$. This gives $(f_{\mathit{f}}(\gamma) - g_{\mathit{f}}(\gamma))/\gamma^{n+1} \leq \xi/((n+1)!)^{\sim}$ to the right of $\tilde{0}$. The lower bound on $f_{\mathit{f}} - g_{\mathit{f}}$, and the bounds to the left of $\tilde{0}$, are established similarly. ||

§8. *Further Properties of Real Numbers*

In this section we extend the results of the preceding sections; in particular, we discuss rational, algebraic, and transcendental numbers and decimal expansions. If the reader finds some of the results here less appealing than those in the earlier sections, it may be because the correspondence with classical mathematics is less close. To increase readability, we have opted for a slightly less formal style of presentation than that to which we have grown accustomed; for instance, if f is a polynomial and r_1 and r_2 are fractions we shall take the liberty of writing $f(r_1+r_2)$ rather than Polyvalue $(f, r_1 \hat{+} r_2)$. It should be clear that all of our results can be formalized in our current theory R_4.

Natural and rational numbers

One of the interesting features of the number system emerging in R_4 is that we have several choices when it comes to defining natural numbers. The real numbers include $\tilde{0}, \tilde{1}, \ldots$, and $\tilde{n}$ for all n such that $\varepsilon_\nu(n)$, and these may appear to be the obvious candidates for the natural numbers. On the other hand, if we really expect a number to be "finite", then it should satisfy not only ε_ν but also $\varepsilon_{\nu+1}, \varepsilon_{\nu+2}, \ldots$. It turns out that for many purposes the definition

8.1) Def α is a natural number $\longleftrightarrow \exists n\, (\varepsilon_{\nu+1}(n)\ \&\ \alpha = \tilde{n})$

makes good sense. We already know, for instance, that $\varepsilon_{\nu+1}(n)$ is required in order for polynomials of degree n, or even the notion of raising a real number to the nth power, to behave properly.

On the heels of (8.1) follows the definition

8.2) Def β is rational $\longleftrightarrow \exists\alpha_1 \exists\alpha_2$ (α_1 and α_2 are natural numbers & $(\beta = \alpha_1/\alpha_2 \vee \beta = -\alpha_1/\alpha_2))$,

or equivalently

8.3) β is rational $\longleftrightarrow \exists r$ ($r \in \beta$ & $\varepsilon_{\nu+1}$ (Numer r) & $\varepsilon_{\nu+1}$ (Denom r)) . ‖

The real number $\sqrt{2}$, which exists by (5.52), is irrational; in fact, it is irrational in the strong sense that it is not represented by any fraction a/b with $\varepsilon_\nu(a)$ & $\varepsilon_\nu(b)$. Indeed, assume such a representation, with a/b in lowest terms. Then $a^2/b^2 = 2+d$ for some infinitesimal d , so $a^2 = 2b^2+db^2$. But db^2 is infinitesimal and a^2 and $2b^2$ are both integers (fractions with denominator 1), so it must be the case that $a^2 = 2b^2$. It follows that a is even, then that b is even, a contradiction.

It is admittedly a bit disturbing that the rational numbers, according to (8.2), are not cofinal in the ordering of the real numbers: if $\neg\varepsilon_{\nu+1}(n)$ & $\alpha > \tilde{n}$, then α is simply too big to be rational. It seems prudent to avoid these dangerous outer reaches of the number line and to restrict our attention to more manageable numbers, say in the unit interval.

Can we use a cardinality argument to prove the existence of transcendental numbers in the unit interval? First we must decide what this means. An algebraic number should be a root of a polynomial whose coefficients are natural numbers and whose degree is finite. That is, the coefficients should satisfy $\varepsilon_{\nu+1}$ and the degree should be subject to even further restriction. With these agreements, we can answer our question in the affirmative; to do so requires a few more facts about polynomials.

Roots of polynomials

Let $f(x) = x^n+a_{n-1}x^{n-1}+\ldots+a_1x+a_0$ be a monic polynomial (strictly speaking, a sequence of fractions whose last term is $\hat{1}$), and let b be a fraction. Let f_b be the monic polynomial

$$f_b(x) = x^{n-1}+c_{n-2}x^{n-2}+\ldots+c_1x+c_0 \text{ , where}$$

$$c_0 = b^{n-1}+a_{n-1}b^{n-2}+\ldots+a_3b^2+a_2b+a_1 \text{ ,}$$

$$c_1 = b^{n-2}+a_{n-1}b^{n-3}+\ldots+a_3b+a_2 \text{ ,}$$

$$\vdots$$

$$c_{n-3} = b^2+a_{n-1}b+a_{n-2} \text{ ,}$$

$$c_{n-2} = b+a_{n-1} \text{ .}$$

If f is not a monic polynomial, or if f is the monic polynomial 1, let f_b be the zero polynomial. This defines a bounded binary function symbol (the arguments are f and b); the idea is that $f_b(x)$ is the quotient when $f(x)$ is divided by $x-b$. In fact, let us establish

8.4) f is a monic polynomial & b is a fraction $\longrightarrow$

$$f(x) = f_b(x)\cdot(x-b)+f(b) \text{ .}$$

Proof. First let us understand the statement. The conclusion of (8.4) is a polynomial identity: the assertion that two sequences coincide. The right side, of course, is the polynomial that results when the fraction $f(b)$ is added to the constant term of the polynomial product $f_b(x)\cdot(x-b)$.

Now that we know what we are to prove, we proceed by bounded induction on f. (The formula *is* bounded.) Let $\hat{f}(x)$ be $x^{n-1}+a_{n-1}x^{n-2}+\ldots+a_2x+a_1$, so that

1) $$f(x) = \hat{f}(x) \cdot x + a_0$$

and

2) $$\hat{f}(b) = c_0 .$$

Note also that

3) $$\hat{f}_b(x)\cdot x = f_b(x) - c_0 .$$

Certainly $\hat{f} < f$, so by the induction hypothesis,

4) $$\hat{f}(x) = \hat{f}_b(x)\cdot(x-b) + \hat{f}(b) .$$

Combining threads, we have

$$\begin{aligned}
f(x) &= \hat{f}(x)\cdot x + a_0 && \text{(by (1))}\\
&= \hat{f}_b(x)\cdot x\cdot(x-b) + \hat{f}(b)\cdot x + a_0 && \text{(by (4))}\\
&= (f_b(x)-c_0)\cdot(x-b) + c_0\cdot x + a_0 && \text{(by (3) and (2))}\\
&= f_b(x)\cdot(x-b) + (c_0\cdot b+a_0)
\end{aligned}$$

$$= f_b(x)\cdot(x-b) + (\hat{f}(b)\cdot b + a_0) \qquad \text{by (2))}$$

$$= f_b(x)\cdot(x-b) + f(b) \qquad \text{by (1))} \ . \ \|$$

If $b_1,\ldots,b_k$ is a sequence of fractions (that is, if b' is a sequence of fractions and $\mathrm{Ln}\ b' = k$) , then we can iterate the above procedure to obtain the polynomial $f_{b_1\ldots b_k}$. This is another bounded binary function symbol; the arguments are the polynomial f and the sequence b' . The analog of (8.4) for $f_{b_1\ldots b_k}$ is

8.5) f is a monic polynomial & $b_1,\ldots,b_k$ is a sequence of fractions $\longrightarrow$

$$\begin{aligned} f(x) = {} & f_{b_1\ldots b_k}(x)\cdot(x-b_1)\cdot(x-b_2)\ldots(x-b_k) \\ & + f_{b_1\ldots b_{k-1}}(b_k)\cdot(x-b_1)\ldots(x-b_{k-1}) \\ & + f_{b_1\ldots b_{k-2}}(b_{k-1})\cdot(x-b_1)\ldots(x-b_{k-2}) \\ & + \ldots \\ & + f_{b_1 b_2}(b_3)\cdot(x-b_1)\cdot(x-b_2) \\ & + f_{b_1}(b_2)\cdot(x-b_1) \\ & + f(b_1) \ . \ \| \end{aligned}$$

It should be clear that (8.5) can be formulated in our theory; the proof is by bounded induction on the sequence b' , using (8.4).

8.6) f is a monic polynomial & $b_1,\ldots,b_k$ is a sequence of fractions & $\forall i\ (1 \le i \le k \longrightarrow |f(b_i)| \le d)$ & $\forall i \forall j\ (1 \le i < j \le k \longrightarrow |b_i - b_j| \ge e > 0) \longrightarrow |f_{b_1 \ldots b_{k-1}}(b_k)| \le (\frac{2}{e})^{k-1} \cdot d$.

Proof by bounded induction on k (really on b') . If $k = 1$, the assertion is $|f(b_1)| \le d$, which is part of the hypothesis. For the induction step, we have

$$|f_{b_1 \ldots b_{k-1}}(b_k)| = \left| \frac{f_{b_1 \ldots b_{k-2}}(b_k) - f_{b_1 \ldots b_{k-2}}(b_{k-1})}{b_k - b_{k-1}} \right|$$

(by (8.4) with $f_{b_1 \ldots b_{k-2}}(b_k)$ and b_{k-1} in place of of $f(x)$ and b)

$$\le \frac{(\frac{2}{e})^{k-2} \cdot d + (\frac{2}{e})^{k-2} \cdot d}{e}$$

(by the induction hypothesis and $|b_k - b_{k-1}| \ge e$)

$$= (\frac{2}{e})^{k-1} \cdot d \ . \ \|$$

Our objective at this point is the theorem that a polynomial f whose coefficients are real numbers and whose degree satisfies $\varepsilon_{\nu+1}$ can have at most Deg f roots. Here is a preliminary version:

8.7) $f(x) = x^k + a_{k-1}x^{k-1} + \ldots + a_0$ is a monic polynomial & Deg $f = k$ & $\varepsilon_{\nu+1}(k)$ & $\forall i\ (0 \leq i \leq k-1 \longrightarrow a_i$ is limited) & $b_1, \ldots, b_k$ is a sequence of limited fractions & $\forall i (1 \leq i \leq k \longrightarrow f(b_i) \sim 0)$ & $\forall i \forall j\ (1 \leq i < j \leq k \longrightarrow b_i \not\sim b_j)$ & b is a limited fraction & $\forall i\ (1 \leq i \leq k \longrightarrow b_i \not\sim b) \longrightarrow f(b) \not\sim 0$.

Proof. Let d be the largest value of any $|f(b_i)|$, $1 \leq i \leq k$, and e the smallest value of any $|b_i - b_j|$, $1 \leq i < j \leq k$. By hypothesis $d \sim 0$ and $e \not\sim 0$. By (8.5),

$$
\begin{aligned}
f(b) = {} & f_{b_1 \ldots b_k}(b) \cdot (b-b_1) \ldots (b-b_k) \\
& + f_{b_1 \ldots b_{k-1}}(b_k) \cdot (b-b_1) \ldots (b-b_{k-1}) \\
& + \ldots \\
& + f_{b_1}(b_2) \cdot (b-b_1) \\
& + f(b_1) .
\end{aligned}
$$

Now $f_{b_1 \ldots b_k}(x)$ is a monic polynomial of degree 0, namely 1.

Since $b-b_1 \not\sim 0, \ldots, b-b_k \not\sim 0$, and $\varepsilon_{\nu+1}(k)$, the first term in the above sum is not infinitesimal. But by (8.6), we have $|f(b_1)| \leq d$, $|f_{b_1}(b_2)| \leq \frac{2}{e} \cdot d, \ldots,$ $|f_{b_1 \ldots b_{k-1}}(b_k)| \leq (\frac{2}{e})^{k-1} \cdot d$, and all of these numbers *are* infinitesimal. Hence every term other than the first is infinitesimal, and $f(b) \not\sim 0$. ||

Of course, the assumption that f is monic is easily eliminated at this point. Also, the result can be reformulated in terms of the sort $\mathcal{S}r$ as follows:

8.8) $\varepsilon_{\nu+1}(\mathrm{Ln}\ f_{\mathcal{S}r})\ \&\ \forall i\ (1 \leq i \leq \mathrm{Ln}\ u_{\mathcal{S}r} \longrightarrow \mathrm{Polyvalue}\ (f_{\mathcal{S}r}, u_{\mathcal{S}r}(i)) = \tilde{0})\ \&$
$\forall i \forall j\ (1 \leq i < j \leq \mathrm{Ln}\ u_{\mathcal{S}r} \longrightarrow u_{\mathcal{S}r}(i) \neq u_{\mathcal{S}r}(j) \longrightarrow \mathrm{Ln}\ u_{\mathcal{S}r} <$
$\mathrm{Ln}\ f_{\mathcal{S}r}$. ||

The existence of transcendental numbers

We are now ready to tackle the problem of the transcendental numbers. What we can prove is that if the degrees of our polynomials are required to satisfy $\varepsilon_{\nu+2}$ and the coefficients $\varepsilon_{\nu+1}$, then most numbers in the unit interval are transcendental. To make this precise, let m be such that $\varepsilon_{\nu}(m)\ \&\neg\ \varepsilon_{\nu+1}(m)$, and consider polynomials whose degrees are at most $\mathrm{Log}\ m$ and whose coefficients are integer fractions between $-m$ and m; these polynomials include all those previously mentioned. All such polynomials can be listed in a sequence in such a way that each f appears $\mathrm{Deg}\ f$ times, and the length of this sequence, say M, satisfies $\varepsilon_{\nu}(M)$. Let a be the set of fractions $\{0, \frac{1}{M}, \frac{2}{M}, \ldots, \frac{M-1}{M}, 1\}$; note that each of these fractions represents a different real number. For each of our polynomials f, let $Z(f)$ be the subset of a with cardinality $\mathrm{Deg}\ f$ whose elements are those $\mathrm{Deg}\ f$ elements of a at which the value of f is smallest in absolute value. By (8.8), at most $\mathrm{Deg}\ f$ of the numbers $0, \frac{1}{M}, \ldots, 1$ can satisfy $f(x) \sim 0$; hence all such x are elements of $Z(f)$. It follows that every algebraic number among $0, \frac{1}{M}, \ldots, 1$ is in $Z(f)$ for some f. But we can form the union over all f of the sets $Z(f)$, and its cardinality is at most M, by the definition of M. Therefore at least one of the $M+1$ elements of a is transcendental. By replacing M with larger numbers in the construction of a, one can easily obtain a sequence of transcendental numbers of any desired length satisfying ε_ν. On the other hand, every sequence of distinct algebraic numbers has length at most M; in fact, by replacing m (and therefore M) with *smaller* numbers, one can see that the length of a sequence of distinct algebraic numbers must always satisfy $\varepsilon_{\nu+1}$.

Let us summarize.

8.9) Def α is algebraic $\longleftrightarrow \exists f_{Sr}(\varepsilon_{\nu+2}(\text{Ln } f_{Sr})$ & $\forall i\ (1 \le i \le \text{Ln } f_{Sr} \longrightarrow$ $\exists k\ (\varepsilon_{\nu+1}(k)$ & $(f_{Sr}(i) = \tilde{k} \vee f_{Sr}(i) = -\tilde{k})))$ & Polyvalue $(f_{Sr}, \alpha) = \tilde{0})$.

8.10) $\varepsilon_\nu(M) \longrightarrow \exists u_{Sr}$ (Ln $u_{Sr} = M$ & $\forall i \forall j\ (1 \le i < j \le M \longrightarrow$ $u_{Sr}(i) \ne u_{Sr}(j))$ & $\forall i\ (1 \le i \le M \longrightarrow \tilde{0} \le u_{Sr}(i) \le \tilde{1}$ & $\neg(u_{Sr}(i)$ is algebraic))). ||

8.11) $\forall i \forall j\ (1 \le i < j \le \text{Ln } u_{Sr} \longrightarrow u_{Sr}(i) \ne u_{Sr}(j))$ & $\forall i\ (1 \le i \le \text{Ln } u_{Sr} \longrightarrow u_{Sr}(i)$ is algebraic) $\longrightarrow$ $\varepsilon_{\nu+1}(\text{Ln } u_{Sr})$. ||

The proofs used above to establish (8.10) and (8.11) are actually rather simple examples of cardinality arguments that we shall investigate more thoroughly in Part Three.

Decimal expansions

With (8.1) as motivation, we define what it means for a sequence of fractions to converge to a real number.

8.12) Def u converges to $\alpha \longleftrightarrow$ u is a sequence of fractions &

$\neg\varepsilon_{\nu+1}$ (Ln u) & $\forall i$ ($i \leq$ Ln u & $\neg\varepsilon_{\nu+1}(i) \longrightarrow u(i) \in \alpha$) .

Now we discuss decimal expansions of numbers in the unit interval.

8.13) Def u is an m-ary expansion of $\alpha \longleftrightarrow m \geq 2$ & u is a sequence of fractions & $\forall i$ ($1 \leq i \leq$ Ln u $\longrightarrow \exists k$ ($k < m$ & $u(i) = \hat{k}$)) & $u(1) = \hat{0}$ & Polyvalue $(u,\hat{1}\hat{/}\hat{m}) \in \alpha$.

8.14) u is an m-ary expansion of $\alpha \longrightarrow \tilde{0} \leq \alpha \leq \tilde{1}$.

Proof. Easy formalization of the argument

$$0 = \sum_{i=1}^{n} 0m^{-i} \leq \sum_{i=1}^{n} u_i m^{-i} \leq \sum_{i=1}^{n} (m-1)m^{-i} < 1 . \quad \|$$

8.15) u is an m-ary expansion of α & $\neg\varepsilon_{\nu+1}$ (Ln u) $\longrightarrow$ Polyvalseq $(u,\hat{1}\hat{/}\hat{m})$ converges to α .

Proof. It suffices to show that if $\neg\varepsilon_{\nu+1}(i)$ but $i \leq$ Ln u = n , then Polyvalue $(u,\hat{1}\hat{/}\hat{m}) \hat{-}$ (Polyvalseq $(u,\hat{1}\hat{/}\hat{m})$)(i) is infinitesimal. This quantity is nonnegative but at most $\sum_{j=i}^{n-1} (m-1)m^{-j} = m^{-i+1}-m^{-n+1}$, which is infinitesimal because $\neg\varepsilon_{\nu+1}(i)$ & $\neg\varepsilon_{\nu+1}(n)$ & $m \geq 2$. $\|$

Now for uniqueness and existence theorems.

8.16) $\varepsilon_{\nu}(m)$ & u and v are m-ary expansions & α & $\neg\varepsilon_{\nu+1}$ (Ln u) & $\neg\varepsilon_{\nu+1}$ (Ln v) $\longrightarrow$ $\forall i\ (i \geq 1\ \&\ \varepsilon_{\nu+1}(i) \longrightarrow u(i) = v(i))\ \vee$

$\exists j\ (\varepsilon_{\nu+1}(j)\ \&\ \ \forall i\ (1 \leq i < j \longrightarrow u(i) = v(i))\ \&$

$((v(j) = u(j)\hat{+}\hat{1}\ \&\ \forall i\ (i > j\ \&\ \varepsilon_{\nu+1}(i) \longrightarrow u(i) = (m-1)\hat{}\ \&$

$v(i) = \hat{0})\ \vee\ (u(j) = v(j)\hat{+}\hat{1}\ \&\ \forall i\ (i > j\ \&\ \varepsilon_{\nu+1}(i) \longrightarrow$

$v(i) = (m-1)\hat{}\ \&\ u(i) = \hat{0})))$.

Proof. If u and v do not agree on all i with $\varepsilon_{\nu+1}(i)$, then by (BLNP) there is a smallest j such that $u(j) \neq v(j)$. We may assume $u(j) \hat{<} v(j)$. If $u(j)\hat{+}\hat{1} \hat{<} v(j)$, then Polyvalue $(u,\hat{1}\hat{/}\hat{m})$ differs from Polyvalue $(v,\hat{1}\hat{/}\hat{m})$ by at least m^{-j+1} . Because $\varepsilon_{\nu+1}(j)$, though, m^{-j+1}, is not infinitesimal, so this is impossible since Polyvalue $(u,\hat{1}\hat{/}\hat{m})$ and Polyvalue $(v,\hat{1}\hat{/}\hat{m})$ represent the same real number. Hence $u(j)\hat{+}\hat{1} = v(j)$. Suppose i is the smallest number $> j$ such that either $u(i) \neq (m-1)\hat{}$ or $v(i) \neq \hat{0}$. Then Polyvalue $(u,\hat{1}\hat{/}\hat{m})$ and Polyvalue $(v,\hat{1}\hat{/}\hat{m})$ differ by at least m^{-i+1} ; for this to be infinitesimal, it must be the case that $\neg\varepsilon_{\nu+1}(i)$. Thus $\forall i\ (i > j\ \&\ \varepsilon_{\nu+1}(i) \longrightarrow u(i) = (m-1)\hat{}\ \&\ v(i) = \hat{0})$. ||

8.17) Def u is an m-ary approximating sequence for r $\longleftrightarrow$

$m \geq 2$ & r is a fraction & u is a sequence of fractions &

$\forall i\ (1 \leq i \leq \text{Ln } u \longrightarrow \exists k\ (k < m\ \&\ u(i) = \hat{k}))\ \&\ u(1) = \hat{0}\ \&$

$\forall i\ (1 \leq i \leq \text{Ln } u \longrightarrow (\text{Polyvalseq } (u,\hat{1}\hat{/}\hat{m}))(i) \hat{\leq} r$

$\hat{<} (\text{Polyvalseq } (u,\hat{1}\hat{/}\hat{m}))(i) \hat{+} (\text{Powerseq } (\hat{1}\hat{/}\hat{m},u))(i))$.

8.18) $m \geq 2$ & r is a fraction & $\hat{0} \hat{\leq} r \hat{<} \hat{1} \longrightarrow \exists u$ (u is an m-ary approximating sequence for r & Ln u = Log n)..

Proof. By bounded induction on n . ||

8.19) u and v are m-ary approximating sequences for r & $\mathrm{Ln}\, u = \mathrm{Ln}\, v \longrightarrow u = v$. ||

8.20) Def Approxseq $(m,r,n) = u \longleftrightarrow u$ is an m-ary approximating sequence for r & $\mathrm{Ln}\, u = \mathrm{Log}\, n$, otherwise $u = 1$.

Since $\mathrm{Ln}\, u = \mathrm{Log}\, n$ and $\mathrm{Sup}\, u \leq (m-1)\hat{}$, it follows that the function symbol Approxseq is bounded.

8.21) $m \geq 2$ & r is a fraction & $\hat{0} \leq r \mathbin{\hat{<}} \hat{1}$ & $r \in \alpha$ & $\neg\varepsilon_\nu(n) \longrightarrow$ Approxseq (m,r,n) is an m-ary expansion of α .

Proof. If $i \leq \mathrm{Log}\, n$ & $\neg\,\varepsilon_{\nu+1}(i)$, then the difference between r and (Polyvalseq(Approxseq (m,r,n) ,$\hat{1}\hat{/}\hat{m}$))(i) is at most m^{-i+1} , which is infinitesimal. ||

Decimal expansions give us a new technique for forming sets of real numbers. For instance, take n such that $\varepsilon_\nu(n)$ & $\neg\,\varepsilon_{\nu+1}(n)$, and let a be the set of all fractions that are sums of ternary expansions of length n all of whose coefficients are 0 or 2 . Then $\mathcal{s}a$ is the Cantor set.

§9. *The p-adic Numbers*

The construction of the real numbers in §5 applies *mutatis mutandis* to the p-adic numbers as well. For the most part the imitation is straightforward, and we shall omit some of the easier proofs.

Another kind of infinitesimal

As before, we begin with a few definitions involving fractions.

9.1) Def a is a power of $p \longleftrightarrow \exists k\ (a = \text{Explog}\ (p,k))$.

9.2) $p \neq 0 \longrightarrow$ (a is a power of $p \longleftrightarrow a \neq 0\ \&\ \exists k\ (a = p^k)$) . ||

9.3) Def Val $(p,a) = r \longleftrightarrow p$ is a prime & $((a = 0\ \&\ r = \hat{0}) \vee$
$(a \neq 0\ \&\ \exists b$ (b is a power of p & $b | a\ \&\ b\cdot p \nmid a\ \&\ r = \hat{1}\hat{/}\hat{b})))$,
otherwise $r = 0$.

9.4) Def Value (p,r) = Val $(p,\text{Numer}\ r)\hat{/}$Val $(p,\text{Denom}\ r)$.

9.5) p is a prime & r is a fraction $\longrightarrow$ Value $(p,r)\ \hat{\geq}\ \hat{0}$. ||

9.6) p is a prime & r is a fraction $\longrightarrow$ (Value $(p,r) = \hat{0} \longleftrightarrow r = \hat{0}$). ||

9.7) p is a prime & r and s are fractions $\longrightarrow$
Value $(p,r\hat{\cdot}s)$ = Value $(p,r)\hat{\cdot}$Value (p,s) . ||

9.8) p is a prime & r and s are fractions & Value (p,r)
$\hat{\leq}$ Value $(p,s) \longrightarrow$ Value $(p,r\hat{+}s)\ \hat{\leq}$ Value (p,s) . ||

9.9) p is a prime & r and s are fractions & Value (p,r)
$\hat{<}$ Value $(p,s) \longrightarrow$ Value $(p,r\hat{+}s)$ = Value (p,s) . ||

The symbol Value is bounded; the following symbols are unbounded.

9.10) Def r is p-limited $\longleftrightarrow$ p is a prime & r is a fraction & Value (p,r) is limited.

9.11) Def r is p-unlimited $\longleftrightarrow$ p is a prime & r is a fraction & Value (p,r) is unlimited.

9.12) Def r is p-infinitesimal $\longleftrightarrow$ p is a prime & r is a fraction & Value (p,r) is infinitesimal.

9.13) Def $\sim(p,r,s)$ $\longleftrightarrow$ p is a prime & r and s are fractions & $r \hat{-} s$ is p-infinitesimal.

We shall generally write $r \sim_p s$ rather than $\sim(p,r,s)$.

The following properties follow easily from (9.5)-(9.9).

9.14) r and s are p-limited $\longrightarrow$ $\hat{-}r$, $r \hat{+} s$, and $r \hat{\cdot} s$ are p-limited. ||

9.15) r and s are p-infinitesimal $\longrightarrow$ $\hat{-}r$ and $r \hat{+} s$ are p-infinitesimal. ||

9.16) r is p-limited & s is p-infinitesimal $\longrightarrow$ $r \hat{\cdot} s$ is p-infinitesimal. ||

9.17) p is a prime & r is a fraction $\longrightarrow$ $r \sim_p r$. ||

9.18) $r \sim_p s \longrightarrow s \sim_p r$. ||

9.19) $r \sim_p s$ & $s \sim_p t \longrightarrow r \sim_p t$. ||

9.20) $r_1 \sim_p s_1$ & $r_2 \sim_p s_2 \longrightarrow r_1 \hat{+} r_2 \sim_p s_1 \hat{+} s_2$. ||

9.21) r_1 and r_2 are p-limited & $r_1 \sim_p s_1$ & $r_2 \sim_p s_2 \longrightarrow$ $r_1 \hat{\cdot} r_2 \sim_p s_1 \hat{\cdot} s_2$. ||

9.22) r_1 is p-limited & $\neg(r_2$ is p-infinitesimal) &
$r_1 \sim_p s_1$ & $r_2 \sim_p s_2 \longrightarrow r_1 \hat{/} r_2 \sim_p s_1 \hat{/} s_2$. ||

9.23) p is a prime & r is a fraction $\longrightarrow$ (r is p-infinitesimal $\longleftrightarrow$ $\hat{1}\hat{/}r$ is p-unlimited). ||

A theory of p-adic numbers, and an interpretation

The most obvious way to parallel the construction of the real numbers in the p-adic case would be to adjoin a new sort for each prime p . Actually, it is possible to handle the p-adics for all p with a single sort p ; the secret is to let each p-adic numer contain a piece of information indicating what p is. Pending a convention to be introduced later, we use Greek letters with the subscript p for variables of sort p . The sort p comes equipped with the following accessories: a function symbol Prime of type (p,n) ; a predicate symbol ϵ of type (n,n,p) (we write $r \epsilon_p \alpha_p$ rather than $\epsilon(p,r,\alpha_p)$) ; and three new nonlogical axioms:

9.24) Ax Prime $(\alpha_p) = p \longrightarrow$ p is a prime & $\exists r$ (r is a p-limited fraction & $\forall s \forall q$ ($s \epsilon_q \alpha_p \longleftrightarrow q = p$ & $s \sim_p r$)) .

9.25) Ax p is a prime & r is a p-limited fraction $\longrightarrow$ $\exists \alpha_p$ (Prime $(\alpha_p) = p$ & $r \epsilon_p \alpha_p$) .

9.26) Ax $r \epsilon_p \alpha_p$ & $r \epsilon_p \beta_p \longleftrightarrow \alpha_p =_p \beta_p$.

Let $R_p^{\mu\nu}$ be the extension of the theory $R_4^{\mu\nu}$ obtained in this way.

Inasmuch as this is not just a simple equivalence-class construction, it must be checked that R_p is interpretable in R_4 or in $\tilde{Q}^\mu$. Actually, it is not difficult to extend the interpretation of R_4 in $\tilde{Q}^\mu$ constructed in §6 to an interpretation of R_p in $\tilde{Q}^\mu$. Define $U_p x \longleftrightarrow \exists p \exists r$ (p is a prime & r is a p-limited fraction & $x = \langle p,r \rangle$) , $x(=_p)_I y \longleftrightarrow \exists p\ (Proj_1 x = Proj_1 y = p\ \&\ Proj_2 x \sim_p Proj_2 y)$, $Prime_I(x) = Proj_1 x$, and $\epsilon_I(p,r,x) \longleftrightarrow p = Proj_1 x\ \&\ r \sim_p Proj_2\ x$. Conditions (3.1)-(3.5) and the interpretations of (9.24)-(9.26) are easily checked.

Arithmetic of p-adic numbers

Many simple theorems follow immediately from axioms (9.24)-(9.26). For instance, by (9.24),

9.27) $r \epsilon_p \alpha_p \longrightarrow Prime\ (\alpha_p) = p$. ||

9.28) Def $0_p =_p \alpha_p \longleftrightarrow$ (p is a prime & $\hat{0}\ \epsilon_p\ \alpha_p) \vee$
$(\neg(p$ is a prime) & $\hat{0}\ \epsilon_2\ \alpha_p)$.

9.29) Def $p(p,r) =_p \alpha_p \longleftrightarrow$ p is a prime & r is a p-limited fraction &
$r\ \epsilon_p\ \alpha_p$, otherwise $\alpha_p = 0_p$.

Now for the promised notational convention. When no confusion is likely, we reduce use of the cumbersome subscript p and function symbol Prime by writing α_p ,β_q,... for p-adic numbers whose primes are p,q,... . A formula of the form $\mathbb{D}[\alpha_p,p]$ can be taken to mean $\mathbb{D}[\alpha_p, Prime\ (\alpha_p)]$; thus the definition

9.30) Def $\alpha_{p_1} + \beta_{p_2} =_p \gamma_{p_3} \longleftrightarrow p_1 = p_2 = p_3$ & $\exists r \exists s (r \in_{p_1} \alpha_{p_1}$ &
$s \in_{p_2} \beta_{p_2}$ & $r \hat{+} s \in_{p_3} \gamma_{p_3})$, otherwise $\gamma_{p_3} = 0_2$

is understood to mean

Def $\alpha_p + \beta_p =_p \gamma_p \longleftrightarrow$ Prime (α_p) = Prime (β_p) = Prime (γ_p) &
$\exists r \exists s$ ($\in$(Prime $(\alpha_p), r, \alpha_p)$ & $\in$(Prime $(\beta_p), s, \beta_p)$ &
$\in$(Prime $(\gamma_p), r \hat{+} s, \gamma_p)$) , otherwise $\gamma_p = 0_2$.

The uniqueness condition for (9.30) follows from (9.20).

9.31) Def $-\alpha_{p_1} =_p \beta_{p_2} \longleftrightarrow p_1 = p_2$ & $\exists r$ $(r \in_{p_1} \alpha_{p_1}$ & $\hat{-} r \in_{p_2} \beta_{p_2})$.

9.32) Def $\alpha_{p_1} \cdot \beta_{p_2} =_p \gamma_{p_3} \longleftrightarrow p_1 = p_2 = p_3$ & $\exists r \exists s$ $(r \in_{p_1} \alpha_{p_1}$ &
$s \in_{p_2} \alpha_{p_2}$ & $r \hat{\cdot} s \in_{p_3} \gamma_{p_3})$, otherwise $\gamma_{p_3} = 0_2$.

9.33) Def $\alpha_{p_1} / \beta_{p_2} =_p \gamma_{p_3} \longleftrightarrow p_1 = p_2 = p_3$ & $((\beta_{p_2} \neq 0_{p_2}$ & $\exists r \exists s (r \in_{p_1} \alpha_{p_1}$ &
$s \in_{p_2} \beta_{p_2}$ & $r \hat{/} s \in_{p_3} \gamma_{p_3})) \vee (\beta_{p_2} = 0_{p_2}$ & $\gamma_{p_3} = 0_{p_3}))$, otherwise
$\gamma_{p_3} = 0_2$.

The field axioms follow immediately. After recording a preliminary theorem, we can now define the p-adic valuation on the p-adic numbers.

9.34) $r \sim_p s \longrightarrow$ Value $(p, r) \sim$ Value (p, s) .

Proof. By (9.12) if r and s are p-infinitesimal; otherwise, by (9.9), Value (p,r) = Value (p,s) . ||

9.35) Def Value $(\alpha_p) =_{\pi} \xi \longleftrightarrow \exists r\ (r \in_p \alpha_p$ & Value $(p,r) \in \xi)$, otherwise $\xi = \tilde{0}$.

The properties corresponding to (9.5)-(9.9) are easily verified.

Analogous to (5.32) is the following result, which says that if p is a *finite* prime, then every p-adic number contains a U_ν-fraction.

9.36) p is a prime & $\varepsilon_\nu(p)$ & r is p-limited $\longrightarrow$ $\exists s$ (s is a U_ν-fraction & $s \sim_p r$) .

Proof. If r is p-infinitesimal, let s be $\hat{0}$. Otherwise, here is the idea. Write $r = \pm p^j \cdot a/b$, where p^j is limited and noninfinitesimal, $a > 0$, $b > 0$, $p \nmid a$, and $p \nmid b$. (Here j may be positive or negative. Note that p^j is limited and noninfinitesimal $\longleftrightarrow \varepsilon_{\nu+1}(|j|)$.) There exists an unlimited power of p , say p^m , such that $p^{m+|j|} \in U_\nu$. Let $p^m \cdot k$ be the greatest multiple of p^m not exceeding a , and let $p^m \cdot \ell$ be the greatest multiple of p^m not exceeding b . Let s be $\pm p^j \cdot (a-p^m \cdot k)/(b-p^m \cdot \ell)$. Then s is a U_ν-fraction since both $a-p^m \cdot k$ and $b-p^m \cdot \ell$ are less than p^m . To show that $s \sim_p r$, note that

$$s-r = \pm\left(p^j \frac{a-p^m \cdot k}{b-p^m \cdot \ell} - p^j \frac{a}{b}\right) = \pm p^{m+j} \frac{a \cdot \ell - b \cdot k}{b \cdot (b-p^m \cdot \ell)} \quad .$$

The exponent on p here is at least $m+j$ (maybe more); since p^m is unlimited and $p^{|j|}$ is limited, it follows that p^{m+j} is unlimited and that $s-r$ is p-infinitesimal. ||

Infinite primes do exist, as a Euclid-style argument shows (9.37); if p is infinite, the p-adic numbers turn out to be nothing more than the integers mod p (9.38).

9.37) $\exists p$ (p is a prime & $\neg\varepsilon_\nu(p)$) .

Proof. If every prime satisfied ε_ν , then the primes would form a sequence u , and the number $(\Pi u)(\text{Ln } u)+1$ could have no prime factors, contrary to the fundamental theorem of arithmetic (itself an easy consequence of (BLNP)). ||

9.38) p is a prime & $\neg\varepsilon_\nu(p) \longrightarrow \forall\alpha_p \exists! i(0 \leq i \leq p-1 \;\&\; \hat{i} \in_p \alpha_p)$.

Proof. Clearly the p-adic numbers $p(p,\hat{0})$, $p(p,\hat{1}),\ldots,$ $p(p,(p-1)\hat{\ })$ are all distinct; it remains to show that every fraction a/b that is p-limited and not p-infinitesimal is p-infinitely close to one of $p(p,\hat{1}),\ldots,p(p,(p-1)\hat{\ })$. Since $\neg\varepsilon_\nu(p^1)$, neither a nor b is divisible by p . Let k be the unique number among $1,\ldots,p-1$ such that $b\cdot k \equiv a \pmod p$. Then $(a/b)-k = (a-b\cdot k)/b$ is p-infinitesimal; that is, a/b is p-infinitely close to $\hat{k}$. ||

In §8 we discussed decimal expansions of real numbers. Analogously, p-adic numbers have p-adic expansions. We list the definitions and theorems, omitting the (easy) proofs.

9.39) Def u p-converges to $\alpha_q \longleftrightarrow p=q$ & u is a sequence of fractions & $\neg\,\varepsilon_{\nu+1}(\text{Ln } u)$ & $\forall i\ (i \leq \text{Ln}\, u\ \&\ \neg\varepsilon_{\nu+1}(i) \longrightarrow u(i)\ \epsilon_q\ \alpha_q)$.

9.40) Def u is a p-adic expansion of $\alpha_q \longleftrightarrow p=q$ & u is a sequence & $\forall i\ (1 \leq i \leq \text{Ln } u \longrightarrow \exists k\ (k < p\ \&\ u(i) = \hat{k}))$ & Polyvalue $(u,\hat{p})\ \epsilon_q\ \alpha_q$.

9.41) u is a p-adic expansion of α_p & $\neg\,\varepsilon_{\nu+1}(\text{Ln } u) \longrightarrow$ Polyvalseq $(u,\hat{p})$ p-converges to α_p . ||

9.42) $\varepsilon_\nu(p)$ & u and v are p-adic expansions of α_p & $\neg\,\varepsilon_{\nu+1}(\text{Ln } u)$ & $\neg\,\varepsilon_{\nu+1}(\text{Ln } v)$ & $i \geq 1$ & $\varepsilon_{\nu+1}(i) \longrightarrow u(i) = v(i)$. ||

9.43) Def u is a p-adic approximating sequence for $r \longleftrightarrow$ r is a fraction & Value $(p,r)\ \hat{\leq}\ \hat{1}$ & u is a sequence of fractions & $\forall i\ (1 \leq i \leq \text{Ln } u \longrightarrow \exists k\ (k < p\ \&\ u(i) = \hat{k})$ &

Value $(p, r\,\hat{-}\,(\text{Polyvalseq } (u,\hat{p}))(i))\ \hat{\leq}\ (\text{Powerseq } (\hat{1}\hat{/}\hat{p},u))(i+1))$.

9.44) r is a fraction & Value $(p,r)\ \hat{\leq}\ \hat{1} \longrightarrow \exists u$ (u is a p-adic approximating sequence for r & Ln u = Log n) . ||

9.45) u and v are p-adic approximating sequences for r & Ln u = Ln v $\longrightarrow$ u = v . ||

9.46) Def p-Approxseq $(r,n) = u \longleftrightarrow$ u is a p-adic approximating sequence for r & Ln u = Log n , otherwise u = 1 .

9.47) Value $(\alpha_p) \leq \tilde{1}$ & $r\ \epsilon_p\ \alpha_p$ & $\neg\,\varepsilon_\nu(n) \longrightarrow$ p-Approxseq (r,n) is a p-adic expansion of α_p . ||

More sorts; Hensel's lemma

As was the case with the real numbers, we can introduce additional sorts for sets, functions and sequences. In probably the most convenient formulation, each set of p-adic numbers x_{sp} has a prime Prime (x_{sp}) associated with it; only p-adic numbers α_p with Prime (α_p) = Prime (x_{sp}) can be elements of x_{sp}. We can even introduce sorts for mixed concepts such as functions from the real numbers to the p-adic numbers or vice versa. One example of a set of p-adic numbers that can be formed is the set of all α_p such that Value $(\alpha_p) \in x_s$, where x_s is a given set of real numbers. Every p-adic set is "p-closed", and every p-adic function "p-continuous". The skeptical reader should have no difficulty providing his own details.

Let us write Sp for the sort "sequences of p-adic numbers". Then each sequence v_{Sp} is determined by a prime p and a sequence u of fractions; we write $v_{Sp} =_{Sp} Sp(p,u)$ (cf. (9.29)). We can define a function symbol Polyvalue of type $(Sp,p;p)$. Just as in the real case, Polyvalue (v_{Sp},α_p) means what one expects it to mean provided $\varepsilon_{\nu+1}$ (Ln v_{Sp}) -- and provided, of course, Prime (v_{Sp}) = Prime (α_p). If v_{Sp} is the polynomial $(\alpha_0)_p+(\alpha_1)_p x+\ldots+(\alpha_k)_p x^k$, then Derivpoly v_{Sp} is the polynomial $(\alpha_1)_p+p(p,\hat{2})\cdot(\alpha_2)_p\cdot x+\ldots+p(p,\hat{k})\cdot(\alpha_k)_p\cdot x^{k-1}$. We conclude this section by giving a formulation and sketching a proof of a standard tool in p-adic analysis, namely Hensel's lemma.

9.48) Prime $(u_{Sp}) = p$ & $\varepsilon_{\nu+1}(\text{Ln } u_{Sp})$ & $\forall i(1 \leq i \leq \text{Ln } u_{Sp} \longrightarrow$
Value $(u_{Sp}(i)) \leq \tilde{1}$) & Value $(\alpha_p) \leq \tilde{1}$ &
Value(Polyvalue $(u_{Sp},\alpha_p)) < \tilde{1}$ &
Value(Polyvalue(Derivpoly $u_{Sp},\alpha_p)) = \tilde{1} \longrightarrow$
$\exists\beta_p$(Value $(\beta_p) \leq \tilde{1}$ & Value $(\beta_p-\alpha_p) < \tilde{1}$ & Polyvalue $(u_{Sp},\beta_p) = 0_p$).

Proof. If $\neg\varepsilon_\nu(p)$, then Value(Polyvalue (u_{Sp},α_p)) , being a power of p less than $\tilde{1}$, must be $\tilde{0}$, whence Polyvalue $(u_{Sp},\alpha_p) = 0_p$, so that $\beta_p = \alpha_p$ satisfies the requirements. Assume therefore that $\varepsilon_\nu(p)$.

There is a sequence $f(x) = c_0+c_1x\ldots+c_kx^k$ of p-limited fractions such that $Sp(p,f) = u_{Sp}$. Then the sequence $f'(x) = c_1+2c_2x+\ldots+kc_kx^{k-1}$ satisfies $Sp(p,f') =$ Derivpoly u_{Sp} . There is also a p-limited fraction a such that $p(p,a) = \alpha_p$. By hypothesis, Value $(p,c_i) \mathrel{\hat{\leq}} \hat{1}$ for $i = 0,\ldots,k$; Value $(p,a) \mathrel{\hat{\leq}} \hat{1}$; Value (p,Polyvalue (f,a)) $\mathrel{\hat{<}} \hat{1}$; and Value (p,Polyvalue (f',a)) $= \hat{1}$. That is, the exponent on p in each c_i and in a is nonnegative, in $f(a)$ strictly positive, and in $f'(a)$ zero.

Assume for the moment that for each n there is a sequence $b_0,\ldots,b_{\text{Log } n}$ such that for $i = 0,\ldots,\text{Log } n$ we have $0 \leq b_i \leq p-1$ and Value $(p,f(b_0+b_1p+\ldots+b_ip^i)) \leq p^{-i-1}$ and such that Value $(p,b_0-a) < 1$. (Observe that the assertion following "for each n" is bounded.) Then let $\neg\varepsilon_\nu(n)$, and let β_p be the p-adic number given by the expansion $b_0+b_1p+\ldots+b_{\text{Log } n}p^{\text{Log } n}$. Then the first two of the three desired conclusions about β_p are clear, and **Polyvalue** $(u_{\mathcal{S}p}, \beta_p) = 0_p$ because $p^{-\text{Log } n}$ is infinitesimal (so that **Value**(**Polyvalue** $(u_{\mathcal{S}p}, \beta_p)) = \tilde{0})$.

It remains to construct the sequence $b_0, \ldots, b_{\text{Log } n}$. This is done, as usual, by "Newton's method". First let b_0 be the unique number among $0, \ldots, p-1$ such that p divides the numerator of $b_0 - a$; we write $b_0 \equiv a \pmod p$. Then $f(b_0) \equiv f(a) \equiv 0 \pmod p$, so that b_0 has the necessary properties. Now proceed by bounded induction, assuming that for $i < \text{Log } n$ we have found $b_0, \ldots, b_i$ such that Value $(p, f(b_0 + \ldots + b_i p^i)) \leq p^{-i-1}$ -- that is, such that $f(b_0 + \ldots + b_i p^i) \equiv 0 \pmod{p^{i+1}}$. Let $r = f(b_0 + \ldots + b_i p^i)/p^{i+1}$, so that the exponent on p in r is nonnegative. Then let b_{i+1} be the unique number among $0, \ldots, p-1$ satisfying $f'(a) \cdot b_{i+1} \equiv -r \pmod p$ (this uses the hypothesis $f'(a) \not\equiv 0 \pmod p$). Since $b_0 + \ldots + b_i p^i \equiv b_0 \equiv a \pmod p$, it follows that $f'(b_0 + \ldots + b_i p^i) \cdot b_{i+1} \equiv f'(a) \cdot b_{i+1} \equiv -r \pmod p$, and therefore that $f'(b_0 + \ldots + b_i p^i) \cdot b_{i+1} \cdot p^{i+1} \equiv -r \cdot p^{i+1} \pmod{p^{i+2}}$. But then mod p^{i+2} we have

$$\begin{aligned} f(b_0 + \ldots + b_i p^i + b_{i+1} p^{i+1}) &\equiv f(b_0 + \ldots + b_i p^i) + f'(b_0 + \ldots + b_i p^i) \cdot b_{i+1} \cdot p^{i+1} \\ &\equiv f(b_0 + \ldots + b_i p^i) - r \cdot p^{i+1} \\ &= 0 , \end{aligned}$$

as desired. $\|$

PART THREE

PREDICATIVE SET THEORY

§10. *Collections*

It is desirable to be able to refer to certain collections of objects in addition to those collections that form "sets" in the strict sense of Nelson's theory Q^0 : the collection of all x such that $\varepsilon_2(x)$, the collection of all limited U_ν-fractions, and so on. This objective is accomplished in this section. By analogy with the well-known "arithmetical hierarchy", we define, for $\lambda = 1,2,\ldots,$ new membership relations $\epsilon(\sum_\lambda)$, $\epsilon(\Pi_\lambda)$, and $\epsilon(\Delta_\lambda)$, and the notion of a "Δ_λ-collection". We then list several properties and examples of Δ_λ-collections, as well as a few unsolved problems.

Preliminaries: extend the definitions of ordered pair, ordered triple (4.1), and cartesian product as follows:

10.1) Def $\langle x_1,x_2,x_3\rangle = \langle x_1,\langle x_2,x_3\rangle\rangle$,

Def $\langle x_1,x_2,x_3,x_4\rangle = \langle x_1,\langle x_2,x_3,x_4\rangle\rangle$

$\vdots$

Def $\langle x_1,x_2,\ldots,x_\lambda\rangle = \langle x_1,\langle x_2,\ldots,x_\lambda\rangle\rangle$

$\vdots$

10.2) Def $a^2 = a\times a$,

Def $a^3 = a\times a^2$,

$\vdots$

Def $a^\lambda = a\times a^{\lambda-1}$,

$\vdots$

Context will prevent (10.2) from conflicting with the notation for exponentiation.

New membership relations

We first define, for $\kappa = 1,2,\ldots$ and $\lambda = 1,2,\ldots,$ binary relations $x \in y\,(\sum_\lambda^\kappa)$ and $x \in y\,(\Pi_\lambda^\kappa)$. We use the following abbreviations: $\exists^\kappa x A$ means $\exists x(\varepsilon_\kappa(x) \& A)$, and $\forall^\kappa x A$ means $\forall x(\varepsilon_\kappa(x) \longrightarrow A)$; Q_λ is $\exists$ if λ is odd and $\forall$ if λ is even, and vice versa for $\overline{Q}_\lambda$.

10.3) Def $x \in y(\sum_\lambda^\kappa) \longleftrightarrow \exists a(a$ is a set $\& \forall w(w \in a \longrightarrow \varepsilon_{\kappa-1}(w))\ \&$
$y \subseteq a^{\lambda+1})\ \&\ \exists^\kappa x_1 \forall^\kappa x_2 \exists^\kappa x_3 \ldots Q_\lambda^\kappa x_\lambda <x_1, x_2, x_3, \ldots, x_\lambda, x> \in y$.

10.4) Def $x \in y(\Pi_\lambda^\kappa) \longleftrightarrow \exists a(a$ is a set $\& \forall w(w \in a \longrightarrow \varepsilon_{\kappa-1}(w))\ \&$
$y \subseteq a^{\lambda+1})\ \&\ \forall^\kappa x_1 \exists^\kappa x_2 \forall^\kappa x_3 \ldots \overline{Q}_\lambda^\kappa x_\lambda <x_1, x_2, x_3, \ldots, x_\lambda, x> \in y$.

The set a that appears in the first part of these definitions could well be U_κ (see (5.2)), but on occasion it will need to be something larger. The interesting part of the definitions is the second part. For instance, let y_1 be $\{z \in U_1^2 : \mathrm{Proj}_1\ z = \mathrm{Proj}_2\ z\}$. Then

$$x \in y_1(\textstyle\sum_1^1) \longleftrightarrow \exists^1 x_1 <x_1, x> \in y_1$$

$$\longleftrightarrow \exists^1 x_1 (x_1 \in U_1 \ \&\ x \in U_1 \ \&\ x_1 = x)$$

$$\longleftrightarrow \exists^1 x_1 (x_1 = x)$$

$$(\text{because } \varepsilon_1(x_1) \longrightarrow x_1 \in U_1)$$

$$\longleftrightarrow \varepsilon_1(x) .$$

Likewise, if y_2 is $\{z \in U_1^2\colon \text{Log Proj}_1\ z = \text{Proj}_2\ z\}$, then $x \in y_2(\sum_1^1) \longleftrightarrow \varepsilon_2(x)$. The point is that by changing the notion of membership (using $\in(\sum_1^1)$ instead of $\in$), we have made sets (y_1 and y_2) represent collections that are *not* sets (the x such that $\varepsilon_1(x)$ and $\varepsilon_2(x)$). The situation can be summarized by saying that ε_1 and ε_2 are "$\sum_1^1$-properties"; in these terms, the following proposition asserts that for $\kappa = 1,2,\ldots,$ ε_κ is a $\sum_1^\kappa$-property.

10.5) $\exists y \forall x(x \in y(\sum_1^\kappa) \longleftrightarrow \varepsilon_\kappa(x))$.

Proof. Let y be $\{z \in U_\kappa^2\colon \text{Proj}_1\ z = \text{Proj}_2\ z\}$ and argue as above. ||

First among the facts that we shall record about the relations $\in(\sum_\lambda^\kappa)$ and $\in(\Pi_\lambda^\kappa)$ is a useful theorem scheme. Let $\underline{f}$ be a bounded unary function symbol in an extension by definitions of $\tilde{\Omega}^\mu$; the assertion is that if a formula $\mathbf{A}[x]$ defines a $\sum_\lambda^\kappa$-property, then so does $\mathbf{A}[\underline{f}x]$. More precisely:

10.6) a is a set & $\forall w(w \in a \longrightarrow \varepsilon_{\kappa-1}(w)) \longrightarrow$
$\exists z \forall x(x \in z(\sum_\lambda^\kappa) \longleftrightarrow (x \in a \ \&\ \underline{f}x \in y(\sum_\lambda^\kappa)))$.

Proof. Let z be $\{\langle x_1,\ldots,x_\lambda,x\rangle\colon x \in a \ \&\ \langle x_1,\ldots,x_\lambda,\underline{f}x\rangle \in y\}$. ||

Similarly:

10.7) a is a set & $\forall w(w \in a \longrightarrow \varepsilon_{\kappa-1}(w)) \longrightarrow$
$\exists z \forall x(x \in z(\Pi_\lambda^\kappa) \longleftrightarrow (x \in a \ \&\ \underline{f}x \in y(\Pi_\lambda^\kappa)))$. ||

Let y be the set of all ordered pairs $\langle k,x \rangle$ such that $k \in U_\nu$, x is a U_ν-fraction, and $|x| \mathbin{\hat{>}} \hat{k}$. One checks easily that $x \in y(\Pi_1^\nu) \longleftrightarrow x$ is an unlimited U_ν-fraction. It then follows from (10.7) (let $\underline{f}$ be the function symbol Recip) that being an infinitesimal U_ν-fraction is also a Π_1^ν-property.

The indices κ and λ that appear in (10.3) and (10.4) are genetic numbers, not formal terms of the theory. Indeed, (10.3) and (10.4) are really whole families of definitions, one for each choice of κ and λ, and similar remarks apply to many of the theorems that follow (including (10.5)-(10.7)). (Strictly speaking, some of the theorems below are valid only for $1 \leq \kappa \leq \mu$, where we are working in the theory $\tilde{\underline{Q}}^\mu$, the reason being that the proofs require ε_κ to respect multiplication. Of course, if we want a larger κ, nothing prevents us from moving to a stronger $\tilde{\underline{Q}}^\mu$; see the remarks at the end of §2.) We shall see at the end of this section that for many purposes the superscript κ may be replaced by a formal variable k. It will then be possible to quantify over k and thereby essentially to remove that index from the notation entirely. Certainly the more important of the two indices is the subscript λ; there seems to be no way to handle all λ at once.

Roughly speaking, κ indicates the "height" of a collection and λ its "complexity". That is, κ describes the size (level of exponentiability) of the collection's largest members, and λ gives the number of quantifiers in its defining formula. Let us now make these ideas more precise by investigating the dependence of the relations $\epsilon(\Sigma_\lambda^\kappa)$ and $\epsilon(\Pi_\lambda^\kappa)$ on κ and λ.

Properties of the relations

We examine λ first. Here the results correspond closely to standard theorems concerning the arithmetical hierarchy. For instance, every $\Sigma^{\kappa}_{\lambda}$-property is a $\Sigma^{\kappa}_{\lambda+1}$-property, and every Π^{κ}_{λ}-property is a $\Pi^{\kappa}_{\lambda+1}$-property:

10.8) $\exists z \forall x(x \in z(\Sigma^{\kappa}_{\lambda+1}) \longleftrightarrow x \in y(\Sigma^{\kappa}_{\lambda}))$.

Proof. Let z be the set

$\{\langle x_1,\ldots,x_\lambda,x_{\lambda+1},x\rangle : \langle x_1,\ldots,x_\lambda,x\rangle \in y \;\&\; x_{\lambda+1} \in U_\kappa\}$. ||

10.9) $\exists z \forall x(x \in z(\Pi^{\kappa}_{\lambda+1}) \longleftrightarrow x \in y(\Pi^{\kappa}_{\lambda+1}))$. ||

To save space, let us use the symbol $*$ for the following "dualization" operation: $\mathbf{A}^*$ is the formula obtained from the formula $\mathbf{A}$ by replacing $\Sigma^{\kappa}_{\lambda}$ with Π^{κ}_{λ} , Π^{κ}_{λ} with $\Sigma^{\kappa}_{\lambda}$, $\exists^{\kappa}$ with $\forall^{\kappa}$, and $\forall^{\kappa}$ with $\exists^{\kappa}$, while leaving all other symbols (including other occurrences of $\exists$ and $\forall$) unchanged. Hence $(10.3)^*$ is (10.4), $(10.6)^*$ is (10.7), and $(10.8)^*$ is (10.9). The reader may check that the proofs of (10.10)-(10.19) can be dualized, so that $(10.10)^*$-$(10.19)^*$ are theorems of $\tilde{Q}^{\mu}$. For instance, (10.10) asserts that every $\Sigma^{\kappa}_{\lambda}$-property is a $\Pi^{\kappa}_{\lambda+1}$-property, and $(10.10)^*$ asserts that every Π^{κ}_{λ}-property is a $\Sigma^{\kappa}_{\lambda+1}$-property.

10.10) $\exists z \forall x(x \in z(\Pi^{\kappa}_{\lambda+1}) \longleftrightarrow x \in y(\Sigma^{\kappa}_{\lambda}))$.

Proof. Let z be $U_\kappa \times y$. Then

$$x \in z(\Pi^\kappa_{\lambda+1}) \longleftrightarrow \forall^\kappa x_1 \exists^\kappa x_2 \ldots \overline{Q}_{\lambda+1}{}^\kappa x_{\lambda+1} \langle x_1, x_2, \ldots, x_{\lambda+1}, x\rangle \in z$$

$$\longleftrightarrow \exists^\kappa x_2 \ldots \overline{Q}_{\lambda+1}{}^\kappa x_{\lambda+1} \langle x_2, \ldots, x_{\lambda+1}, x\rangle \in y$$

$$\longleftrightarrow x \in y(\textstyle\sum^\kappa_\lambda) \ . \ \|$$

10.11) $\exists z \forall x(x \in z(\Pi^\kappa_{\lambda+1}) \longleftrightarrow \forall^\kappa x' \langle x', x\rangle \in y(\sum^\kappa_\lambda))$.

Proof. Let z be $\{\langle x', x_1, \ldots, x_\lambda, x\rangle : \langle x_1, \ldots, x_\lambda, \langle x', x\rangle\rangle \in y\}$. ‖

10.12) $\exists z \forall x(x \in z(\sum^\kappa_\lambda) \longleftrightarrow \exists^\kappa x' \langle x', x\rangle \in y(\sum^\kappa_\lambda))$.

Proof. Since y is a subset of $a^{\lambda+1}$ for some set a all of whose elements satisfy $\varepsilon_{\kappa-1}$, the set $z = \{\langle\langle x', x_1\rangle, x_2, \ldots, x_\lambda, x\rangle : \langle x_1, \ldots, x_\lambda, \langle x', x\rangle\rangle \in y\}$ satisfies this requirement as well; this is because $\varepsilon_{\kappa-1}$ respects ordered pairs. Then

$$x \in z(\textstyle\sum^\kappa_\lambda) \longleftrightarrow \exists^\kappa w \forall^\kappa x_2 \ldots Q^\kappa_\lambda x_\lambda \langle w, x_2, \ldots, x_\lambda, x\rangle \in z$$

$$\longleftrightarrow \exists^\kappa x' \exists^\kappa x_1 \forall^\kappa x_2 \ldots Q^\kappa_\lambda x_\lambda \langle\langle x', x_1\rangle, x_2, \ldots, x_\lambda, x\rangle \in z$$

($\longrightarrow$ by the definition of z , $\longleftarrow$ because ε_κ respects ordered pairs)

$$\longleftrightarrow \exists^\kappa x' \exists^\kappa x_1 \forall^\kappa x_2 \ldots Q^\kappa_\lambda x_\lambda \langle x_1, \ldots, x_\lambda, \langle x', x\rangle\rangle \in y$$

$$\longleftrightarrow \exists^\kappa x' \langle x', x\rangle \in y(\textstyle\sum^\kappa_\lambda) \ . \ \|$$

Formulas (10.12) and (10.12)* are of course "contraction of quantifiers." Now we prove the expected theorems about complements, intersections, and unions.

10.13) a is a set & $\forall w(w \in a \longrightarrow \varepsilon_{\kappa-1}(w)) \longrightarrow$

$$\exists z \forall x(x \in z(\Pi^{\kappa}_{\lambda}) \longleftrightarrow (x \in a \ \&\ x \notin y(\Sigma^{\kappa}_{\lambda}))) .$$

Proof. Let $y \subseteq b^{\lambda+1}$; we may assume that $U_{\kappa} \subseteq b$. Let z be $\{\langle x_1,\ldots,x_{\lambda},x\rangle : \langle x_1,\ldots,x_{\lambda}\rangle \in b^{\lambda}\ \&\ x \in a\ \&\ \langle x_1,\ldots,x_{\lambda},x\rangle \notin y\}$. Then

$$x \in z(\Pi^{\kappa}_{\lambda}) \longleftrightarrow \forall^{\kappa} x_1 \ldots \overline{Q}^{\kappa}_{\lambda} x_{\lambda} \langle x_1,\ldots,x_{\lambda},x\rangle \in z$$

$$\longleftrightarrow x \in a\ \&\ \forall^{\kappa} x_1 \ldots \overline{Q}^{\kappa}_{\lambda} x_{\lambda} \neg \langle x_1,\ldots,x_{\lambda},x\rangle \in y$$

$$\longleftrightarrow x \in a\ \& \neg \exists^{\kappa} x_1 \ldots Q^{\kappa}_{\lambda} x_{\lambda} \langle x_1,\ldots,x_{\lambda},x\rangle \in y$$

$$\longleftrightarrow x \in a\ \&\ x \notin y(\Sigma^{\kappa}_{\lambda}) \ . \|$$

We noted earlier that "unlimited U_{ν}-fraction" is a Π^{ν}_1-property; it now follows from (10.13)* that "limited U_{ν}-fraction" is a Σ^{ν}_1-property. Of course, it is easy to see this directly.

10.14) $\exists z \forall x(x \in z(\Sigma^{\kappa}_{\lambda}) \longleftrightarrow (x \in y_1(\Sigma^{\kappa}_{\lambda})\ \&\ x \in y_2(\Sigma^{\kappa}_{\lambda})))$.

Proof. First let z_1 be

$\{\langle w_{\lambda},x_{\lambda},w_{\lambda-1},x_{\lambda-1},\ldots,w_1,x_1,x\rangle : \langle x_1,\ldots,x_{\lambda},x\rangle \in y_1\ \&\ \langle w_1,\ldots,w_{\lambda},x\rangle \in y_2\}$.

By the prenex operations,

$$x \in y_1(\Sigma^{\kappa}_{\lambda})\ \&\ x \in y_2(\Sigma^{\kappa}_{\lambda})$$

$$\longleftrightarrow \exists^{\kappa} x_1 \exists^{\kappa} w_1 \forall^{\kappa} x_2 \forall^{\kappa} w_2 \ldots Q^{\kappa}_{\lambda} x_{\lambda} Q^{\kappa}_{\lambda} w_{\lambda} \langle w_{\lambda},x_{\lambda},\ldots,w_1,x_1,x\rangle \in z_1 \ .$$

Now use (10.11) and (10.12) and their duals. For instance, if λ is 3 , then

$\exists^\kappa x_1 \exists^\kappa w_1 \forall^\kappa x_2 \forall^\kappa w_2 \exists^\kappa x_3 \exists^\kappa w_3 \langle w_3, x_3, w_2, x_2, w_1, x_1, x\rangle \in z_1$

$\longleftrightarrow \exists^\kappa x_1 \exists^\kappa w_1 \forall^\kappa x_2 \forall^\kappa w_2 \exists^\kappa x_3 \langle x_3, w_2, x_2, w_1, x_1, x\rangle \in z_1(\Sigma^\kappa_1)$

$\longleftrightarrow \exists^\kappa x_1 \exists^\kappa w_1 \forall^\kappa x_2 \forall^\kappa w_2 \langle w_2, x_2, w_1, x_1, x\rangle \in z_2(\Sigma^\kappa_1)$

(for some z_2 , by (10.12))

$\longleftrightarrow \exists^\kappa x_1 \exists^\kappa w_1 \forall^\kappa x_2 \langle x_2, w_1, x_1, x\rangle \in z_3(\Pi^\kappa_2)$

(for some z_3 , by (10.11))

$\longleftrightarrow \exists^\kappa x_1 \exists^\kappa w_1 \langle w_1, x_1, x\rangle \in z_4(\Pi^\kappa_2)$

(for some z_4 , by $(10.12)^*$)

$\longleftrightarrow \exists^\kappa x_1 \langle x_1, x\rangle \in z_5(\Sigma^\kappa_3)$ (for some z_5 , by $(10.11)^*$)

$\longleftrightarrow x \in z(\Sigma^\kappa_3)$ (for some z , by (10.12)) . ||

10.15) $\exists z \forall x(x \in z(\Sigma^\kappa_\lambda) \longleftrightarrow (x \in y_1(\Sigma^\kappa_\lambda) \vee x \in y_2(\Sigma^\kappa_\lambda)))$.

Proof. Like (10.14); alternatively, use (10.13), $(10.14)^*$, and $(10.13)^*$. ||

The next proposition says that the cartesian product of two Σ^κ_λ-properties is a Σ^κ_λ-property.

10.16) $\exists z \forall x(x \in z(\Sigma^\kappa_\lambda) \longleftrightarrow \exists w_1 \exists w_2(x=\langle w_1, w_2\rangle \,\&\, w_1 \in y_1(\Sigma^\kappa_\lambda) \,\&\, w_2 \in y_2(\Sigma^\kappa_\lambda)))$.

Proof. We may assume that both y_1 and y_2 are subsets of $a^{\lambda+1}$ and that $U_\kappa \subseteq a$. First we construct, as a $\sum_\lambda^\kappa$-property, the cartesian product of the $\sum_\lambda^\kappa$-property y_1 with the set a. To do this, let z_1 be
$\{<x_1,\ldots,x_\lambda,<x',x''>> : <x_1,\ldots,x_\lambda,x'> \in y_1 \ \& \ x'' \in a\}$; one checks easily that $x \in z_1(\sum_\lambda^\kappa) \longleftrightarrow \exists w_1 \exists w_2(x=<w_1,w_2> \& \ w_1 \in y_1(\sum_\lambda^\kappa) \ \& \ w_2 \in a)$. Similarly, we can construct z_2 such that $x \in z_2(\sum_\lambda^\kappa) \longleftrightarrow \exists w_1 \exists w_2(x=<w_1,w_2> \ \& \ w_1 \in a \ \& \ w_2 \in y_2(\sum_\lambda^\kappa))$. The desired z is then the $\sum_\lambda^\kappa$-intersection (cf. (10.14)) of z_1 and z_2. ||

We now briefly discuss the role of the index κ. Briefly, the greatest generality is obtained at the first level, $\kappa = 1$. This is because every $\sum_\lambda^{\kappa+1}$-property is a $\sum_\lambda^\kappa$-property:

10.17) $\exists z \forall x(x \in z(\sum_\lambda^\kappa) \longleftrightarrow x \in y(\sum_\lambda^{\kappa+1}))$.

Proof. Let z be $\{<x_1,\ldots,x_\lambda,x> : <\mathrm{Log}\ x_1,\ldots,\mathrm{Log}\ x_\lambda,x> \in y\}$. Then

$$x \in z(\textstyle\sum_\lambda^\kappa) \longleftrightarrow \exists^\kappa x_1 \ldots Q_\lambda^\kappa x_\lambda <x_1,\ldots,x_\lambda,x> \in z$$

$$\longleftrightarrow \exists^\kappa x_1 \ldots Q_\lambda^\kappa x_\lambda <\mathrm{Log}\ x_1,\ldots,\mathrm{Log}\ x_\lambda,x> \in y$$

$$\longleftrightarrow \exists^{\kappa+1} w_1 \ldots Q_\lambda^{\kappa+1} w_\lambda <w_1,\ldots,w_\lambda,x> \in y$$

$$\longleftrightarrow x \in y(\textstyle\sum_\lambda^{\kappa+1}) \ . \quad ||$$

It can be seen using arguments of the sort used in proving (10.17) that the quantifiers in the defining formula of a $\sum_\lambda^\kappa$- or Π_λ^κ-property may be restricted to $\varepsilon_{\kappa+1}$, $\varepsilon_{\kappa+2}$, ... rather than always ε_κ. For instance, if $y \subseteq U_2^4$, then $\forall^2 x_1 \exists^5 x_2 \forall^3 x_3 \langle x_1, x_2, x_3, x\rangle \in y$ defines a Π_3^2-property; indeed, it is equivalent to

$\forall^2 x_1 \exists^2 x_2 \forall^2 x_3 \langle x_1, \text{Log Log Log } x_2, \text{Log } x_3, x\rangle \in y$.

According to our next proposition, $\sum_\lambda^\kappa$-properties are not *much* more general than $\sum_\lambda^{\kappa+1}$-properties; they can simply extend "higher". That is, if we have a $\sum_\lambda^\kappa$-property and intersect it with $U_{\kappa+1}$ or some other set all of whose elements satisfy ε_κ, we get a $\sum_\lambda^{\kappa+1}$-property.

10.18) a is a set & $\forall w(w \in a \longrightarrow \varepsilon_\kappa(w)) \longrightarrow$

$$\exists z \forall x(x \in z(\textstyle\sum_\lambda^{\kappa+1}) \longleftrightarrow (x \in a \ \&\ x \in y(\sum_\lambda^\kappa)))\ .$$

Proof. The set

$z = \{\langle \text{Log } x_1, \ldots, \text{Log } x_\lambda, x\rangle : x \in a \ \&\ \langle x_1, \ldots, x_\lambda, x\rangle \in y\}$ consists of $(\lambda+1)$-tuples of numbers satisfying ε_κ, and fulfills the desired property. ||

The reason we bother with the index κ at all rather than working only with the relations $\in(\sum_\lambda^1)$ and $\in(\Pi_\lambda^1)$ is that we may have occasion to jump to a larger "universe". For instance, if $\kappa \geq 2$, then all ordered pairs $\langle w_1, w_2\rangle$ such that $w_1 \in w_2(\sum_\lambda^\kappa)$, restricted to certain "universe" a, define not a $\sum_\lambda^\kappa$-property but a $\sum_\lambda^{\kappa-1}$-property:

10.19) a is a set & $\forall w(w \in a \longrightarrow \varepsilon_{\kappa-1}(w)) \longrightarrow$

$$\exists z \forall x(x \in z(\textstyle\sum_\lambda^{\kappa-1}) \longleftrightarrow \exists w_1 \exists w_2(x = \langle w_1, w_2\rangle \ \&\ w_2 \subseteq a^{\lambda+1}\ \&$$

$$w_1 \in w_2(\textstyle\sum_\lambda^\kappa)))\ .$$

Proof. Let z_1 be the set $\{\langle x_1,\ldots,x_\lambda, \langle w_1,w_2\rangle\rangle : w_2 \subseteq a^{\lambda+1} \;\&\; \langle x_1,\ldots,x_\lambda,w_1\rangle \in w_2\}$. This is a set of $(\lambda+1)$-tuples of numbers satisfying $\varepsilon_{\kappa-2}$ (not necessarily $\varepsilon_{\kappa-1}$) , and clearly

$$w_2 \subseteq a^{\lambda+1} \;\&\; w_1 \in w_2(\textstyle\sum_\lambda^\kappa) \longleftrightarrow \exists^\kappa x_1\ldots Q_\lambda^\kappa x_\lambda \langle x_1,\ldots,x_\lambda,\langle w_1,w_2\rangle\rangle \in z_1 .$$

By earlier remarks, the right-hand side of this last formula defines a $\sum_\lambda^{\kappa-1}$-property of the ordered pair $\langle w_1,w_2\rangle$. ||

The definition of a Δ_λ^κ-collection

As expected, we call a property Δ_λ^κ if it can be expressed both as a $\sum_\lambda^\kappa$-property and as a Π_λ^κ-property. To be precise:

10.20) Def y is a Δ_λ^κ-collection $\longleftrightarrow$ $\exists y_1 \exists y_2 \exists a(y=\langle y_1,y_2\rangle \;\&$
a is a set $\&\; \forall w(w \in a \longrightarrow \varepsilon_{\kappa-1}(w)) \;\&\; y_1 \subseteq a^{\lambda+1} \;\&\; y_2 \subseteq a^{\lambda+1} \;\&$
$\forall x(x \in y_1(\textstyle\sum_\lambda^\kappa) \longleftrightarrow x \in y_2(\Pi_\lambda^\kappa)))$.

10.21) Def $x \in y(\Delta_\lambda^\kappa) \longleftrightarrow y$ is a Δ_λ^κ-collection $\&\; x \in \mathrm{Proj}_1\, y(\textstyle\sum_\lambda^\kappa)$.

The notion of a Δ_λ^κ-collection is somewhat easier to work with than the relations $\in(\sum_\lambda^\kappa)$ and $\in(\Pi_\lambda^\kappa)$ by themselves. Of course, many simple properties of Δ_λ^κ-collections are immediate from (10.6)-(10.18) and $(10.10)^*$-$(10.18)^*$. For instance, let $\underline{f}$ be a bounded unary function symbol in an extension by definitions of $\tilde{Q}^\mu$. Then:

10.22) a is a set $\&\; \forall w(w \in a \longrightarrow \varepsilon_{\kappa-1}(w)) \longrightarrow$
$\exists z \forall x(x \in z(\Delta_\lambda^\kappa) \longleftrightarrow x \in a \;\&\; \underline{f}x \in y(\Delta_\lambda^\kappa))$.

Proof. We may assume that $y = \langle y_1, y_2 \rangle$ is a Δ^κ_λ-collection. By (10.6), there is a z_1 such that $x \in z_1(\Sigma^\kappa_\lambda) \longleftrightarrow (x \in a \;\&\; \underline{f}x \in y_1(\Sigma^\kappa_\lambda))$; by (10.7), there is a z_2 such that $x \in z_2(\Pi^\kappa_\lambda) \longleftrightarrow (x \in a \;\&\; \underline{f}x \in y_2(\Pi^\kappa_\lambda))$. Let z be $\langle z_1, z_2 \rangle$. ||

The proofs of the following are not much harder.

10.23) $\exists z \forall x(x \in z(\Delta^\kappa_{\lambda+1}) \longleftrightarrow x \in y(\Sigma^\kappa_\lambda))$. ||

10.24) $\exists z \forall x(x \in z(\Delta^\kappa_{\lambda+1}) \longleftrightarrow x \in y(\Pi^\kappa_\lambda))$. ||

10.25) $\exists z \forall x(x \in z(\Delta^\kappa_{\lambda+1}) \longleftrightarrow \exists^\kappa x' \langle x', x \rangle \in y(\Delta^\kappa_\lambda))$. ||

10.26) $\exists z \forall x(x \in z(\Delta^\kappa_{\lambda+1}) \longleftrightarrow \forall^\kappa x' \langle x', x \rangle \in y(\Delta^\kappa_\lambda))$. ||

10.27) a is a set & $\forall w(w \in a \longrightarrow \varepsilon_{\kappa-1}(w)) \longrightarrow$
$\exists z \forall x(x \in z(\Delta^\kappa_\lambda) \longleftrightarrow (x \in a \;\&\; x \notin y(\Delta^\kappa_\lambda)))$. ||

10.28) $\exists z \forall x(x \in z(\Delta^\kappa_\lambda) \longleftrightarrow (x \in y_1(\Delta^\kappa_\lambda) \;\&\; x \in y_2(\Delta^\kappa_\lambda)))$. ||

10.29) $\exists z \forall x(x \in z(\Delta^\kappa_\lambda) \longleftrightarrow (x \in y_1(\Delta^\kappa_\lambda) \vee x \in y_2(\Delta^\kappa_\lambda)))$. ||

10.30) $\exists z \forall x(x \in z(\Delta^\kappa_\lambda) \longleftrightarrow \exists w_1 \exists w_2(x = \langle w_1, w_2 \rangle \;\&\; w_1 \in y_1(\Delta^\kappa_\lambda) \;\&$
$w_2 \in y_2(\Delta^\kappa_\lambda)))$. ||

10.31) $\exists z \forall x(x \in z(\Delta^\kappa_\lambda) \longleftrightarrow x \in y(\Delta^{\kappa+1}_\lambda))$. ||

10.32) a is a set & $\forall w(w \in a \longrightarrow \varepsilon_\kappa(w)) \longrightarrow$
$\exists z \forall x(x \in z(\Delta^{\kappa+1}_\lambda) \longleftrightarrow (x \in a \;\&\; x \in y(\Delta^\kappa_\lambda)))$. ||

The $\Delta^{\kappa}_{\lambda}$-hierarchy

What kinds of properties can be formalized as $\Delta^{\kappa}_{\lambda}$-collections? First of all, it is fairly clear that every *set* all of whose elements satisfy $\varepsilon_{\kappa-1}$ forms a Δ^{κ}_{1}-collection:

10.33) a is a set & $\forall w(w \in a \longrightarrow \varepsilon_{\kappa-1}(w)) \longrightarrow$
$\exists z \forall x(x \in z(\Delta^{\kappa}_{1}) \longleftrightarrow x \in a)$.

Proof. Let y be $U_{\kappa} \times a$; then clearly $x \in y(\Sigma^{\kappa}_{1}) \longleftrightarrow x \in a$ and $x \in y(\Pi^{\kappa}_{1}) \longleftrightarrow x \in a$. Let z be $\langle y,y \rangle$. ||

A somewhat more sophisticated result is the converse:

10.34) $\exists b(b$ is a set & $\forall x(x \in b \longleftrightarrow x \in y(\Delta^{\kappa}_{1})))$.

Proof. Assume that $y = \langle y_1,y_2 \rangle$, with $y_1 \subseteq a^2$ and $y_2 \subseteq a^2$. Then $x \in y(\Delta^{\kappa}_{1}) \longleftrightarrow \exists^{\kappa} x_1 \langle x_1,x \rangle \in y_1$. But also $x \in y(\Delta^{\kappa}_{1}) \longleftrightarrow \forall^{\kappa} x_1 \langle x_1,x \rangle \in y_2$, so that $x \notin y(\Delta^{\kappa}_{1}) \longleftrightarrow \exists^{\kappa} x_1 \langle x_1,x \rangle \notin y_2$. For every x either $x \in y(\Delta^{\kappa}_{1})$ or $x \notin y(\Delta^{\kappa}_{1})$ holds, so in particular for every x in a there is an x_1 such that $\varepsilon_{\kappa}(x_1)$ and either $\langle x_1,x \rangle \in y_1$ or $\langle x_1,x \rangle \notin y_2$. By bounded replacement (see §1) there exists a function $f = \{\langle x,x_1 \rangle : x \in a \ \& \ \min_{x_1}(\langle x_1,x \rangle \in y_1 \vee \langle x_1,x \rangle \notin y_2)\}$ with the property that $\varepsilon_{\kappa}(f(x))$ for every x in a . Let b be $\{x \in a: \langle f(x),x \rangle \in y_1\}$. If $x \in b$, then $\langle f(x),x \rangle \in y_1$, so $\exists^{\kappa} x_1 \langle x_1,x \rangle \in y_1$ and therefore $x \in y(\Delta^{\kappa}_{1})$. If $x \notin b$, then either $x \notin a$, in which case certainly $x \notin y(\Delta^{\kappa}_{1})$, or else $x \in a$ and $\langle f(x),x \rangle \notin y_2$ -- that is, $\exists^{\kappa} x_1 \langle x_1,x \rangle \notin y_2$, whence $x \notin y(\Delta^{\kappa}_{1})$. ||

It follows from observations made earlier in this section that the x such that $\varepsilon_\kappa(x)$ form a Δ_2^κ-collection and that the limited U_ν-fractions, unlimited U_ν-fractions, and infinitesimal U_ν-fractions all form Δ_2^ν-collections. None of these collections forms a set, of course, so it follows from (10.34) that we have before us several examples of Δ_2^κ-collections that are *not* Δ_1^κ-collections.

It is a well-known theorem of recursion theory [2,chapter 7] that every step in the arithmetical hierarchy is nontrivial: for every $\lambda \geq 1$, there are $\sum_\lambda$ relations that are not Δ_λ and vice versa, and therefore there are $\Delta_{\lambda+1}$ relations that are not Δ_λ. It is not at all clear how to duplicate this result in the present situation. In the usual proof, one enumerates all $\sum_\lambda$ relations of one variable by a single $\sum_\lambda$ relation of two variables and then "diagonalizes"; the closest we can come to such an enumeration, however, seems to be (10.19), and the jump from κ to $\kappa-1$ precludes the possibility of diagonalizing.

Another open question is whether bounded quantifiers $\exists x'(x' \leq c \& \ldots)$ or $\forall x'(x' \leq c \longrightarrow \ldots)$ can affect the smallest value of λ for which a formula defines a $\sum_\lambda^\kappa$- or Π_λ^κ-property. Of course, the bounded separation principle implies that bounded quantifiers may be used to define sets, so such quantifiers occurring *inside* a string of unbounded quantifiers ($\exists^\kappa$ and $\forall^\kappa$) will have no effect. That the same is true of bounded quantifiers preceding exactly *one* unbounded quantifier is the content of the following proposition (and its dual, which follows from (10.13)). The proof appears not to generalize, however, and it is by no means certain that a formula of the form

$\forall x'(x' \leq c \longrightarrow \exists^{\kappa} x_1 \forall^{\kappa} x_2 \ldots Q^{\kappa}_{\lambda} x_{\lambda} \langle x', x_1, x_2, \ldots, x_{\lambda}, x\rangle \in y)$ always defines a $\sum^{\kappa}_{\lambda}$-property.

10.35) $\varepsilon_2(c) \longrightarrow \exists z \forall x(x \in z(\sum^{\kappa}_1) \longleftrightarrow \forall x'(x' \leq c \longrightarrow$
$\langle x', x\rangle \in y(\sum^{\kappa}_1)))$.

Proof. Write $\theta(x)$ for the formula $\forall x'(x' \leq c \longrightarrow \langle x', x\rangle \in y(\sum^{\kappa}_1))$; we wish to show that $\theta(x)$ is a $\sum^{\kappa}_1$-property. We may assume that $y \subseteq a^2$ & $\forall w(w \in a \longrightarrow \varepsilon_{\kappa-1}(w))$, and also that $\varepsilon_{\kappa-1}(c)$: otherwise no x will satisfy $\theta(x)$.

Certainly $\theta(x)$ is equivalent to $\forall x'(x' \leq c \longrightarrow \exists^{\kappa} x_1 \langle x_1, \langle x', x\rangle\rangle \in y)$. If this holds for a particular x , then there exists a function $f = \{\langle x', x_1\rangle : x' \leq c \ \& \ \min_{x_1} \langle x_1, \langle x', x\rangle\rangle \in y\}$. If $x' \leq c$, then $\varepsilon_{\kappa}(f(x'))$; since $\varepsilon_{\kappa-1}(c)$, it follows that $\varepsilon_{\kappa-2}(f)$. (If κ is 1 or 2 , then the hypothesis $\varepsilon_2(c)$ implies $\varepsilon_1(f)$. We leave to the reader the slight modifications of our argument for those two cases.) Let m be the largest value attained by f , so $\varepsilon_{\kappa}(m)$; let M be $2\wedge(2\wedge m)$, so $\varepsilon_{\kappa-2}(M)$. Let g be $f \cup \{\langle c+1, M\rangle\}$. Then g is a function whose domain is $\{0,1,\ldots,c+1\}$, and $\varepsilon_{\kappa-2}(g)$ holds; moreover, if $x' \leq c$, then $g(x') \leq \text{Log Log } g(c+1)$.

We have shown that $\theta(x)$ implies $\exists^{\kappa-2} g\langle g, x\rangle \in z_1$, where z_1 is the set $\{\langle g, x\rangle \in U_{\kappa-2}$: g is a function & Dom $g = \{0,1,\ldots,c+1\}$ & $\forall x'(x' \leq c \longrightarrow (g(x') \leq \text{Log Log } g(c+1) \ \& \ \langle g(x'), \langle x', x\rangle\rangle \in y))\}$. Conversely, suppose $\exists^{\kappa-2} g\langle g, x\rangle \in z_1$, and suppose $x' \leq c$. Let x_1 be $g(x')$. Then $x_1 \leq \text{Log Log } g(c+1)$ and $\langle x_1, \langle x', x\rangle\rangle \in y$. Since

$\varepsilon_{\kappa-2}(g)$ by assumption, it follows that $\varepsilon_{\kappa-2}(g(c+1))$ and therefore that $\varepsilon_{\kappa}(x_1)$. Thus $\exists^{\kappa} x_1 \langle x_1, \langle x', x\rangle\rangle \in y$, and thus $\theta(x)$. So $\theta(x)$ is a $\sum_1^{\kappa-2}$-property: it is equivalent to $x \in z_1(\sum_1^{\kappa-2})$. But the $\sum_1^{\kappa-2}$-elements of z_1 are all bounded above by $\sum_1^{\kappa}$-elements of y, which are elements of a; it follows by (10.18) that $\theta(x)$ is in fact a $\sum_1^{\kappa}$-property. ||

A refinement: Δ_λ-collections

We conclude this section by describing briefly how the genetic index κ may be replaced by a formal variable k. The idea is that the definitions (2.1) of unary predicate symbols $\varepsilon_1(x), \varepsilon_2(x), \ldots$ can all be encompassed in a single binary predicate symbol:

10.36) Def $\varepsilon_k(x) \longleftrightarrow \exists u(u$ is a sequence & $\mathrm{Ln}\ u = k+2$ & $\forall i(1 \leq i \leq k+1 \longrightarrow u(i+1) = \mathrm{Log}\ u(i))$ & $u(k+2) = x)$.

In the same vein (cf. (5.2)):

10.37) Def $U_k = a \longleftrightarrow \exists u(u$ is a sequence & $\mathrm{Ln}\ u = k$ & $u(1) = N$ & $\forall i(1 \leq i \leq k-1 \longrightarrow u(i+1) = \mathrm{Log}\ u(i))$ & $a = \mathrm{Setlog}\ u(k))$, otherwise $a = 1$.

(Aside from the case $k = 0$, the "otherwise" clause in (10.37) comes into play precisely if $\neg\varepsilon(k)$, for in that case there can be no sequence of length k. For the same reason, $\neg\varepsilon(k) \longrightarrow \neg\varepsilon_k(x)$.)

It should now be clear how to convert (10.3) and (10.4) into definitions of ternary predicate symbols $x \in y(\sum_\lambda^k)$ and $x \in y(\Pi_\lambda^k)$ (the variables being x, y, and k). Actually, it proves useful to be able to recover k from y; this can be accomplished by making y

an ordered pair whose first element is k . That is, we define

10.38) Def $x \in y(\sum_{\lambda}^{k}) \longleftrightarrow k \geq 1 \ \& \ \exists y' \exists a(y = \langle k,y' \rangle \ \& \ a$ is a set &
$\forall w(w \in a \longrightarrow \epsilon_{k-1}(w)) \ \& \ y' \subseteq a^{\lambda+1} \ \&$
$\exists^{k} x_1 \forall^{k} x_2 \ldots Q_{\lambda}^{k} x_{\lambda} \langle x_1, x_2, \ldots, x_{\lambda}, x \rangle \in y')$,

and dually for $\epsilon(\Pi_{\lambda}^{k})$. Then we can simplify the notation by defining

10.39) Def $x \in y(\sum_{\lambda}) \longleftrightarrow \exists k \exists y'(y = \langle k,y' \rangle \ \& \ x \in y(\sum_{\lambda}^{k}))$;

notions of Δ_{λ}^{k}-collection, $\epsilon(\Delta_{\lambda}^{k})$, Δ_{λ}-collection, and $\epsilon(\Delta_{\lambda})$ follow close behind. Theorems (10.17) and (10.18) can be generalized to

10.40) $0 < j < k \longrightarrow \exists z \forall x(x \in z(\sum_{\lambda}^{j}) \longleftrightarrow x \in y(\sum_{\lambda}^{k}))$

and

10.41) $0 < k < j \ \& \ a$ is a set & $\forall w(w \in a \longrightarrow \epsilon_{j-1}(w)) \longrightarrow$
$\exists z \forall x(x \in z(\sum_{\lambda}^{j}) \longleftrightarrow (x \in a \ \& \ x \in y(\sum_{\lambda}^{k})))$;

the earlier proofs carry over straightforwardly, making use of a bounded binary function symbol for the (Log m)-fold iterated logarithm of x .

It should be noted, however, that some of the other results of this section necessarily remain theorem *schemes*, even for fixed λ . For instance, (10.12) is certainly a theorem if in place of κ we write any of the formal terms 1,2,...; on the other hand, we cannot prove (10.12) with the unrestricted variable k in place of κ . The reason is that the proof requires ϵ_{κ} to respect multiplication --

something we cannot expect from the general ε_k . The reason for *that* follows from our discussion at the beginning of §5. If, for instance, $\varepsilon_k(x)$ were known to be inductive in x for all k such that $\varepsilon(k)$, then the formula $\forall k(\varepsilon(k) \longrightarrow \varepsilon_k(x))$ would be inductive in x *and would respect exponentiation*; as we noted in §5, this is impossible.

§11. *Infinite Cardinals*

Traditionally, the notion of cardinality has been approached through one-to-one correspondences. Such a method seems unsatisfactory in dealing with our Δ_λ-collections, for several reasons. Requiring our bijections to be functions (in the strict sense of $\tilde{\mathfrak{Q}}^\mu$) would certainly be too restrictive. On the other hand, more general "Δ_λ-mappings" are quite unmanageable: as noted in §10, it seems impossible to treat all λ at once, and any fixed λ would be arbitrary and unproductive inasmuch as simple operations on collections can make the functions involved much more complex.

We therefore adopt a different approach: we try to approximate Δ_λ-collections by *sets* (in the sense of $\tilde{\mathfrak{Q}}^\mu$), both from below and from above. In general, a collection will have subsets of all sufficiently small cardinalities and supersets of all sufficiently large cardinalities (subject, of course, to the restriction ε(Card a)). We declare that the cardinality of a Δ_λ-collection is determined by the cardinalities of its subsets and supersets.

Before we make these ideas precise, a comment about Δ_λ-collections is in order. As they were presented in §10, it may well happen that many different Δ_λ-collections have exactly the same Δ_λ-elements. For instance, in determining whether $x \in y(\Sigma_2^3)$ or $x \notin y(\Sigma_2^3)$, the presence or absence in y of ordered pairs $\langle x_1, x\rangle$ with $\neg\varepsilon_3(x_1)$ is completely irrelevant. This annoyance can be circumvented -- as problems involving equivalence relations always can -- by considering objects of new sorts: in this case, for each λ a sort for equivalence classes of

Δ_λ-collections modulo the relation of having the same elements. The reader who so desires may imagine that we are working with these sorts from the start; we prefer simply to disregard the problem and assume that a Δ_λ-collection is uniquely determined by its elements. If we know that a formula $\mathbb{A}[x]$ defines a Δ_2-property, we shall not hesitate to refer to "the Δ_2-collection $\{x: \mathbb{A}[x]\}$". It is hoped that the reader will forgive an increasingly informal style in other ways as well.

Cardinality relations

The following definitions actually depend on λ , but there is probably no harm in suppressing reference to λ .

11.1) Def a is an n-subset of $y \longleftrightarrow y$ is a Δ_λ-collection & a is a set & $\forall x(x \in a \longrightarrow x \in y(\Delta_\lambda))$ & Card a = n .

11.2) Def b is an n-superset of $y \longleftrightarrow y$ is a Δ_λ-collection & b is a set & $\forall x(x \in y(\Delta_\lambda) \longrightarrow x \in b)$ & Card b = n .

We define two size relations for Δ_λ-collections: "smaller according to subsets" ($\preceq$) and "smaller according to supersets" ($\overline{\prec}$) .

11.3) Def $y_1 \preceq y_2 \longleftrightarrow y_1$ and y_2 are Δ_λ-collections & $\forall n(\exists a_1(a_1$ is an n-subset of $y_1) \longrightarrow \exists a_2(a_2$ is an n-subset of $y_2))$.

11.4) Def $y_1 \overline{\prec} y_2 \longleftrightarrow y_1$ and y_2 are Δ_λ-collections & $\forall n(\exists b_2(b_2$ is an n-superset of $y_2) \longrightarrow \exists b_1(b_1$ is an n-superset of $y_1))$.

Now we say that y_1 and y_2 are the same size ($\approx$) if they have subsets and supersets of exactly the same cardinalities.

11.5) Def $y_1 \approx y_2 \longleftrightarrow y_1 \preceq y_2 \;\&\; y_1 \bar{\prec} y_2 \;\&\; y_2 \preceq y_1 \;\&\; y_2 \bar{\prec} y_1$.

We shall allow ourselves to use these symbols even if y_1 is a Δ_{λ_1}-collection and y_2 a Δ_{λ_2}-collection.

As an example, consider the Δ_2-collections $z_1 = \{x: \varepsilon_4(x)\}$ and $z_2 = \{x: \varepsilon_5(x)\}$. Since z_2 is a subcollection of z_1 , it is obvious that $z_2 \preceq z_1$ and $z_2 \bar{\prec} z_1$. On the other hand, it is not the case that $z_1 \approx z_2$. To see this, let K be a number such that $\varepsilon_4(K)$ but $\neg\varepsilon_5(K)$. (We shall use this same K for several examples in the course of this section.) Then z_1 has K-subsets but no K-supersets, and z_2 has K-supersets but no K-subsets. In this way, we regard z_1 as strictly larger than z_2 . If z_3 is the complement of z_2 in z_1 -- that is, the Δ_2-collection $\{x: \varepsilon_4(x) \;\&\; \neg\varepsilon_5(x)\}$ -- then it is easy to see that $z_1 \preceq z_3$ (if $\varepsilon_4(n)$, then $\{K+1,\ldots,K+n\}$ is an n-subset of z_3) and $z_1 \bar{\prec} z_3$ (if $\varepsilon_4(n)$, then z_3 cannot possibly have an n-superset, since it has an (n+1)-subset); hence $z_1 \approx z_3$.

The reader with one-to-one correspondences on his mind may wonder how the relation $\approx$ compares with more traditional definitions. If K is as above, then certainly $\{1,\ldots,K\} \not\approx \{1,\ldots,K-1\}$. On the other hand, there *is* a "Δ_2-mapping" -- a Δ_2-collection of ordered pairs -- that puts these two sets in one-to-one correspondence: just subtract 1 from every x such that $\neg\varepsilon_5(x)$.

If, however, f is a *function* mapping some superset of a Δ_λ-collection y_1 bijectively to some superset of a Δ_λ-collection y_2 in such a way that elements of y_1 correspond exactly to elements of y_2 via f (this is the most one could ask for, since the domain and range of a function must be sets), then subsets of y_1 correspond to subsets of y_2, and (sufficiently small) supersets of y_1 correspond to (sufficiently small) supersets of y_2 ; hence $y_1 \approx y_2$. Therefore the relation $\approx$ is at least as general as bijective correspondence via functions; it is, in fact, more general, as the following example shows.

Let z_1, z_3, and K be as above, and suppose f is, as above, a function mapping some superset of z_1 bijectively to some superset of z_3 in such a way that elements of z_1 and z_3 correspond. We may certainly assume that every element of Dom f $\cup$ Ran f satisfies ε_3. Then

$$\begin{aligned} y \in z_3(\Delta_2) &\longleftrightarrow \exists x(x \in z_1(\Delta_2) \;\&\; f(x) = y) \\ &\longleftrightarrow \exists^4 x \langle x,y\rangle \in f \\ &\longleftrightarrow x \in f(\textstyle\sum_1^4) \, . \end{aligned}$$

Hence "$y \in z_3(\Delta_2)$" is a $\sum_1^4$-property of y, and so therefore is "$y \in z_3(\Delta_2)$ & $y \leq K$". But this property is equivalent to "$\neg\varepsilon_5(y)$ & $y \leq K$", which is Π_1^5 (and therefore Π_1^4) since "$\varepsilon_5(y)$" is $\sum_1^5$. It follows that all y such that $y \in z_3(\Delta_2)$ & $y \leq K$ form a Δ_1^4-collection, hence a *set* -- which is absurd inasmuch as there is no smallest such y. Thus no such function f can exist, even though $z_1 \approx z_3$ as was shown earlier.

The following properties of $\precsim$, $\precsim$, and $\approx$ are obvious from the definitions.

11.6) y is a Δ_λ-collection $\longrightarrow y \preceq y$. ‖

11.7) y is a Δ_λ-collection $\longrightarrow y \bar{\prec} y$. ‖

11.8) $y_1 \preceq y_2 \;\&\; y_2 \preceq y_3 \longrightarrow y_1 \preceq y_3$. ‖

11.9) $y_1 \bar{\prec} y_2 \;\&\; y_2 \bar{\prec} y_3 \longrightarrow y_1 \bar{\prec} y_3$. ‖

11.10) y_1 and y_2 are Δ_λ-collections $\longrightarrow y_1 \preceq y_2 \vee y_2 \preceq y_1$. ‖

11.11) y_1 and y_2 are Δ_λ-collections $\longrightarrow y_1 \bar{\prec} y_2 \vee y_2 \bar{\prec} y_1$. ‖

11.12) y is a Δ_λ-collection $\longrightarrow y \approx y$. ‖

11.13) $y_1 \approx y_2 \longrightarrow y_2 \approx y_1$. ‖

11.14) $y_1 \approx y_2 \;\&\; y_2 \approx y_3 \longrightarrow y_1 \approx y_3$. ‖

11.15) $y_1 \approx z_1 \;\&\; y_2 \approx z_2 \;\&\; y_1 \preceq y_2 \longrightarrow z_1 \preceq z_2$. ‖

11.16) $y_1 \approx z_1 \;\&\; y_2 \approx z_2 \;\&\; y_1 \bar{\prec} y_2 \longrightarrow z_1 \bar{\prec} z_2$. ‖

A theory of infinite cardinals

Since (11.12)-(11.14) say that $\approx$ is an equivalence relation on Δ_λ-collections, it is natural to examine the equivalence classes, or "cardinals", by adjoining a new sort to our theory. Actually, one new sort is not enough: we must adjoin sorts c_2 ("Δ_2-cardinals", or equivalence classes of Δ_2-collections), c_3 ("Δ_3-cardinals"),... . (A sort c_1 would serve no useful purpose, since "Δ_1-cardinals", or "set-cardinals", are nothing more than numbers satisfying ε.)

Let $S^{\mu\rho}$ be the theory obtained by adjoining $c_2, c_3, \ldots, c_\rho$ to $\tilde{Q}^\mu$; by §3, $S^{\mu\rho}$ is interpretable in $\tilde{Q}^\mu$. In $S^{\mu\rho}$ can be defined the "quotient map" function symbols $\mathrm{Card}_2, \mathrm{Card}_3, \ldots, \mathrm{Card}_\rho$ with the property that if y is a Δ_λ-collection, then $\mathrm{Card}_\lambda\, y$ is its Δ_λ-cardinality: its $\approx$-equivalence class.

As remarked in §10, every Δ_λ-collection is a $\Delta_{\lambda+1}$-collection: just add a dummy quantifier. More precisely, there is a function symbol Dummy_λ such that if y is Δ_λ-collection, then $\mathrm{Dummy}_\lambda\,(y)$ is a $\Delta_{\lambda+1}$-collection and $\forall x(x \in \mathrm{Dummy}_\lambda\,(y)(\Delta_{\lambda+1}) \longleftrightarrow x \in y(\Delta_\lambda))$. If $y_1 \approx y_2$, then clearly $\mathrm{Dummy}_\lambda\,(y_1) \approx \mathrm{Dummy}_\lambda\,(y_2)$ (in fact, with cross-level use of the symbol $\approx$, we can even write $y \approx \mathrm{Dummy}_\lambda\,(y)$) ; hence Dummy_λ induces a function symbol D_λ of type $(c_\lambda ; c_{\lambda+1})$. We shall make a practice of suppressing explicit mention of D_λ , pretending instead that every Δ_λ-cardinal really *is* a $\Delta_{\lambda+1}$-cardinal, a $\Delta_{\lambda+2}$-cardinal,... . In the same way, we write simply $\mathrm{Card}\, y$ for the cardinality (at levels $\lambda, \lambda+1, \ldots$) of the Δ_λ-collection y , and $c(n)$ for the cardinality (at levels $2,3,\ldots$) of the set $\{1,2,\ldots,n\}$ (assuming $\varepsilon(n)$) .

Hereafter let us reserve the letter c , with and without subscripts, for use as a variable of any of the sorts $c_2, c_3, \ldots,$ and the letter e , with subscripts, for the constant symbols $e_1 = \mathrm{Card}\,\{x\colon \varepsilon_1(x)\}$, $e_2 = \mathrm{Card}\,\{x\colon \varepsilon_2(x)\}, \ldots$.

11.17) Def $c_1 \le c_2 \longleftrightarrow \exists y_1 \exists y_2(\mathrm{Card}\, y_1 = c_1 \;\&\; \mathrm{Card}\, y_2 = c_2 \;\&\; y_1 \preceq y_2 \;\&\; y_1 \bar{\prec} y_2)$.

By (11.15) and (11.16) an equivalent definition would be

11.18) $c_1 \leq c_2 \longleftrightarrow \forall y_1 \forall y_2 (\text{Card } y_1 = c_1 \ \& \ \text{Card } y_2 = c_2 \longrightarrow$
$y_1 \preceq y_2 \ \& \ y_1 \bar{\prec} y_2)$. ||

11.19) Def $c_1 < c_2 \longleftrightarrow c_1 \leq c_2 \ \& \ c_1 \neq c_2$.

For example, $e_5 < c(K) < e_4$.

11.20) $c \leq c$.

Proof. By (11.6) and (11.7). ||

11.21) $c_1 \leq c_2 \ \& \ c_2 \leq c_1 \longrightarrow c_1 = c_2$.

Proof. By the definition (11.5). ||

11.22) $c_1 \leq c_2 \ \& \ c_2 \leq c_3 \longrightarrow c_1 \leq c_3$.

Proof. By (11.8) and (11.9). ||

The content of (11.20)-(11.22) is that $\leq$ partially orders all cardinals. The question arises whether this ordering is total; a related question is whether the relations $\preceq$ and $\bar{\prec}$ are the same.

Pseudosets

The answer to both questions is no. Where K is our favorite number such that $\varepsilon_4(K) \ \& \neg \varepsilon_5(K)$, let z be the collection of all x such that $x \leq K \ \& \neg \varepsilon_5(K-x)$ together with all *even* x such that either $x \leq K \ \& \ \varepsilon_5(K-x)$ or $x \geq K \ \& \ \varepsilon_5(x-K)$. Clearly z is a Δ_2-collection. We claim that z has neither a subset nor a superset of cardinality K ; it follows that $z \preceq \{1,\ldots,K\}$ but $\neg z \bar{\prec} \{1,\ldots,K\}$, that $\{1,\ldots,K\} \bar{\prec} z$ but $\neg\{1,\ldots,K\} \preceq z$, and that neither $\mathrm{Card}\ z \leq c(K)$ nor $c(K) \leq \mathrm{Card}\ z$ holds.

First suppose a is a subset of z. Then a contains a largest odd element m and a largest even element n. By the definition of z, $m \leq K$ and in fact $\neg\varepsilon_5(K-m)$; we may assume that $n \geq K$, but in any case $\varepsilon_5(n-K)$ and therefore $\varepsilon_5(n-K+3)$. It follows that $n-K+3 < K-m$, so

$$\begin{aligned} \mathrm{Card}\ a &\leq (m+1) + \frac{1}{2}(n+1-m) \\ &= m + \frac{1}{2}((n-K+3)+(K-m)) \\ &< m + (K-m) \\ &= K . \end{aligned}$$

Thus z has no K-subsets.

Now suppose b is a superset of z. Then there is a smallest odd number m *not* in b and there is a smallest even number n *not* in b. By the definition of z, $n \geq K$ and in fact $\neg\varepsilon_5(n-K)$, and therefore $\neg\varepsilon_5(n-K-3)$; we may assume that $m \leq K$, but in any case $\varepsilon_5(K-m)$. It follows that $K-m < n-K-3$, so

$$\begin{aligned} \mathrm{Card}\ b &\geq (m-1) + \frac{1}{2}(n-1-m) \\ &= m + \frac{1}{2}((n-K-3) + (K-m)) \\ &> m + (K-m) \\ &= K . \end{aligned}$$

Thus z has no K-supersets.

The problem with z is that its cardinality is imprecisely determined; there are some numbers, like K , that are too big to be the cardinality of a subset of z and at the same time too small to be the cardinality of a superset of z . Collections with "precisely determined" cardinalities we shall call *pseudosets*.

11.23) Def y is a Δ_λ-pseudoset $\longleftrightarrow$ y is a Δ_λ-collection & $\forall n(\varepsilon(n) \longrightarrow \exists a($a is an n-subset of y $\vee$ a is an n-superset of y$))$.

11.24) Def c is a pseudoset-cardinal $\longleftrightarrow \exists y($y is a Δ_λ-pseudoset & Card y = c$)$.

It is easy to see that $\{x: \varepsilon_1(x)\}$, $\{x: \varepsilon_2(x)\}, \ldots$ are pseudosets, so that $e_1, e_2, \ldots$ are pseudoset-cardinals. In fact, more can be said.

11.25) Def y is hereditary $\longleftrightarrow$ y is a Δ_λ-collection & $\forall w \forall x(x \in y(\Delta_\lambda)$ & $w \leq x \longrightarrow w \in y(\Delta_\lambda))$.

11.26) y is hereditary $\longrightarrow$ y is a Δ_λ-pseudoset.

Proof. If $\varepsilon(n)$ & $n \in y(\Delta_\lambda)$, then $\{1,\ldots,n\}$ is an n-subset of y . If $\varepsilon(n)$ & $n \notin y(\Delta_\lambda)$, then $\{0,\ldots,n-1\}$ is an n-superset of y . ||

It should go without saying that every *set* all of whose elements satisfy ε is a pseudoset, so that every set-cardinal is a pseudoset-cardinal.

On pseudosets, the relations $\preceq$ and $\bar{\prec}$ really *are* the same; as a consequence, pseudoset cardinals *are* totally ordered. The details follow.

11.27) $y_1 \preceq y_2$ & y_1 is a pseudoset $\longrightarrow y_1 \bar{\prec} y_2$.

Proof. Let b_2 be an n-superset of y_2 . Since y_1 is a pseudoset, y_1 has either an n-subset or an n-superset. If the latter, there is nothing more to prove, so let a_1 be an n-subset of y_1 . Since $y_1 \preceq y_2$, there is an n-subset a_2 of y_2 . Now $a_2 \subseteq y_2 \subseteq b_2$ and $\text{Card } a_2 = \text{Card } b_2 = n$, so $y_2 = a_2 = b_2$ is a *set* with n elements. We claim that $y_1 = a_1$, so that y_1 is also a set with n elements and therefore trivially has an n-superset. To see this, assume some x satisfies $x \in y_1(\Delta_\lambda)$ & $x \notin a_1$. Then $a_1 \cup \{x\}$ is an (n+1)-subset of y_1 . But this implies that y_2 has an (n+1)-subset, which is clearly impossible; it must therefore be the case that $y_1 = a_1$. ||

Similarly:

11.28) $y_1 \bar{\prec} y_2$ & y_2 is a pseudoset $\longrightarrow y_1 \preceq y_2$. ||

In combination, (11.27) and (11.28) imply

11.29) y_1 and y_2 are pseudosets $\longrightarrow (y_1 \preceq y_2 \longleftrightarrow y_1 \bar{\prec} y_2)$. ||

The next two propositions are now immediate in light of (11.10).

11.30) y_1 and y_2 are pseudosets $\longrightarrow ((y_1 \preceq y_2$ & $y_1 \bar{\prec} y_2) \vee$
$(y_2 \preceq y_1$ & $y_2 \bar{\prec} y_1))$. ||

11.31) c_1 and c_2 are pseudoset-cardinals $\longrightarrow c_1 < c_2 \vee$
$c_1 = c_2 \vee c_2 < c_1$. ||

Two more little pieces of mumbo-jumbo: a collection with the same cardinality as a set *is* a set, and a collection with the same cardinality as a pseudoset *is* a pseudoset.

11.32) $\text{Card } y = c(n) \longrightarrow \exists a(a$ is a set & $\text{Card } a = n$ &
$\forall x(x \in a \longleftrightarrow x \in y(\Delta_\lambda)))$.

Proof. If $\text{Card } y = c(n)$, then y has an n-subset a and an n-superset b , so necessarily $y = a = b$. ||

11.33) Card y is a pseudoset-cardinal $\longrightarrow$ y is a pseudoset.

Proof. If $\varepsilon(n)$, then Card y and $c(n)$ are comparable by (11.31); hence y has either an n-subset or an n-superset. ||

Properties of pseudoset-cardinals

For a given cardinal c , there can of course be many different Δ_λ-collections y such that $\text{Card } y = c$. If c is a pseudoset-cardinal, however, there is a canonical representative for c : the unique *hereditary* y such that $\text{Card } y = c$.

11.34) c is a pseudoset-cardinal $\longrightarrow \exists! y(y$ is hereditary &
$\text{Card } y = c)$.

Proof. Let z be any pseudoset with $\text{Card } z = c$, and let y be the collection of all n such that z has an $(n+1)$-subset. Clearly y is hereditary. If z has an n-subset, then so does y (namely $\{0,\ldots,n-1\}$) ; hence $z \preceq y$. If z has an n-superset, then z cannot have an $(n+1)$-subset, so $\{0,\ldots,n-1\}$ is an n-superset of y ; hence $y \preceq z$. Because y and z are pseudosets, we have $y \approx z$ and therefore $\text{Card } y = \text{Card } z = c$.

To prove uniqueness, let y' be another hereditary collection with cardinality c. If $n \in y(\Delta_\lambda)$, then $\{0,\ldots,n\}$ is an $(n+1)$-subset of y. Therefore y' has an $(n+1)$-subset. The largest element of a set with $n+1$ elements must be at least n ; since y' is hereditary, it follows that $n \in y'(\Delta_\lambda)$. Likewise $y' \subseteq y$. ‖

(Actually, y is "unique" only up to the relation of having the same elements; see the introductory remarks to this section.)

11.35) Def $H(c) = y \longleftrightarrow c$ is a pseudoset-cardinal & y is hereditary & $\text{Card } y = c$, otherwise $y = 0$.

Open question: if c is the cardinality of a Δ_λ-pseudoset z, must $H(c)$ be a Δ_λ-collection (for the same λ) ? From the fact that n is in $H(c)$ if and only if

$$\exists a(a \text{ is a set } \& \text{ Card } a = n+1 \ \& \ \forall x(x \in a \longrightarrow x \in z(\Delta_\lambda))) ,$$

it is clear that the answer is yes *if* it is true that bounded quantifiers do not affect the complexity of a collection; see the discussion leading up to (10.35). In any case $H(c)$ is at worst a $\Delta_{\lambda+2}$-collection.

Here are some simple but useful definitions and observations.

11.36) Def c is infinite $\longleftrightarrow \neg\exists n(c = c(n))$.

11.37) Def y is inductive $\longleftrightarrow y$ is a Δ_λ-collection & $\forall m(m \in y(\Delta_\lambda) \longrightarrow m+1 \in y(\Delta_\lambda))$.

11.38) y is inductive $\longrightarrow$ Card y is infinite.

Proof. If Card $y = c(n)$, then y has an n-subset a . Let the largest element of a be m ; then since y is inductive, $a \cup \{m+1\}$ is an $(n+1)$-subset of y , a contradiction. ||

11.39) c is an infinite pseudoset-cardinal $\longrightarrow H(c)$ is inductive.

Proof. If $m \in H(c)$ but $m+1 \notin H(c)$, then $H(c) = \{0,\ldots,m\}$ is a set and c is not infinite. ||

There is a natural way to define the sum of two pseudoset-cardinals, but it does *not* involve unions. In fact, if $y_1 = \{x: \varepsilon_5(x)\}$, $y_2 = \{x \leq K \mathbin{\&} \neg\varepsilon_5(x)\}$, and $y_3 = \{x: x \leq K+1 \mathbin{\&} \neg\varepsilon_5(x)\}$, then $y_2 \approx y_3$ but $y_1 \cup y_2 \not\approx y_1 \cup y_3$, even though y_1 is disjoint from both y_2 and y_3 . Before presenting the *right* definition, we modify the definition of $H(c)$ slightly.

11.40) Def $H'(c) = y \longleftrightarrow c$ is a pseudoset-cardinal & $y = \{n+1: n \in H(c)\}$, otherwise $y = 0$.

Clearly Card $H'(c)$ = Card $H(c) = c$. If $c = c(n)$, then $H(c) = \{0,\ldots,n-1\}$ and $H'(c) = \{1,\ldots,n\}$. If c is infinite, then by (11.39) the only difference between $H(c)$ and $H'(c)$ is the presence of 0 in the former.

11.41) Def $c_1+c_2 = c \longleftrightarrow c_1$ and c_2 are pseudoset-cardinals & $c = \text{Card}\{j: j \geq 1 \;\&\; \exists m \exists n(m \in H'(c_1) \;\&\; n \in H'(c_2) \;\&\; j \leq m+n)\}$, otherwise $c = c(0)$.

It is easy to see that + behaves properly on set-cardinals; that is, $c(m)+c(n) = c(m+n)$. On infinite cardinals, this addition operation, unlike its counterpart in Cantorian set theory, is non-trivial: one can check that

$$c(K)+e_5 < c(2K)+e_5 = (c(K)+e_5) + (c(K)+e_5) .$$

On the other hand, it is certainly true that $e_4+e_4 = e_4$ (e_4 "respects addition") and that $e_5+e_4 = e_4$. These examples show that $c_1 < c_2$ does *not* imply $c_1+c < c_2+c$. (That implication is valid, however, if c is a set-cardinal.) Further properties: + is commutative and associative, and if at least one of c_1 and c_2 is infinite, then c_1+c_2 is infinite.

The above approach makes it clear how to define

11.42) Def $c_1 \cdot c_2 = c \longleftrightarrow c_1$ and c_2 are pseudoset-cardinals & $c = \text{Card}\{j: j \geq 1 \;\&\; \exists m \exists n(m \in H'(c_1) \;\&\; n \in H'(c_2) \;\&\; j \leq m \cdot n)\}$, otherwise $c = c(0)$,

as well as $c_1 \# c_2$, $c_1 \#_1 c_2, \ldots$. One can define also a subtraction operation,

11.43) Def $c_1-c_2 = c \longleftrightarrow c_1$ and c_2 are pseudoset-cardinals & $c = \text{Card}\{m: m \geq 1 \;\&\; \forall n(n \in H'(c_2) \longrightarrow m+n \in H'(c_1))\}$, otherwise $c = c(0)$, and likewise division and inverse smashes. Warning: clearly $(c_1-c_2)+c_2 \leq c_1$ and $c_1 \leq (c_1+c_2)-c_2$, but equality does not always hold. (Example: $(c(K)-e_5)+e_5 = c(K)-e_5 < c(K) < c(K)+e_5 = (c(K)+e_5)-e_5$.) Equality *does* hold if c_2 is a set-cardinal.

The following curious theorem gives some idea of what infinite pseudoset-cardinal arithmetic can accomplish, and at the same time sheds some light on the ordering relation $<$. The statement is that between every two infinite pseudoset-cardinals at least one of which respects addition there is one that does *not* respect addition.

11.44) c_1 and c_2 are infinite pseudoset-cardinals & $(c_1+c_1 = c_1 \vee c_2+c_2 = c_2)$ & $c_1 < c_2 \longrightarrow \exists c(c$ is an infinite pseudoset-cardinal & $c+c \neq c$ & $c_1 < c < c_2)$.

Proof. There is some k such that $c_1 < c(k) < c_2$; replacing k by $k+1$ if necessary, we may assume k is even. First suppose $c_1+c_1 = c_1$. Since $k \notin H'(c_1)$, it follows that $\frac{1}{2}k \notin H'(c_1)$. Let c be $c(k)-c_1$. Then c is infinite (certainly $H'(c)$ is inductive) and $c_1 < c(\frac{1}{2}k) < c < c(k) < c_2$. Moreover, $H'(c)$ contains $\frac{1}{2}k$ but not k ; thus $c < c+c$.

Now suppose $c_2+c_2 = c_2$, so $2k \in H'(c_2)$; let c be $c(k)+c_1$. Then $k \in H'(c)$ but $2k \notin H'(c)$; thus $c_1 < c(k) < c < c(2k) < c_2$, c is infinite, and $c < c(2k) < c+c$. ||

In particular, if c_1 is any cardinal that does respect addition (for instance, e_5), then there cannot be a smallest infinite cardinal c_2 with $c_1 < c_2$; infinite pseudoset-cardinals, therefore, are *not* well-ordered. It is natural to ask whether there may be a *first* infinite pseudoset-cardinal. If there were, it would be quite surprising. In the absence of a definitive answer, we can at least say that if such a cardinal exists, it respects addition:

11.45) c_1 is an infinite pseudoset-cardinal & $c_1+c_1 \neq c_1 \longrightarrow$
$\exists c$(c is an infinite pseudoset-cardinal & $c < c_1$) .

Proof. Let y be $\{m: m \geq 1 \text{ \& } m+m \in H'(c_1)\}$. If c_1 does not respect addition, then y is a proper subcollection of $H'(c_1)$, so Card y (which is actually the cardinal $c_1/c(2)$) is strictly smaller than c_1 . Since $H'(c_1)$ is inductive, so is y ; hence Card y is infinite. ||

From (11.45) and preceding remarks, it follows that while for all we know there may be a smallest infinite pseudoset-cardinal, there definitely cannot be a second-smallest as well.

An application

Let us revisit (9.37), the existence of primes p with $\neg\varepsilon_\nu(p)$, and give a new proof using the techniques of this section. To be concrete, we prove the existence of p such that $\varepsilon_4(p)$ & $\neg\varepsilon_5(p)$. In fact, we prove something stronger: the primes p with $\varepsilon_4(p)$ form a pseudoset whose cardinality is at least e_5 . (If there were no p with $\varepsilon_4(p)$ & $\neg\varepsilon_5(p)$, then $\{p: \varepsilon_4(p)\} = \{p: p \leq K\}$ would be a *set* whose cardinality would be some $c(n) < e_5$.)

Let $\neg\varepsilon_4(M)$, and let $p_1 = 2$, $p_2 = 3,\ldots,p_m$ be a sequence enumerating the primes $\leq M$. It is clear that for each $n \leq m$, $\{p_1,\ldots,p_n\}$ is either an n-subset or an n-superset of $y = \{p\colon \varepsilon_4(p)\}$ and that in particular $\{p_1,\ldots,p_m\}$ is an m-superset of y ; hence y is a pseudoset. (A generalization of (11.26) lurks here: the intersection of a hereditary collection with a set is always a pseudoset.) To show that $e_5 \leq \text{Card } y$, it suffices to show that if $\varepsilon_5(n)$, then $\{p_1,\ldots,p_n\}$ is a subset of y .

Suppose not; then $\{p_1,\ldots,p_n\}$ is a superset of y . By the fundamental theorem of arithmetic, every number satisfying ε_4 can be written as a product of powers of elements of y ; hence as a product of powers of $p_1,\ldots,p_n$; hence in the form $p_1^{j_1}p_2^{j_2}\ldots p_n^{j_n}$ where every exponent j_i satisfies $\varepsilon_5(j_i)$. (After all, $\neg\varepsilon_5(j) \longrightarrow \neg\varepsilon_4(2^j)$.) In particular, $\{x\colon \varepsilon_4(x)\}$ is a subcollection of $\{p_1^{j_1}p_2^{j_2}\ldots p_n^{j_n}\colon j_1 < K \;\&\; j_2 < K \&\ldots\& j_n < K\}$. But this last collection is a set, and its cardinality is K^n . Since $\varepsilon_4(K) \;\&\; \varepsilon_5(n)$, it follows that $\varepsilon_4(K^n)$; hence no set of this cardinality can be a superset of $\{x\colon \varepsilon_4(x)\}$. From this contradiction, we conclude that $\{p_1,\ldots,p_n\}$ is a subset of y , and thus $e_5 \leq \text{Card } y$.