WHAT IS NONSTANDARD ANALYSIS?

W. A. J. LUXEMBURG, California Institute of Technology

1. Introduction. The subject referred to in the title with which we shall deal may seem perhaps at first sight to be far removed from the general topic "The Foundations of Mathematics" of the Symposium. This relatively new field which was created by Abraham Robinson (see [7]) may be looked upon, however, as a major contribution to the foundations of analysis. Furthermore, it is another splendid example of an application of mathematical logic.

The development of mathematical analysis by using infinitely small and infinitely large numbers has been a subject of constant interest and controversy in the history of mathematics. Going back in history we discover that Leibniz was one of the strongest advocates of a method involving infinitely small and infinitely large numbers in the early stages of the development of the calculus. The reason why the theory of infinitesimals gradually fell into disrepute and was replaced later by the ε, δ-method must be sought in the fact that neither Leibniz nor his successors were able to state with sufficient precision just what rules were supposed to govern their system of infinitely large and infinitely small numbers. Although Leibniz stated the principle that what holds for the finite numbers should also hold for the numbers in the extended system, which includes the infinitely small and infinitely large numbers, it is not at all clear in his writings what sort of laws about numbers his principle was supposed to apply to.

It was Abraham Robinson's recent discovery, mentioned above, that the notions of model theory can clarify the notions of infinitely small and infinitely large. Robinson shows that mathematical analysis can be developed by imbedding the real number system R in a proper extension $^*R$ of R which possesses in a certain sense the same properties as R. It is well known that such an extension $^*R$ must be non-Archimedean and this is the fact that enabled Robinson to define in $^*R$ the infinitely small and infinitely large numbers whose existence was taken for granted by Leibniz and his followers. From the well-known result that there exist systems of axioms for the real number system which are categorical, that is, determine the real number systems uniquely up to isomorphism it may seem at first very paradoxical that such systems $^*R$ exist. This sort of paradox has been one of the main sources of the condemnation of the theory of infinitesimals and infinitely large numbers as a tool in analysis. The paradox vanishes completely, however, if we follow Robinson's idea to restrict the statement "the same properties" to a specified collection of properties of R which can be formulated in a specified formal language with the appropriate interpretation in R as well as in $^*R$, and in which the classical isomorphism theorem for the real number system cannot be formulated. Of course it is at this point that model theory comes into play which by means of the compactness principle guarantees the existence of such systems $^*R$.

There is, however, another way to establish the existence of $^*R$. This method is known as the construction of models in the form of ultraproducts. It has the advantage that it can be developed within the framework of axiomatic set theory. We shall follow this procedure here. Sections 2, 3 and 4 are entirely devoted to a discussion of the existence of $^*R$. In our approach we follow very closely the development as given by Abraham Robinson and Elias Zakon in their paper entitled *A set-theoretical characterization of enlargements* and which appeared in **[6]**. In the remaining six sections it is illustrated by means of examples in which sense the theory of infinitely small and infinitely large numbers can be used as a tool in analysis. The topics which were selected for this purpose include the theory of limits, Euler's product formula for the sine, and the existence of functions which are not measurable in the sense of Lebesgue.

The ideas of nonstandard analysis were subsequently successfully applied to other branches of mathematics. These developments are not taken up here as they are beyond the scope of the present introductory paper. But we like to refer the interested reader, who for instance would like to know with what great success this method was used by A. Robinson and A. Bernstein to solve the invariant subspace problem for a certain class of bounded operators on a Hilbert space, to Robinson's book **[8]** and the papers **[1]**, **[2]** and **[15]**. Furthermore, we would like to draw the readers' attention to reference **[6]** which is the *Proceedings of the International Symposium on Nonstandard Analysis*, which was held at the California Institute of Technology in 1967. Its contents, consisting of more than twenty papers, gives the latest developments in this field.

Finally, the author would like to state that the present paper is mainly expository in nature. It is particularly directed to those mathematicians who would like to get acquainted with this new tool in analysis. We do hope, however, that also the specialists in the field will find something new and of interest in this paper.

2. Definition of the structure $\hat{R}$ and some of its properties. The earlier version of nonstandard analysis (see **[7]** and **[3]**) rests on the formulation of the properties of R which can be formulated in a first order language, which means briefly that quantification in the formal language is permitted only on variables ranging over real numbers. One need not go far in analysis, however, to realize the need for a richer language in which statements containing expressions such as for example "For all nonempty sets of natural numbers..." or "There exists a continuous function..." can be formulated. In this connection it is also good to observe that even some of the axioms of the real number system are outside the language of the lower predicate calculus. For example, Dedekind's completion axiom involving quantification with respect to ordered pairs of sets (Dedekind cuts) is such an axiom. In order to cope with this difficulty we shall use the framework of axiomatic set theory in terms of which the theory of real numbers can be developed. The formal language will be a lower order language whose constants will range over sets and numbers. We shall now present this development here in some detail. We shall assume that the reader is familiar with the elements of naive set theory and with some of the definitions and results concerning the lower predicate calculus.

Let R denote as usual the set of real numbers. Then we define inductively the sets $R_0 = R$ and $R_{n+1} = P(\bigcup_{k=0}^{n} R_k)$ $(n = 0, 1, 2, \cdots)$, where $P(X)$ denotes the set of all subsets of X. The union of all the sets R_n, $\bigcup_{n\geqq 0} R_n$ is called the **superstructure** on R and will be denoted by $\hat{R}$. The elements of $\hat{R}$ are called the **entities** of the superstructure $\hat{R}$. The elements of $R_0 = R$, that is the real numbers, on which the superstructure is based are sometimes also referred to as the **individuals** of $\hat{R}$.

We shall assume that an ordered pair (a, b) is defined in the sense of Kuratowski by $(a, b) = \{\{a\}, \{a, b\}\}$ and that n-tuples $(a_1, \cdots, a_n)$ are defined inductively by $(a) = a$, $(a_1, \cdots, a_n) = ((a_1, \cdots, a_{n-1}), a_n)$. Then it follows immediately that relations defined as sets of n-tuples $(n = 1, 2, \cdots)$ are all entities of $\hat{R}$. Since the algebraic operations of R can be defined in terms of three place relations as follows: $ab = c$ if and only if $(a, b, c) \in P \in \hat{R}$ and $a + b = c$ if and only if $(a, b, c) \in S \in \hat{R}$ and the order relation is a binary relation it follows that the axioms and the properties of R can be expressed in terms of certain entities of $\hat{R}$. The remaining part of this section will now be devoted to making this more precise.

The entities of $R_n - R_{n-1}$ $(n \geqq 1)$ are called of rank n in $\hat{R}$. The individuals are given the rank 0. The reader should observe that by means of this definition, the empty set gets assigned rank 1. If $a \in \hat{R}$ is not empty, then the rank of a is the smallest natural number n such that $a \in R_n$. It is also easy to see that if $a_1, \cdots, a_n \in \hat{R}$, then rank $(a_1, \cdots, a_n) = \max(\text{rank}\, a_1, \cdots, \text{rank}\, a_n) + 2n$.

Some minor set-theoretical properties of $\hat{R}$ are collected, for later references, in the following lemma.

LEMMA 2.1. (i) *$R_p \subset R_n$ for all $n \geqq p \geqq 1$.*

(ii) *$\bigcup_{k=0}^{n} R_k = R_0 \cup R_n$ for all $n \geqq 1$.*

(iii) *$R_k \in R_{n+1}$ for all $0 \leqq k \leqq n$ and for all $n \geqq 0$.*

(iv) *If $x \in y \in R_n$ $(n \geqq 1)$, then $x \in R_0 \cup R_{n-1}$.*

(v) *If $(x_1, \cdots, x_n) \in y \in R_p$ $(p \geqq 1)$, then $x_1, \cdots, x_n \in R_0 \cup R_{p-1}$. In particular, if an entity $\Phi \in \hat{R}$ is a binary relation, then its domain,* $\text{dom}\,\Phi = \{x : (\exists y)(x, y) \in \Phi\}$ $\in \hat{R}$, *and its range,* $\text{ran}\,\Phi = \{y : (\exists x)(x, y) \in \Phi\} \in \hat{R}$.

Proof. (i) If $x \in R_p$, then $x \subset \bigcup_{k=0}^{p-1} R_k$, and so $x \subset \bigcup_{k=0}^{q} R_k$ for all $q \geqq p - 1$. Hence, $x \in P(\bigcup_{k=0}^{q} R_k) = R_{q+1}$ for all $q + 1 \geqq p$.

(ii) For $n \geqq 1$, $R_n \subset R_{n+1}$, and so since R_0 is disjoint from all R_n $(n \geqq 1)$ it follows that for all $n \geqq 1$ we have $\bigcup_{k=0}^{n} R_k = R_0 \cup R_n$.

(iii) Since by (ii) we have that $R_k \subset R_0 \cup R_n$ $(0 \leqq k \leqq n)$ we obtain that $R_k \in P(R_0 \cup R_n) = R_{n+1}$.

(iv) If $y \in R_n$, $(n \geqq 1)$, then $y \subset R_0 \cup R_{n-1}$, and so $x \in y$ implies that $x \in R_0 \cup R_{n-1}$.

(v) If $(x_1, \cdots, x_n) \in y \in R_p$ $(p \geqq 1)$, then $(x_1, \cdots, x_n) \in R_0 \cup R_{p-1}$. Hence, $\{\{x_1\}, \{x_1, (x_2, \cdots, x_n)\}\} \in R_0 \cup R_{p-1} = R_0 \cup P(R_0 \cup R_{p-2})$ implies $x_1 \in R_0 \cup R_{p-2} \subset R_0 \cup R_{p-1}$, and similarly for the entities $x_2, \cdots, x_n$.

The formal language will now be introduced.

The **atomic symbols** of L are: (i) The **connectives** $\wedge$, $\vee$, $\Rightarrow$, $\Leftrightarrow$, $\neg$, for "and", "or", "implies", "if and only if", "not" respectively. (ii) The **variables**, a countably infinite sequence usually denoted by $x, y, \cdots$ with or without subscripts. (iii) The **quantifiers** $(\exists\ \cdot)$-existential, and $(\forall\ \cdot)$-universal. (iv) Brackets [ ], used for grouping formulas as usual in mathematics. (v) **The basic predicate**, $\in$ read "member of" with one open place to the left and to the right of it. (vi) **Extra logical constants** (briefly, constants). This is a set of symbols of which there are enough to be put in one-to-one correspondence with the entities of whatever structure may be under consideration. This set of constants is usually infinite but fixed. Furthermore, constants are usually denoted by Roman letters with or without subscripts from the beginning of the alphabet, and other symbols such as the numerals $0, 1, 2, \cdots$.

We shall now assume that the set of constants of L is brought in one-to-one correspondence with all the entities of the structure $\hat{R}$ and we shall from now on identify the constants of L with the entities of $\hat{R}$ so that $\hat{R}$ is part of L. If such an identification has been established, then we refer to $\hat{R}$ as an *L-structure.*

The interpretation of the basic predicate $\in$ of L in $\hat{R}$ will be the membership relation of axiomatic set theory.

From the atomic formulas $\alpha \in \beta$, where the symbols α and β may denote constants and variables, the well-formed formulas (wff) are obtained in successive stages by applying the connectives and quantifiers. At the same time brackets are introduced in such a way that the formation of the formula can be unambiguously determined. More precisely, if V is an atomic formula, then $[V]$ is a wff, if V, W are wff, then $[V \wedge W]$, $[V \vee W]$, $[\neg V]$, $[V \Rightarrow W]$, $[V \Leftrightarrow W]$ are wff; and if V is a wff, then $[(\forall x)V]$ and $[(\exists x)(V)]$ are wff, where x denotes an arbitrary variable, provided x does not already appear in V under the sign of a quantifier. Furthermore, we shall adhere to the terminology that in $[(\forall x)V]$ and $[(\exists x)V]$, V is called the **scope** of the quantifier and in all the wff which can be obtained from these by the further repeated applications of connectives and quantifiers. A variable x is called *free* in a wff V if x is not in $(\exists x)$ or $(\forall x)$ or in the scope of a quantifier in V. A wff is called a **sentence** if every variable is in the scope of a quantifier, otherwise it is called a **predicate.** A wff V in L is said to be in **prenex normal form**, if in the formation of V from atomic formulas the quantifiers are applied after the connectives, that is, if the connectives are in the scope of *all* quantifiers. In symbols, $V = (qx_n) \cdots (qx_1)W$, where $(q\ \cdot)$ denotes either $(\exists\ \cdot)$ or $(\forall\ \cdot)$ and where W is a wff without quantifiers, is a wff in prenex normal form. One of the basic results of the lower predicate calculus states that every wff is equivalent to a wff which is in prenex normal form (see [**8**], p. 10).

For our purpose we shall only consider those wff of L which have the property that all quantifiers are of the form "$(\forall x)[[x \in A] \Rightarrow \cdots]$" and $(\exists x)[[x \in A] \wedge \cdots]$" where A is an entity of $\hat{R}$ and which are called the **admissible** wff. Thus a wff is admissible whenever the domain of every quantifier occurring in it is a specific entity of $\hat{R}$. The set of admissible wff of L will be denoted by $K = K(L)$ and the subset of K of all admissible sentences which hold in $\hat{R}$ will be denoted by $K_0 = K_0(L)$.

At this point the reader should do well to observe that all statements in analysis dealing with numbers, sets of numbers, relations between numbers, relations between sets and numbers, and so on, and which hold in R can be expressed as admissible sentences of L which are in K_0. For instance, the sentence of K_0

$$(\forall a)(\forall b)(\forall c)[a, b, c \in R] \Rightarrow [P(a, b, c) \Rightarrow P(b, a, c)]$$

expresses that R is commutative (P is the constant denoting the three place relation of multiplication).

Any $^*L$-structure $^*(\hat{R})$ in which the L-structure $\hat{R}$ can be properly imbedded and for which all admissible sentences of $\hat{R}$ which hold in $\hat{R}$ with appropriate interpretation of the symbols in $^*(\hat{R})$ also hold in $^*(\hat{R})$ will be called a **higher order nonstandard model** of $\hat{R}$. In that case, it turns out that the set $^*R$ of individuals of $^*(\hat{R})$ is a totally ordered field of which R is a proper subfield. But $^*(\hat{R})$ is not the superstructure determined by $^*R$. In fact, if $A = P(R)$ is the constant which denotes the entity of $\hat{R}$ of all subsets of R, then under the imbedding of $\hat{R}$ in $^*(\hat{R})$ this constant will not denote the set of *all* subsets of $^*R$ as might be expected at first, but only a subsystem of the power set of $^*R$, and so on. How this all will come about will be explained in detail in the next section.

3. Models of $\hat{R}$ that are ultrapowers. We begin by recalling some definitions and elementary results from the theory of filters.

Let I denote a nonempty set. By a **filter** over I we mean a nonempty set $\mathscr{F}$ of of subsets of I such that the empty set $\phi \notin \mathscr{F}$, $\mathscr{F}$ is closed under **finite** intersections, and $F \subset G$ and $F \in \mathscr{F}$ implies $G \in \mathscr{F}$. In particular $\mathscr{F} \neq \emptyset$ implies that $I \in \mathscr{F}$. A filter $\mathscr{F}_1$ is called **finer** than a filter $\mathscr{F}_2$ ($\mathscr{F}_2 \leqq \mathscr{F}_1$) whenever $F \in \mathscr{F}_2$ implies $F \in \mathscr{F}_1$. This relation orders the set of all filters over I and the filter $\{I\}$ is its smallest element. A filter $\mathscr{F}$ is called an **ultrafilter** whenever it is not properly contained in any other filter, that is, the ultrafilters are the maximal elements of the ordered set of filters. Concerning ultrafilters we have the following important characterization. *A filter $\mathscr{F}$ is an ultrafilter if and only if for every $F \subset I$ either $F \in \mathscr{F}$* or *$I - F \in \mathscr{F}$*. The latter statement is easily seen to be equivalent to: If

$$\bigcup_{i=1}^{n} F_i \in \mathscr{F}(F_i \subset I, \ i = 1, 2, \cdots, n),$$

then $F_i \in \mathscr{F}$ for at least one index i, and so, is itself a characterization of the concept of an ultrafilter.

A filter $\mathscr{F}$ is called **δ-incomplete,** whenever there exists a sequence $F_n \in \mathscr{F}$ $(n = 1, 2, \cdots)$ such that $\bigcap_{n=1}^{\infty} F_n \notin \mathscr{F}$, and a filter $\mathscr{F}$ is called **δ-complete** whenever it is not δ-incomplete. A filter $\mathscr{F}$ is called **free** whenever $\cap (F : F \in \mathscr{F}) = \varnothing$. It is not known whether δ-complete free ultrafilters exist. This problem is known as Ulam's measure problem. It is easy to see, however, that a *δ-incomplete ultrafilter is free*. It follows from the following simple result, the proof of which we leave to the reader as an exercise.

An ultrafilter $\mathscr{U}$ is δ-incomplete if and only if there exists a countable partition $\{I_n : n = 1, 2, \cdots\}$ of the set I over which $\mathscr{U}$ is defined such that $I_n \notin \mathscr{U}$ for all $n = 1, 2 \cdots$.

From this result in conjunction with Zorn's lemma it follows now also easily that on every infinite set there exist plenty of δ-incomplete ultrafilters. For further information on filters we refer the reader to the paper of the author: *A general theory of monads*; which appeared in [**6**].

We shall now turn to a description of a structure which is an ultrapower of $\hat{R}$.

Let I be an infinite set, let $\mathscr{U}$ be a δ-incomplete ultrafilter of subsets of I and let $\{I_n : n = 1, 2, \cdots\}$ be a countable partition of I satisfying $I_n \notin \mathscr{U}$ for all $n = 1, 2, \cdots$ which will be kept fixed.

By $\hat{R}^I$ we denote as usual the set of all mappings of I into $\hat{R}$. There exists a natural imbedding $a \to {}^*a$ of $\hat{R}$ into $\hat{R}^I$ defined by ${}^*a(i) = a$ for all $i \in I$, that is $\hat{R}$ is identified in $\hat{R}^I$ by the constant mappings. The undefined basic predicates "$=$" and "$\in$" of $\hat{R}$ can be extended to $\hat{R}^I$ by means of the following $\mathscr{U}$-dependent definitions.

DEFINITION 3.1. *If $a, b \in \hat{R}^I$, then $a =_{\mathscr{U}} b$ if and only if $\{i : a(i) = b(i)\} \in \mathscr{U}$, and $a \in_{\mathscr{U}} b$ if and only if $\{i : a(i) \in b(i)\} \in \mathscr{U}$.*

Since it is an immediate consequence of $I \in \mathscr{U}$ that if $a, b \in \hat{R}$, then $a = b$ if and only if ${}^*a =_{\mathscr{U}} {}^*b$, and $a \in b$ if and only if ${}^*a \in_{\mathscr{U}} {}^*b$ it follows that the relations "$=_{\mathscr{U}}$" and "$\in_{\mathscr{U}}$" are $\mathscr{U}$-extensions of "$=$" and "$\in$" of $\hat{R}$. For the sake of simplicity we shall from now on retain the original notation "$=$" for "$=_{\mathscr{U}}$" and "$\in$" for "$\in_{\mathscr{U}}$".

In order to justify the definition we are going to show that for all $a, b \in \hat{R}^I$ either $a = b$ or not $(a = b)(a \neq b)$ holds, and $a \in b$ or not $(a \in b)(a \notin b)$ holds. Since the proof for both cases is the same we shall only verify it for "$=$". If $a, b \in \hat{R}^I$, then we set

$$U_1 = \{i : a(i) = b(i)\} \text{ and } U_2 = \{i : a(i) \neq b(i)\}.$$

From $U_1 \cup U_2 = I \in \mathscr{U}$ it follows from the basic property of an ultrafilter that either $U_1 \in \mathscr{U}$ and $U_2 \notin \mathscr{U}$ or $U_1 \notin \mathscr{U}$ and $U_2 \in \mathscr{U}$, that is, by Definition 3.1., either $a = b$ or not $(a = b)$ holds.

Having justified the definition we can justify further the suggestion that the relations "=", "$\in$" in $\hat{R}^I$ behave like equality and membership of set theory. Since the individuals of $\hat{R}$ are without members but different from $\emptyset$, that is, set theory in $\hat{R}$ is based on a set of so-called urelements, equality of sets in terms of $\in$ should read "$a = b$" if and only if $a \in c$ and $b \in c$ for all $c \in \hat{R}^I$. But this can now be immediately verified by observing that if $a = b$ and $a \in c$, then

$$U_1 = \{i: a(i) = b(i)\} \in \mathscr{U} \text{ and } U_2 = \{i: a(i) \in c(i)\} \in \mathscr{U}$$

implies by the filter properties that $U_1 \cap U_2 \in \mathscr{U}$, and so $i \in U_1 \cap U_2$ implies that $b(i) \in c(i)$, that is, $b \in c$. Conversely, we have that if $a, b \in \hat{R}^I$, then $a \in \{x: x = a$ and $x \in \hat{R}^I\} = \{a\}$, implies that $b \in \{a\}$, that is $b = a$. That the relation of equality, as defined in Definition 3.1, is an equivalence relation is immediately clear. That it satisfies the rule of substitution in $\in$, namely,

$$(\forall a)(\forall b)(\forall c)(\forall d)[[a \in b] \wedge [a = c] \wedge [b = d]] \Rightarrow [c \in d]$$

can be verified in the same way by using the properties of $\mathscr{U}$.

Continuing this process we can show, by using the basic properties of $\mathscr{U}$, that one by one the statements which hold in $\hat{R}$ hold in $\hat{R}^I$ under the defined interpretation of the basic predicates. We shall of course not follow this procedure but present in a general fashion that a certain substructure of $\hat{R}^I$ has the same properties as $\hat{R}$.

For this purpose we shall assume that the elements of $\hat{R}^I$ are identified in a one-to-one manner with the constants of a formal language $^*L$. Furthermore, $^*L$ is assumed to have two basic predicates "=" (equality) and "$\in$" (membership) which are identified with the corresponding relations of $\hat{R}^I$. Thus we obtain an $^*L$-structure $\hat{R}^I$ whose set of true sentences depends on $\mathscr{U}$. A certain substructure of our $^*L$-structure will be singled out which we shall show to satisfy, in a certain sense, the sentences of K_0.

In the following lemma, however, we shall first list for later reference, some of the basic properties of the imbedding $a \to {}^*a$ of $\hat{R}$ into $\hat{R}^I$.

LEMMA 3.2. (i) $^*\emptyset = \emptyset$.

(ii) *If* $a, b \in \hat{R}$, *then* $a \subset b$ *implies* $^*a \subset {}^*b$.

(iii) *If* $a, b \in \hat{R}$, *then* $a \in b$ *if and only if* $^*a \in {}^*b$.

(iv) *For all* $a \in \hat{R}$ *we have* $^*\{a\} = \{^*a\}$.

(v) *If* $a_1, \cdots, a_n \in \hat{R}$, *then* $^*(\bigcup_{i=1}^n a_i) = \bigcup_{i=1}^n {}^*a_i$, $^*(\bigcap_{i=1}^n a_i) = \bigcap_{i=1}^n {}^*a_i$, $^*\{a_1, \cdots, a_n\} = \{^*a_1, \cdots, {}^*a_n\}$, $^*(a_1, \cdots, a_n) = ({}^*a_1, \cdots, {}^*a_n)$, and $^*(a_1 \times \cdots \times a_n) = {}^*a_1 \times \cdots \times {}^*a_n$.

(vi) *For all* $a, b \in \hat{R}$ *we have* $^*(a - b) = {}^*a - {}^*b$.

(vii) *If* $b \in \hat{R}$ *is a binary relation, then* $^*(\operatorname{dom} b) = \operatorname{dom} {}^*b$, $^*(\operatorname{ran} b) = \operatorname{ran} {}^*b$, *and for all* $a \in \hat{R}$ *we have*

$$^*(b(a)) = {}^*\{y: (\exists x)(x \in a \wedge (x, y) \in b)\} = {}^*b({}^*a) = \{y: (\exists x)(x \in {}^*a \wedge (x, y) \in {}^*b)\}.$$

Proof. We shall only prove (vi) since the proofs of the other statements are similar. For these proofs we refer the reader to the proofs of Theorems 7.1 and 7.7. of [3].

(vi) If $c \in {}^*(a-b)$, then $U_1 = \{i: c(i) \in a-b\} \in \mathscr{U}$ implies, using $U_1 \subset U_2 = \{i: c(i) \in a\} \in \mathscr{U}$ that $c \in {}^*a$ and using $U_1 \subset U_3 = \{i: c(i) \notin b\} \in \mathscr{U}$ which, since $\mathscr{U}$ is an ultrafilter, is equivalent to $\{i: c(i) \in b\} \notin \mathscr{U}$ that $c \notin {}^*b$, and so $c \in {}^*a - {}^*b$. For the converse reverse the steps.

DEFINITION 3.3. *An entity a of the $^*L$-structure $\hat{R}^I$ is called internal whenever there exists a natural number $n \geqq 0$ such that $a \in {}^*R_n$. An internal entity a is called a standard entity whenever there exists an entity $b \in \hat{R}$ such that $a = {}^*b$. All entities which are not internal are called external.*

The set $\bigcup_{n \geqq 0} {}^*R_n$ of all internal entities is called the ultrapower of $\hat{R}$ with respect to the ultrafilter $\mathscr{U}$ and will be denoted by $^*(\hat{R})$.

The $\mathscr{U}$-ultrapower of $\hat{R}$ is usually denoted by $\mathscr{U}$-prod $\hat{R}$ but we shall not employ this notation in this paper.

Observe that the mapping $a \to {}^*a$ of $\hat{R}$ into $\hat{R}^I$ imbeds $\hat{R}$ into the substructure $^*(\hat{R})$ of $\hat{R}^I$.

The notion of rank extends immediately to the internal entities. An internal entity $a \in {}^*(\hat{R})$ is said to be of rank n ($n \geqq 1$) whenever $a \in {}^*R_n - {}^*R_{n+1}$; and the entities of $^*R = {}^*R_0$ are said to be of rank 0. The entities of rank 0 are also referred to as the **individuals** of $^*(\hat{R})$. Again, by means of this definition the empty set $^*\varnothing$ has rank 1. The rank of an internal entity can be further specified. If a is nonempty and internal, then $a \in {}^*R_p$ for some $p \geqq 0$, and so, by Definition 3.1, we have that $U = \{i: a(i) \in R_p\} \in \mathscr{U}$. Then $\bigcup_{k=0}^{p} \{i: \text{rank}\, s(i) = k\} = U \in \mathscr{U}$ implies, using the fact that $\mathscr{U}$ is an ultrafilter, that there exists exactly one index n such that $0 \leqq n \leqq p$ and $U_1 = \{i: \text{rank}\, s(i) = n\} \in \mathscr{U}$. Then for all $i \in U_1 \in \mathscr{U}$ we have $a(i) \in R_n - R_{n-1}$, and so $a \in {}^*(R_n - R_{n-1}) = {}^*R_n - {}^*R_{n-1}$ (Lemma 3.2(vi)), that is, rank $a = n$.

If $a = {}^*b$, $b \in \hat{R}$, is a standard entity of $^*(\hat{R})$, then its rank remains unchanged.

At this point it seems natural to ask the question whether there are internal entities which are not standard. Fortunately, the answer to this question is affirmative and as we shall see in the following theorem it is a consequence of the hypothesis that the ultrafilter $\mathscr{U}$ is δ-incomplete, a hypothesis which we have not used so far.

THEOREM 3.5. *There exist internal entities which are not standard. In fact, if $a \in \hat{R}$ is an entity which has infinitely many elements, then there exists an entity $b \in {}^*a$ such that b is not standard.*

Proof. Since a is an infinite set there exists a sequence $\{b_n: n = 1, 2, \cdots\}$ of elements of a such that $b_n \neq b_m$ for all $n, m = 1, 2, \cdots$ and $n \neq m$. Let b be the mapping of I into a such that $b(i) = b_n$ for all $i \in I_n$ ($n = 1, 2, \cdots$). Then $b \in {}^*a$ but b is not equal to any standard element of $^*(\hat{R})$, and the proof is complete.

The internal entities, defined to be the elements of the special standard sets $*R_n$, can also be characterized as follows. *An entity a is internal if and only if a is an element of a standard entity.* In order to see this we need only to show that if $a \in *b$, $b \in \hat{R}$, then a is internal. Now from $b \in \hat{R}$ it follows that $b \in R_n$ for some n which implies that $b \subset R_0 \cup R_{n-1}$, and so, by Lemma 3.2(v), $a \in *b \subset *R_0 \cup *R_{n-1}$ implies $a \in *R_0 \cup *R_{n-1}$ which shows that a is internal. In view of Theorem 3.5, we may ask the question, what about the nature of the entities which are elements of internal entities? The answer is that they are internal, as the following theorem shows. The converse, however, is not true. In fact, we shall see later in Section 5 that a set of internal entities need not be internal.

THEOREM 3.6. *If $a \in b \in *R_n$ $(n \geqq 1)$, then $a \in *R_{n-1}$, that is, the elements of an internal entity are internal.*

Proof. From $b \in *R_n$ it follows that $U = \{i : b(i) \subset R_0 \cup R_{n-1}\} = \{i : b(i) \in R_n\} \in \mathscr{U}$, and so for all $i \in U$ we have $a(i) \in R_0 \cup R_{n-1}$. Hence, by Lemma 3.2(v) and Definition 3.1, $a \in *(R_0 \cup R_{n-1}) = *R_0 \cup *R_{n-1}$, and the proof is finished.

As in the case of the L-structure $\hat{R}$ we shall call an $*L$-wff **admissible** whenever all the quantifiers occurring in it are of the form "$(\forall x)[[x \in a] \Rightarrow \cdots]$" and "$(\exists x)[[x \in a] \wedge \cdots]$", where a is a constant denoting an entity of $\hat{R}^I$.

An admissible wff of $*L$ is called internal whenever all the constants occurring in it denote internal entities. An admissible wff of $*L$ is called standard whenever all the constants occurring in it denote standard entities. Thus a standard wff is internal.

The set of all **internal sentences** of $*L$ will be denoted by $*K = *K(*L)$, and the subset of all **internal sentences** which hold in $*(\hat{R})$ will be denoted by $*K_0 = *K_0(*L)$.

If V is an admissible wff of L, then its **$*$-transform** $*V$ is defined to be that **standard wff** of $*L$ which is obtained from V by replacing in V *all* the constants, say, $a_1, \cdots, a_p$, occurring in it, by $*a_1, \cdots, *a_p$ but leaving the variables and bracketing unchanged.

We shall now prove that the $*$-imbedding has the following important property.

THEOREM 3.7. *Let $V = V(x_1, \cdots, x_p)$ be an admissible L-wff with the free variables $x_1, \cdots, x_p$, an let $A = \{(x_1, \cdots, x_p) : (x_1, \cdots, x_p) \in a$ and $V(x_1, \cdots, x_p)\}$, where a is an arbitrary entity of $\hat{R}$. Then $A \in \hat{R}$ and*

$$*A = \{(y_1, \cdots, y_p) : (y_1, \cdots, y_p) \in *a \text{ and } *V(y_1, \cdots, y_p)\}.$$

Proof. That $A \in \hat{R}$ is trivial. If $V = V(x_1, \cdots, x_p, a_1, \cdots, a_q)$ is atomic, that is, V has the form $(x_1, \cdots, x_p, a_1, \cdots, a_{q-1}) \in a_q$ or $(x_1, \cdots, x_{p-1}, a_1, \cdots, a_q) \in x_p$ with possible permutation of the variables, then the result follows immediately from

Definition 3.1. In order to show that the result holds for all wff V of L without quantifiers we have to show that if it holds for two such wff V and W, then it also holds for $[V \wedge W]$ and $[\neg V]$. As is well known this will take care of all the logical connectives. Assume that ${}^*A = \{(x_1, \cdots, x_p) \colon (x_1, \cdots, x_p) \in {}^*a \text{ and } {}^*V(x_1, \cdots, x_p)\}$, then we have to show that

$${}^*B = \{(x_1, \cdots, x_p) \colon (x_1, \cdots, x_p) \in {}^*a \text{ and } \neg\, {}^*V(x_1, \cdots, x_p)\},$$

where $B = a - A$. Since, by Lemma 3.2(vi), ${}^*B = {}^*a - {}^*A$ the result follows. Assume now that $V = V(x_1, \cdots, x_p, y_1, \cdots, y_q)$ and $W = W(x_1, \cdots, x_p, z_1, \cdots, z_r)$ be two L-wff without quantifiers for which the result holds, and let

$$A = \{(x_1, \cdots, x_p, y_1, \cdots, y_q, z_1, \cdots, z_r) \colon (x_1, \cdots, x_p, y_1, \cdots, y_q, z_1, \cdots, z_r) \in a \text{ and } [V \wedge W]\}.$$

Then $A = \{(x_1, \cdots, z_r) \colon (x_1, \cdots, z_r) \in a \text{ and } V\} \cap \{(x_1, \cdots, z_r) \colon (x_1, \cdots, z_r) \in a \text{ and } W\}$ implies, by Lemma 3.2(v), that ${}^*A = {}^*\{\cdots\} \cap {}^*\{\cdots\} = \{(x_1, \cdots, z_r) \colon (x_1, \cdots, z_r) \in {}^*a$ and $[V \wedge W]\}$, and so the result holds for all wff without quantifiers.

For admissible wff with quantifiers we shall use induction on the number n of quantifiers. For $n = 0$ the result was shown above. Assume now that the result holds for all admissible wff with less than or equal n quantifiers. Let V be an admissible wff with $(n + 1)$-quantifiers which is written in its prenex normal form $(qx_{n+1}) \cdots (qx_1) W(x_1, \cdots, x_{n+1}, y_1, \cdots, y_q)$, where W has no quantifiers and $y_1, \cdots, y_q$ are the free variables occurring in V. Without loss of generality we may assume that (qx_{n+1}) is the existential quantifier $(\exists x_{n+1})$ otherwise we consider not V. Let b denote the domain of $(\exists x_{n+1})$. Then since V is admissible, $b \in \hat{R}$. Let

$$B = \{((y_1, \cdots, y_p), x_{n+1}) \colon ((y_1, \cdots, y_p), x_{n+1}) \in a \times b \text{ and } (qx_n) \cdots (qx_1) W\},$$

where $a \in \hat{R}$. Then, by the induction hypothesis and Lemma 3.2(v), we obtain that

$${}^*B = \{((y_1, \cdots, y_p), x_{n+1}) \colon ((y_1, \cdots, y_p), x_{n+1}) \in {}^*a \times {}^*b \text{ and } (qx_n) \cdots (qx_1) {}^*W\}.$$

The domain of the binary relation B is the set

$$\begin{aligned} A = \; & \{(y_1, \cdots, y_p) \colon (y_1, \cdots, y_p) \in a \text{ and} \\ & (\exists x_{n+1})(x_{n+1} \in b \wedge (qx_n) \cdots (qx_1) W)\} \\ = \; & \{(y_1, \cdots, y_p) \colon (y_1, \cdots, y_p) \in a \text{ and } V(y_1, \cdots, y_p)\}. \end{aligned}$$

The domain of the binary relation ${}^*B$ is, however, the set

$$\begin{aligned} & \{(y_1, \cdots, y_p) \colon (y_1, \cdots, y_p) \in {}^*a \text{ and } (\exists x_{n+1})(x_{n+1} \in {}^*b \wedge (qx_n) \cdots (qx_1) {}^*W)\} \\ & \quad = \{(y_1, \cdots, y_p) \colon (y_1, \cdots, y_p) \in {}^*a \text{ and } {}^*V\}. \end{aligned}$$

Then, by Lemma 3.2(vii), we obtain the desired result that

$$*A = \{(y_1, \cdots, y_p): (y_1, \cdots, y_p) \in *a \text{ and } *V\},$$

and the proof is finished.

We are now in a position to prove the *Fundamental Theorem* about ultrapowers which we shall refer to throughout the rest of the paper by F.T.

THEOREM 3.8. *$*(\hat{R})$ is a higher order nonstandard model of $\hat{R}$, that is, an admissible sentence V of $K(L)$ holds in $\hat{R}$ if and only if $*V$ holds in $*(\hat{R})$, and $\hat{R}$ is properly imbedded in $*(\hat{R})$.*

Proof. Theorem 3.5 tells us that the imbedding $a \to *a$ of $\hat{R}$ into $*(\hat{R})$ is proper. We have to show that if $V \in K(L)$, then $V \in K_0$ if and only if $*V \in *K_0$. If V has no quantifiers, then it follows immediately from Definition 3.1. Assume that $V \in K$ has the prenex normal form $V = (qx_n) \cdots (qx_1)W$, where W has no quantifiers. There is no loss in generality to assume that (qx_n) is the existential quantifier $(\exists x_n)$. Then $V \in K_0(L)$ is equivalent to "the set $A = \{x_n: x_n \in a \text{ and } (qx_{n-1}) \cdots (qx_1)W\} \neq \varnothing$, where a is the domain of $(\exists x_n)$. Then, by Theorem 3.7 and Lemma 3.2(i), we see that $A \neq \varnothing$ is equivalent to $*A = \{x_n: x_n \in *a \text{ and } (qx_{n-1}) \cdots (qx_1)*W\} \neq *\varnothing$ which itself is equivalent to $*V \in *K_0$, and the proof is finished.

An important aspect of the method of nonstandard analysis is to use the F.T. repeatedly to transform the true statements of $\hat{R}$ into true statements about the internal entities of $*(\hat{R})$. To illustrate this we shall give a number of examples dealing with the set theory of $\hat{R}$.

EXAMPLES 3.9. (i). The individuals of $\hat{R}$ are the "urelements" of the set theory of $\hat{R}$ in the sense that although they are different from the empty set $\varnothing$ there are no entities of $\hat{R}$ which are elements of individuals. This true statement can be expressed by the following infinite list of sentences of K_0.

$$(\forall x)(\forall y)[x \in R] \wedge [y \in R_n] \Rightarrow [\neg y \in x], \; n = 0, 1, 2, \cdots.$$

From the F.T. we conclude that $*K_0$ contains the following list of sentences

$$(\forall x)(\forall y)[x \in *R] \wedge [y \in *R_n] \Rightarrow [\neg y \in x], \; n = 0, 1, 2, \cdots$$

In words, *there are no internal entities which are elements of the individuals of $*(\hat{R})$.*

(ii) One of the axioms of set theory states that the union of the elements of a set is a set. For the set theory of $\hat{R}$ this means that K_0 contains the following infinite list of sentences.

$$(\forall z)[z \in R_n] \Rightarrow (\exists y)[y \in R_n] \wedge (\forall x)[x \in R_n] \Rightarrow [[x \in y]$$
$$\Leftrightarrow (\exists u)[u \in R_n] \wedge [u \in z] \wedge [x \in u]].$$

Thus from the F.T. we have the following result: *The union of the elements of an internal entity is an internal entity.*

(iii) The power set axiom of set theory states that for every set there exists a set whose elements are the subsets of this set. Thus K_0 contains the following infinite list of sentences.

$$(\forall x)[x \in R_n] \Rightarrow (\exists y)[y \in R_{n+1}] \wedge (\forall z)[z \in R_n] \Rightarrow [[z \in y] \Leftrightarrow [z \subset x]], \; n = 1,2,\cdots.$$

Then the F.T. implies that *the set of all internal entities which are subsets of an internal entity is an internal entity.*

(iv) Lemma 2.1(v) states that the domain and range of every entity of $\hat{R}$ which is a binary relation is an entity of $\hat{R}$. This again can be expressed by an infinite list of sentences of K_0.

$$(\forall b)[b \in B_n] \Rightarrow (\exists z)[z \in R_n] \wedge (\forall x)[x \in R_n] \Rightarrow [[x \in z] \Leftrightarrow (\exists y)[y \in R_n] \wedge [(x, y) \in b]]$$

$(n = 3,4,\cdots)$, where B_n denotes the entity of all binary relations of rank $\leqq n$. The F.T. then implies that *the domain and range of any internal binary relation is internal.*

Another remark which is of importance is that if $b \in \hat{R}$ is a binary relation, then any property which b possesses and which can be expressed by sentences of K_0 also holds for $^*b$. For instance, if b is an order relation or function or equivalence relation, then $^*b$ is an order relation or function or equivalence relation. If, however, $b \in R$ wellorders its domain, then $^*b$ wellorders its domain in the sense that every *nonempty internal* subset of the domain of $^*b$ has a first element.

(v) From the axioms of set theory it follows that the image of a set under a binary relation is a set. Thus in $\hat{R}$ the following statement holds. If $b \in \hat{R}$ is a binary relation and $a \in \hat{R}$, then $\{y: (\exists x)(x \in a \wedge (x,y) \in b)\} \in \hat{R}$. We leave it now to the reader to show that this statement can be expressed by sentences of K_0. The F.T. tells us that the following results holds.

The image of an internal entity under an internal binary relation is internal.

In Theorem 3.6 we have shown that the entities of R^I which are elements of an internal entity are internal, and we remarked that a set of internal entities need not be internal (see Section 5). One of the problems in nonstandard analysis is to decide whether certain sets of internal entities are internal or not. As we shall see in the subsequent sections, one of the methods used to decide such a question involves F.T., by showing that the set in question violates a certain property which it should possess, according to the F.T., if it had been internal. Another useful and helpful result in this respect is the following theorem.

THEOREM 3.10. *Let $V = V(x_1, \cdots, x_n)$ be an internal wff with the free variables $x_1, \cdots, x_n$, and let $a \in {}^*(\hat{R})$ be an internal entity. Then the set $\{(x_1, \cdots, x_n) : (x_1, \cdots, x_n) \in a$ and $V(x_1, \cdots, x_n)\}$ is internal.*

Proof. If V has no quantifiers, that is, $V = V(x_1, \cdots, x_n, a_1, \cdots, a_p)$, where $a_1, \cdots, a_p$ are the constants occurring in V which by hypothesis, denote internal entitles. Since a is internal, it follows immediately that the mapping $i \to E(i) = \{(x_1, \cdots, x_n): (x_1, \cdots, x_n) \in a(i)$ and $V(x_1, \cdots, x_n, a_1(i), \cdots, a_p(i))\}$ is a mapping of T into R_n for some n, and so determines an internal entity which we shall denote by E. Then it is easy to see that $E = \{(x_1, \cdots, x_n): (x_1, \cdots, x_n) \in a$ and $V\}$. This proves the result for internal wff without quantifiers. For general internal wff we shall use again induction on the number of quantifiers. Thus assume that the theorem holds for all internal wff with $\leqq n$ quantifiers. Let $V = (qx_{n+1}) \cdots (qx_1)W$ be an internal wff with the free variables $y_1, \cdot, y_p$. There is no loss in generality to assume that $(qx_{n+1}) = (\exists x_{n+1})$ with domain $b \in {}^*(\hat{R})$. Since b is internal it follows from the induction hypothesis that the binary relation

$$B = \{((y_1, \cdots, y_p), x_{n+1}): ((y_1, \cdots, y_p), x_{n+1}) \in a \times b \text{ and } (qx_n) \cdots (qx_1)W(y_1, \cdots, y_p, x_{n+1})\}$$

is internal, and so, by Example 3.9(iv), its domain

$$\{(y_1, \cdots, y_p): (y_1, \cdots, y_p) \in a \text{ and } (\exists x_{n+1})(qx_n) \cdots (qx_1)W\}$$

is internal, and the proof is finished.

4. The nonstandard real number system $^*R$. The set $^*R$ of individuals of the $\mathscr{U}$-ultrapower $^*(\hat{R})$ of the superstructure $\hat{R}$, where $\mathscr{U}$ is a δ-incomplete ultrafilter, has according to the F.T. the same properties as R as far as they can be expressed by sentences of K_0.

Since R is a totally ordered field and since it is easy to see that this can be expressed by sentences of K_0 it follows that $^*R$ is a totally ordered field. The imbedding $a \to {}^*a$ of R into $^*R$ imbeds R into a subfield of $^*R$. In order to simplify our notation we shall denote the extensions of the algebraic operation and order when passing from R to $^*R$ by the same symbols. Thus $a + b = c$ in $^*R$ means in terms of $\mathscr{U}$ that $\{i: a(i) + b(i) = c(i)\} \in \mathscr{U}$, and similarly for subtraction and multiplication. Furthermore, $a \leqq b$ in $^*R$ means $\{i: a(i) \leqq b(i)\} \in \mathscr{U}$. As an illustration the statement that the order relation "$\leqq$" totally orders R can be expressed by the following sentence of K_0

$$(\forall x)(\forall y)[x \in R \wedge y \in R] \Rightarrow [x < y] \vee [x = y] \vee [x > y],$$

and so, as already mentioned above it follows from the F.T. that the extension of the order relation to $^*R$ totally orders R.

The unit element $e \in {}^*R$ has the property that for all $0 \neq r \in R$, $^*r(^*r)^{-1} = e$, and so $e = {}^*1$, where 1 denotes the real number one.

The reader will appreciate that we shall simplify our notation further by no longer using the *-notation to denote the **standard individuals** of $^*R$. Thus we shall from now on identify R with the subfield of the standard numbers of $^*R$, and we shall feel free to write $R \subset {}^*R$.

The absolute value $|r|$ of a real number $r \in R$ defined by $|r| = r$ whenever $r > 0$ and $|r| = -r$ whenever $r < 0$ can be considered to be a mapping of R into $R^+ = \{r: r \in R \text{ and } r \geqq 0\}$ the set of all nonnegative real numbers. The constant of L denoting this mapping extends by passing from $\hat{R}$ to $^*(\hat{R})$ to a mapping $^*|\cdot|$ of $^*R$ into $^*(R^+)$ which according to the F.T. has the property that $^*|a| = a$ for all $^*R \ni a \geqq 0$ and $^*|a| = -a$ for all $^*R \ni a < 0$. Also in this case we shall drop the *-notation and write $|a|$ to denote the absolute value of a real number $a \in {}^*R$. Similarly, we shall write $\max(a, b)$ and $\min(a, b)$, $a, b \in {}^*R$, for the extensions $^*\max(\,,\,)$ and $^*\min(\,,\,)$ of the mappings $\max(r, s)$ and $\min(r, s)$ of $R \times R$ into R respectively.

This liberalization of the notation and some additional notation later on will help a great deal to simplify the mechanics of the subject and can hardly be expected to cause confusion.

Let the constant S denote a subset of R. Then on passing to $^*(\hat{R})$, $^*S$ denotes a subset of $^*R$ which is a standard entity and which by the F.T. has the same properties as S as far as they can be expressed by sentences of K_0. More precisely the substructure $^*(\hat{S})$ of $^*(\hat{R})$, where $\hat{S}$ denotes the superstructure defined by S, is an ultrapower nonstandard model of $\hat{S}$. On the basis of Lemma 3.2(iii) and the present notation, we feel free to write $S \subset {}^*S$. Furthermore, by Lemma 3.2(v), $S = {}^*S$ *if and only if* S *is a finite set.*

If the constant N denotes the set of natural numbers of R, that is, $N = \{1, 2, \cdots\}$, then the standard entity $^*N$ denotes a set of numbers of $^*R$ which again has the same properties as N as far as they can be expressed by sentences of K_0. More precisely, $^*(\hat{N})$ is an ultrapower higher order nonstandard model of arithmetic.

From Theorem 3.5 it follows that $^*R$ is a proper extension of R, and so, according to a result from algebra to the effect that every Archimedean field is isomorphic to a subfield of R, we conclude that $^*R$ is non-Archimedean. But $^*R$ has the same properties as R and R is Archimedean. Let us now examine this apparent paradox. The fact that R is Archimedean can be expressed by the following sentence of K_0:

$$(\forall x)[x \in R] \Rightarrow (\forall n)[n \in N] \Rightarrow [[nx \leqq 1] \Leftrightarrow [x \leqq 0]],$$

and so, by the F.T., the following statement holds for $^*R$.

$$(\forall x)[x \in {}^*R] \Rightarrow (\forall n)[n \in {}^*N] \Rightarrow [[nx \leqq 1] \Leftrightarrow [x \leqq 0]],$$

that is, with the proper interpretation of the constants, $^*R$ is Archimedean with respect to $^*N$. It is not Archimedean in the sense of the metalanguage, that is, if $0 < a \in {}^*R$, then there exists a natural number n in the metalanguage such that $a + \cdots + a > 1$, n-times $+$.

Up till now we have only considered some properties of R and their extensions which can be formulated in a lower order language, that is, sentences in which quantification is over numbers only. Let us now examine a few of the higher order type properties of R. One of the important higher order properties which R possesses and which we have already referred to in the beginning of Section 3 is the so-called **Dedekind completeness property** of R which states that *every nonempty subset of R which is bounded above has a least upper bound.* This statement about R can easily be expressed by a sentence of K_0 which will contain a universal quantifier ranging over subsets of R. Then it follows from the F.T. that $^*R$ satisfies a Dedekind completeness property of the following kind.

(4.1) *Every nonempty internal subset of* $^*R$ *which is bounded above has a least upper bound.*

Since $^*(\hat{N})$ is a higher order nonstandard model of arithmetic, it follows that under the appropriate interpretation of the F.T. the model $^*(\hat{N})$ satisfies all the axioms of Peano. For instance, the principle of induction stating that every nonempty set of natural numbers has a first element, being a higher order property of $\hat{N}$, has to be interpreted in $^*(\hat{N})$ in the following sense.

(4.2) *Every nonempty internal subset of* $^*N$ *has a first element.*

From Theorem 3.5 it also follows that $^*N - N \neq \emptyset$. More precisely, we shall now show that there *exists a natural number* $\omega \in {}^*N$ *such that* $|r| < \omega$ *for all* $r \in R$. Indeed, if $\omega(i) = n$ for all $i \in I_n$ $(n = 1, 2, \cdots)$, where $\{I_n\}$ denotes the partition of I such that $I_n \notin \mathscr{U}$ for all $n = 1, 2, \cdots$, then ω is a mapping of I into N with the property that for all $0 < r \in N$ the set $\{i : \omega(i) < r\} \notin \mathscr{U}$, and so $\omega \in {}^*N$ and $|r| < \omega$ for all $r \in R$. This proves on the basis that $\mathscr{U}$ is δ-incomplete that $^*N$ contains a number which is larger than any positive real number, that is a number which could be called infinitely large. The reader will find it easy now to appreciate the following definition and facts about $^*R$.

DEFINITION 4.3. *A real number* $a \in {}^*R$ *is called finite whenever there exists a standard real number* $0 < r \in R$ *such that* $|a| < r$. *A real number* $a \in {}^*R$ *which is not finite will be called infinite.*

A real number $a \in {}^*R$ *is called an infinitesimal or infinitely small whenever* $|a| < r$ *for all* $0 < r \in R$.

The set of all finite real numbers of $^*R$ will be denoted by M_0 and the set of all infinitesimals by M_1.

Observe that $R \subset M_0$, $M_1 \subset M_0$ and $R \cap M_1 = \{0\}$, that is, 0 ("null") being regarded also as an infinitesimal is the *only* standard infinitesimal.

A real number $a \in {}^*R$ is infinite if and only if $|a| > r$ for all $0 < r \in R$. Thus the natural number ω defined above is infinite. Its reciprocal, however, is an infinitesimal. More generally, a real number $0 \neq a \in {}^*R$ is an infinitesimal if and only if its reciprocal $1/a$ is infinite.

The finite natural numbers are determined in the following theorem.

THEOREM 4.4. *A natural number $n \in {}^*N$ is finite if and only if n is a standard natural number. In symbols,* ${}^*N \cap M_0 = N$.

Proof. It is obvious that $N \subset M_0$. If $n \in {}^*N$ is finite, then there exists a standard real number $0 < r \in R$ such that $n < r$. However, K_0 contains the sentence

$$(\forall x)[x \in N] \Rightarrow [x \leqq r] \Leftrightarrow [x = 1] \vee [x = 2] \vee \cdots \vee [x = p],$$

where r and p are constants and $p = [r]$ is the integral part of r. Thus by the F.T. we obtain that $n = 1$ or $n = 2$ or $\cdots$ or $n = [r]$, and the proof is complete.

From Theorem 4.4 it follows that the set of all infinitely large natural numbers is given by ${}^*N - N$. It is not uncustomary to denote infinitely large natural numbers by lower case greek letters, such as ω, with or without subscripts.

The mapping $r \to [r]$ of R_+ into the set $N \cup \{0\}$, where $[r]$ denotes the largest nonnegative integer less than or equal to r, extends on passing from $\hat{R}$ to ${}^*(\hat{R})$ to a mapping ${}^*[\cdot]$ of ${}^*(R^+)$ into ${}^*N \cup \{0\}$. From the F.T. it follows that for all $0 \leqq a \in {}^*R$, ${}^*[a]$ is the largest nonnegative integer $\leqq a$. Also in this case we shall drop the *-notation and simply write $[a]$ for the integral part of a.

We shall now turn to a discussion of the properties of the finite numbers of ${}^*R$.

It is easy to see that M_0 is a subring of ${}^*R$, and in fact is an integral domain, that is, M_0 has no divisors of zero. The set of infinitesimals constitutes a subring of M_0 with the property that if $h \in M_1$ and $a \in M_0$, then $ah \in M_1$, that is M_1 is an ideal in M_0. In fact, it is easy to see that M_1 is a maximal ideal. Indeed, observe that if $a \in M_0$ and $a \notin M_1$, then there exist positive real numbers r_1, $r_2 \in R$ such that $0 < r_1 < |a| < r_2$, and so $1/a \in M_0$ shows that any ideal which properly contains M_1 must contain the unit element 1 of M_0 and so is all of M_0.

If $a, b \in {}^*R$ and $a - b$ is infinitesimal, then we shall say that b is **infinitely close** to a and we write $a =_1 b$.

Consider the quotient ring M_0/M_1. Then since M_1 is a maximal ideal in M_0, the quotient ring M_0/M_1 is a field. We claim it is isomorphic to the field of standard real numbers. The precise result and details are the subject of the following important theorem.

THEOREM 4.5. *The quotient ring M_0/M_1 is order isomorphic to the field R of the standard real numbers.*

Proof. First observe that if A is an equivalence class in M_0 modulo M_1, then A cannot contain two different standard real numbers r_1 and r_2. Indeed, in that case $|r_1 - r_2| =_1 0$, and so $r_1 \neq r_2$ implies by Definition 4.4 that $|r_1 - r_2| < |r_1 - r_2|$ and a contradiction is obtained. This shows that R is a subfield of M_0/M_1. To complete the proof we have to show that to every $a \in M_0$ there corresponds a standard real number r, which is then unique, such that $a - r =_1 0$. To this end, observe that if $a \in M_0$, then the sets $D = \{r : r \in R \text{ and } r \leqq a\}$ and $D' = R - D$ define a Dedekind cut (D, D') in R. Let $r \in R$ be the real number in R which determines the same cut (D, D'). Then we shall show that $a =_1 r$. If not, then by Definition 4.4 there exists a positive real number $0 < \varepsilon \in R$ such that $|a - r| \geqq \varepsilon$. If $a > r$, then $|a - r| \geqq \varepsilon$ implies that $r + \varepsilon/2 < a$, and contradicts the fact that a and r determine the same cut. Similarly, if $r > a$, then $r - \varepsilon/2 > a$ gives rise to the same contradiction. Thus M_0/M_1 is order isomorphic to R and the proof is finished.

The unique ring and order isomorphism of M_0 onto R with kernel M_1 plays a very important role in the theory of infinitely small and infinitely large numbers. We shall firmly establish it in the following definition.

DEFINITION 4.6. *The ring and order homomorphism of M_0 onto R_0 with kernel M_1 will be called the standard part homomorphism and will be denoted by st.*

In the next theorem, we shall summarize the basic properties of the homomorphism st for later reference.

THEOREM 4.7. (i) $st(a + b) = st(a) + st(b)$, $st(ab) = st(a)st(b)$ *and* $st(a - b) = st(a) - st(b)$ *for all* $a, b \in M_0$.

(ii) *If* $a, b \in M_0$, *then* $a \leqq b$ *implies* $st(a) \leqq st(b)$.

(iii) $st(|a|) = |st(a)|$, $st(\max(a, b)) = \max(st(a), st(b))$ *and* $st(\min(a, b)) = \min(st(a), st(b))$ *for all* $a, b \in M_0$.

(iv) $st(a) = 0$ *if and only if* $a \in M_1$.

(v) *For all standard* $r \in R$ *we have* $st(r) = r$.

(vi) *If* $a \in M_0$ *and* $st(a) \geqq 0$, *then* $|a| =_1 st(a)$.

(vii) *For all* $a, b \in M_0$ *we have* $a =_1 b$ *if and only if* $st(a) = st(b)$.

It is now customary to call the equivalence classes of M_0 with respect to M_1 the **monads** of the standard numbers determined by them. The monads are denoted by $\mu(r)$, $r \in R$. Thus, in particular, $\mu(0) = M_1$.

We shall conclude this section with a number of remarks which are of interest in themselves.

REMARKS. (i) (*The standard part operation defined as a limit*). The standard part operation "st" can also be defined as follows. If $a \in M_0$, then, by Definition 4.4 and Definition 3.1, there is a set $U \in \mathscr{U}$ and a positive standard real number $0 < r \in R$ such that $i \in U$ implies $|a(i)| < r$. Hence, the image of the ultrafilter $\mathscr{U}$ under the mapping $i \to a(i)$ of I into R is a basis of a bounded ultrafilter of subsets of R, and so, by the local compactness of R, it converges to a unique real number r. A simple observation shows that $r = st(a)$. Thus, $st(a) = \lim_{\mathscr{U}} a$ for all $a \in M_0$.

(ii) (*A nonstandard construction of the real number systems*). The proof of Theorem 4.5 suggests immediately the following alternative construction of the real number system. Let the constant Q of L denote the field of rational numbers. Then $*(\hat{Q})$ is a higher order nonstandard model of the superstructure Q. Thus the set of individuals $*Q \subset *R$ is a subfield of $*R$ which has the same properties as Q as far as they can be expressed by sentences of K_0. From Theorem 3.5 we know that

$^*Q \neq Q$, and in fact $^*Q$ contains an element which is larger than any standard real number. It is an easy and interesting exercise for the reader to transform the properties of Q to $^*Q$. We shall show here only that $^*Q$ can be used to define the real number system. To this end, we single out the rationals of $^*Q$ which are finite, that is, $q \in {}^*Q$ is finite whenever $|q| <$ some positive standard rational number. The set of all finite rationals will be denoted by Q_0. Observe that $Q_0 = {}^*Q \cap M_0$. A rational $q \in {}^*Q$ is called infinitesimal whenever $|q|$ is smaller than all positive standard rationals. The set of all infinitely small rationals will be denoted by Q_1. Thus $Q_1 = {}^*Q \cap M_1$. Then it is easy to see that Q_1 is a maximal ideal in the integral domain Q_0. Thus the quotient ring Q_0/Q_1 is ring and order isomorphic to a field. The proof of Theorem 4.5 shows us, however, that this field is isomorphic to the field of Dedekind cuts of Q, and so, by definition, Q_0/Q_1 is isomorphic to the real number system.

(iii) (*The nonstandard complex number system*). Within the framework of axiomatic set theory the complex number system C may be regarded as a subtheory of the theory of the superstructure $\widehat{R \times R}$ determined by $R \times R$. The algebraic operations of addition and multiplication are denoted by constants which correspond to certain six-place relations; and so $^*(\widehat{R \times R})$ may be looked upon as a higher order non-standard model of the complex number system.

It is advisable also in this case to employ the familiar notation $z = x + iy$ for complex numbers, where now $x, y \in {}^*R$ and $i^2 = -1$. The set $^*C = {}^*R \times {}^*R$ of the extended complex number system has of course the same properties as C, and so is, in particular, a field. If $z_1 = x_1 + iy_1$ and $z_2 = x_2 + iy_2$, then also in $^*C$ we have

$$z_1 + z_2 = x_1 + x_2 + i(y_1 + y_2) \text{ and } z_1 z_2 = (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + x_2 y_1).$$

Furthermore, $z = x + iy$, then x is called the real part of z and y is called the imaginary part of z. A complex number $z = x + iy$ is finite whenever x and y are finite, otherwise it is infinite. If x and y are both infinitely small, then $z = x + iy$ is called an infinitely small complex number. From Theorem 4.6 it follows that every finite complex number is infinitely close to a unique standard complex number. For further details concerning nonstandard complex function theory we refer the reader to [**8**] and [**10**].

5. Definitions and properties of some external entities. We pointed out that the converse of Theorem 3.6 need not hold, that is, a set of internal entities need not be internal. In the preceding section we introduced a number of sets of individuals, namely, the set of all infinitely large natural numbers $^*N - N$, the set of finite numbers M_0, the set of infinitesimals M_1, and the monads $\mu(r)$, $r \in R$. It is now natural to ask the question whether these sets are internal or not? To decide this we shall use the following procedure. We assume the set in question is internal and then show that it violates a property which it should have possessed on the basis of the assumption that it is internal and the F.T. The details are contained in the following theorem.

THEOREM 5.1. *The nonempty sets* $^*N - N$, M_0, M_1, $\mu(r)(r \in R)$, *and the set of infinitely large real numbers* $^*R_\infty = {}^*R - M_0$ *are all external.*

Proof. Assume that $^*N - N$ is internal. Then since $^*N - N \neq \emptyset$ (Theorem 3.5) we have by (4.2) that $^*N - N$ has a first element, say, ω_0. But the set of infinitely large natural numbers does not have a first element. Indeed, if $\omega \in {}^*N - N$, then $k + 1 < \omega$ for all $k \in N$ implies that $\omega - 1 \in {}^*N - N$, and so $\omega_0 - 1 < \omega_0$ shows that $^*N - N$ has no first element. Thus $^*N - N$ is external.

Assume that the set M_1 is internal. Since $M_1 \neq \emptyset$ and $h \in M_1$ implies $|h| < 1$ it follows from (4.1) that M_1 has a least upperbound, say, a_0. From $0 \in M$, it follows that $a_0 \geqq 0$. Furthermore, $a_0 \neq M_1$ since M_1 contains elements other than 0. But then $a_0/2$ is also a least upper bound of M_1 and a contradiction is obtained, and so M_1 is external.

Similarly on the basis of (4.1) we can show that M_0 is external. We leave it to the reader as an exercise.

If $^*R_\infty = {}^*R - M_0$ is internal, then also $M_0 = {}^*R - {}^*R_\infty$ is internal, and a contradiction is obtained. Thus $^*R_\infty$ is external.

Since the translation mappings of $^*R$ are internal (check this) it follows immediately from $\mu(0)$ is external that $\mu(r) = \mu(0) + r(r \in R)$ is external. This completes the proof.

REMARKS. (i) If $D \subset M_1$ is internal and nonempty, then according to (4.1) it has a least upper bound. The above proof shows that this least upper bound is an infinitesimal. Similarly, the least upper bound of a nonempty internal set of finite numbers is finite. The greatest lower bound of a nonempty internal set of infinite numbers is of course infinite.

(ii) The standard part operation is a mapping of M_0 onto R. It, is, however, *not an internal mapping*. Indeed, if it were internal, then according to Example 3.9(v) its domain M_0 would have to be internal which contradicts the preceding theorem, and we conclude that the standard part operation is an external operation.

Let $A \in \hat{R}$ be infinite. Then according to Theorem 3.5 the set of all the nonstandard entities of $^*A$ is not empty. More precisely, we have the following result.

THEOREM 5.2. *If* $A \in \hat{R}$, *then the set* $^*A - \{^*a : a \in A\}$ *of all the nonstandard elements of* $^*A$ *is either empty or external, and in the latter case the set* $\{^*a : a \in A\}$ *is also external.*

Proof. If $A \in \hat{R}$, then $^*A - \{^*a : a \in A\} = \emptyset$ if and only if A is finite. (Theorem 3.5 and Lemma 3.2(v)). Assume therefore that A is infinite. Then there is a one-to-one mapping f of a subset of A onto the set $N = \{1, 2, \cdots\}$. If $B = {}^*A - \{^*a : a \in A\}$ is internal, then $B \cap \text{dom}(^*f)$ is internal also (Theorem 3.10). Hence, by Example 3.9(v), we have that $^*N - N = {}^*f(B \cap \text{dom}\, ^*f)$ is internal which contradicts Theorem 5.1 and the proof is finished.

Although the preceding theorem shows that the set of nonstandard elements of the extension of an infinite set of $\hat{R}$ is external there are plenty of internal sets whose elements are all internal entities which are not standard. Indeed, if $\omega \in {}^*N - N$, then the set $\{\omega\}$ is internal but its element is not a standard entity. More generally any *finite* set of internal entities which are not standard is internal. This statement can be generalized as follows. We begin with a definition.

DEFINITION 5.3. *A set D of internal entities of $^*(\hat{R})$ is called *-finite whenever there exists a natural number $\omega \in {}^*N - N$ and an internal one-to-one mapping of D onto the internal set $\{1, 2, \cdots, \omega\}$. In that case, we shall say that the internal cardinal of D is ω or shortly that D has ω-elements.*

If D is *-finite, then it is clear that its external cardinal is at least as big as $\aleph_0$. Concerning *-finite sets we have the following result.

THEOREM 5.4. *Every *-finite set of internal entities is internal. A *-finite set of real numbers has a largest and a smallest element.*

Proof. Since, by Example 3.9(iv), the domain of an internal function is internal it follows immediately from Definition 5.3 that a *-finite set is internal.

If D is a *-finite set of real numbers, then from the sentence of K_0 stating that every finite set of real numbers of R has a largest and a smallest element it follows from the F.T. that every *-finite set of real numbers in $^*R$ has a largest and a smallest element. This completes the proof.

REMARK. If the internal set D is *-finite, then it must contain at least one internal entity which is not standard, and so at least externally infinitely many of those. This can be shown as follows. If the entities of D are all standard, then there exists a standard set $A \in \hat{R}$ such that $D = \{^*a : a \in A\}$ (use first part of Theorem 3.6). Since the cardinal of A is infinite it follows from Theorem 5.2 that the set $D = \{^*a : a \in A\}$ is external, and so a contradiction is obtained.

6. The theory of limits. As a first example and also for later reference we shall illustrate what kind of effect the theory of infinitely small and infinitely large numbers has on the theory of limits.

We recall that a (standard) sequence $\{s_n : n = 1, 2, \cdots\}$ can be regarded as a mapping of N into R, and so being a subset of $N \times R$ it is an entity of $\hat{R}$ which we shall denote for obvious reasons by s. On passing from $\hat{R}$ to $^*(\hat{R})$ the entity s extends to an entity $^*s$ which according to the F.T. and Lemma 3.2(vii) is a mapping of $^*N$ into $^*R$. Furthermore, for all finite $n \in N$ we have $^*s_n = s_n$ as follows from the fact that $^*(\text{ran}\, s) = \text{ran}\, ^*s$ and the convention of dropping the *-notation for individuals. The standard sequence $*s$ in $*(\hat{R})$ has the same properties as the sequence s as far as they can be expressed by sentences of K_0. With this fundamental principle in mind we shall now prove the following theorems.

THEOREM 6.1. *A sequence* $\{s_n: n = 1,2,\cdots\}$ *in* R *is bounded if and only if* $*s_\omega$ *is finite for all infinitely large natural numbers* $\omega \in *N - N$.

Proof. This follows immediately from the remark following Theorem 5.1 to the effect that the least upper bound of an internal set of finite numbers is finite. Hence, if $(\text{ran } *s) \subset M_0$, then $|*s_n| \leqq a$ for all $n \in *N$ and some $a \in M_0$, that is, $|s_n| \leqq st(a)$ for all $n \in N$, and the proof is finished.

In the classical sense a sequence $\{s_n: n = 1,2,\cdots\}$ is said to be convergent with limit s if and only if

$$(*) \quad (\forall \varepsilon)[0 < \varepsilon \in R] \Rightarrow (\exists x)[x \in N] \wedge (\forall y)[y \in N \wedge x \leqq y] \Rightarrow [|s_y - s| < \varepsilon].$$

In nonstandard analysis this is expressed in a more intuitive fashion as follows.

THEOREM 6.2. *Let* $\{s_n: n = 1,2,\cdots\}$ *be a sequence of numbers of* R, *and let* $s \in R$. *Then* $\lim_{n\to\infty} s_n = s$ *if and only if* $*s_\omega =_1 s$ *for all* $\omega \in *N - N$.

Proof. Assume first that $\lim_{n\to\infty} s_n = s$. Then from the sentence (*) of K_0 the following is a sentence of K_0.

$$(\forall x)[x \in N \wedge x > n] \Rightarrow |s_x - s| < \varepsilon, \quad \text{where } \varepsilon > 0 \text{ and } n \in N$$

are constants. Thus the following $*L$-sentence holds.

$$(\forall x)[x \in *N \wedge x > n] \Rightarrow |*s_x - s| < \varepsilon.$$

In particular, for all $\omega \in *N - N$ we have that $|*s_\omega - s| < \varepsilon$. The latter statement holds, however, for all $\varepsilon > 0$, that is, $*s_\omega =_1 s$ for all $\omega \in *N - N$.

In order to see that the condition is sufficient we observe that if ε is a constant denoting a positive number of R, the following sentence holds in $*(\hat{R})$.

$$(\exists y)[y \in *N] \wedge (\forall x)[x \in *N \wedge y < x] \Rightarrow |*s_x - s| < \varepsilon.$$

Indeed, we need to take for y only an infinitely large natural number. Observe now that this sentence is the $*$-transform of the sentence

$$(\exists y)[y \in N] \vee (\forall x)[x \in N \wedge y < x] \Rightarrow |s_x - s| < \varepsilon,$$

and so by F.T. holds in $\hat{R}$. This means that there is an index $n_0 \in N$ such that $|s_n - s| < \varepsilon$ for all $n > n_0$. Since this holds for all $\varepsilon > 0$ we obtain that $\lim_{n\to\infty} s_n = s$ and the proof is finished.

The condition $*s_\omega =_1 s$ for all $\omega \in *N - N$ is equivalent to $st(*s_\omega) = s$ for all $\omega \in *N - N$.

Theorem 6.2 also tells us immediately that if the limit exists it is unique. Furthermore, Theorem 6.1 shows that every convergent sequence is bounded.

EXAMPLES 6.3. (i) If one wishes to show that $\lim_{n\to\infty}\sqrt[n]{n} = 1$, then set $s_n = \sqrt[n]{n} - 1$ $(n = 1,2,\cdots)$ and observe that

$$n = (1 + s_n)^n = \sum_{k=0}^{n} \binom{n}{k} s_n^k \geqq \binom{n}{2} s_n^2 \quad \text{for all } n = 1,2,\cdots.$$

Hence, $0 \leqq s_n \leqq \sqrt{2/(n-1)}$ for all $n > 1$, and so, also $0 \leqq {}^*s_m \leqq \sqrt{2/(m-1)}$ for all $1 < m \in {}^*N$. In particular if $\omega \in {}^*N - N$ is infinitely large, then $0 \leqq {}^*s_\omega \leqq \sqrt{2/(\omega-1)}$ and $\sqrt{2/(\omega-1)} \in M_1$ implies that ${}^*s_\omega = {}_1 0$, and so, by Theorem 6.2, $\lim_{n\to\infty} s_n = 0$, and the proof is finished.

(ii) (*Algebra of limits*). The usual rules for calculating with limits are now easily obtained. For, if $\lim s_n = s$ and $\lim t_n = t$, then ${}^*(s+t)_\omega = {}^*s_\omega + {}^*t_\omega =_1 s + t$ for all $\omega \in {}^*N - N$ and so $\lim_{n\to\infty}(s_n + t_n) = s + t$. Similarly, $st({}^*(st)_\omega) = st({}^*s_\omega {}^*t_\omega) = st({}^*s_\omega)st({}^*t_\omega) = st$ for all $\omega \in {}^*N - N$ shows that $\lim_{n\to\infty} s_n t_n = st$. In the same way one shows that if $t \neq 0$, then $\lim_{n\to\infty} s_n/t_n = s/t$.

(iii) It is well known that if $\lim_{n\to\infty} s_n = s$, then

$$\lim_{n\to\infty} \frac{s_1 + \cdots + s_n}{n} = s.$$

The proof of this result in nonstandard analysis reads as follows. From $\lim_{n\to\infty} s_n = s$ it follows first of all that for some $0 < r \in R$, $|{}^*s_n - s| < r$ for all $n \in {}^*N$ (Theorem 6.1) and ${}^*s_n - s =_1 0$ for all $n \in {}^*N - N$. Now let $\omega \in {}^*N - N$ and let $\omega_0 = [\sqrt{\omega}]$. Then the following simple estimation gives the required result:

$$\left|\frac{{}^*s_1 + \cdots + {}^*s_\omega}{\omega} - s\right| \leqq \frac{|{}^*s_1 - s| + \cdots + |{}^*s_{\omega_0} - s|}{\omega_0} \frac{1}{\sqrt{\omega}}$$

$$+ \frac{|{}^*s_{\omega_0+1} - s| + \cdots + |{}^*s_\omega - s|}{\omega}$$

$$\leqq \frac{r}{\sqrt{\omega}} + \frac{(\omega - \omega_0)}{\omega} \cdot \max(|{}^*s_n - s| : \omega_0 < n \leqq \omega) =_1 0,$$

by Theorem 5.4.

Cauchy's criterion for convergence in analysis takes on the following form.

THEOREM 6.4. *A sequence $\{s_n : n = 1,2,\cdots\}$ of real numbers of R is convergent if and only if ${}^*s_\omega =_1 {}^*s_{\omega'}$ for all $\omega, \omega' \in {}^*N - N$.*

Proof. From Cauchy's criterion $|s_n - s_m| < \varepsilon$ for all n, m sufficiently large it follows as in the proof of Theorem 6.2 that the condition is necessary. In order to prove that the condition is sufficient we have only to show in view of Theorem 6.2 that $^*s_\omega$ is finite for all $\omega \in {}^*N - N$. To this end, assume that there exists an infinitely large natural number $\omega_0 \in {}^*N - N$ such that $^*s_{\omega_0}$ is infinite. We define now the following set $A = \{n: {}^*N \text{ and } |{}^*s_{\omega_0} - {}^*s_n| < 1\}$ of natural numbers. From Theorem 3.10 it follows that A is internal. Furthermore, by hypothesis $^*N - N \subset A$. If $n \in N$ is finite, then $|{}^*s_{\omega_0}| \leqq |{}^*s_{\omega_0} - {}^*s_n| + |{}^*s_n| \in M_0$ shows that $n \notin A$, and so $A = {}^*N - N$. Contradicting the fact that $^*N - N$ is not internal (Theorem 5.1), and so $^*s_\omega$ is finite for all $\omega \in {}^*N - N$, and the proof is finished.

REMARK. The above proof shows also that an infinite sequence $\{s_n: n = 1, 2, \cdots\}$ is bounded if and only if $^*s_\omega - {}^*s_{\omega'}$ is finite for all $\omega, \omega' \in {}^*N - N$.

The following result of A. Robinson (see [9]) concerning internal sequences will be used in Section 9.

THEOREM 6.5. *Let $\{a_n: n \in {}^*N\}$ be an internal sequence of real numbers such that a_n is infinitely small for all finite $n \in N$. Then there exists an infinitely large natural number $\omega \in {}^*N - N$ such that $a_n =_1 0$ for all $n \leqq \omega$.*

Proof. Consider the internal sequence $\{na_n: n \in {}^*N\}$ and let $A = \{n: n \in {}^*N$ and $\forall k[k \in {}^*N \wedge k \leqq n] \Rightarrow k|a_k| \leqq 1\}$. Then, by Theorem 3.10, A is internal. Since the hypothesis $a_n =_1 0$ for all finite $n \in N$ implies $na_n =_1 0$ for all finite $n \in N$ it follows that $N \subset A$. Since, by Theorem 5.2, the set N is external and since A is internal, $A - N \neq \varnothing$. Hence, there exists an infinitely large natural number $\omega \in A$. Then for all infinitely large $n \leqq \omega$ the condition $n|a_n| \leqq 1$ implies that $0 \leqq |a_n| \leqq 1/n =_1 0$, and the proof is finished.

7. Sequences that are asymptotically linear. A standard sequence of real numbers $\{s_n: n = 1, 2, \cdots\}$ is called **asymptotically linear** whenever there exists a real constant $\sigma \in R$ such that $s_n = n\sigma + o(n)$, $n \in N$.

A now classical result of Pólya and Szegö states if a *sequence* $\{s_n: n = 1, 2, \cdots\}$ *is almost additive, that is, there exists a constant s such that* $|s_{n+m} - s_n - s_m| \leqq s$ *for all* $n, m = 1, 2, \cdots$, *then* $\{s_n\}$ *is asymptotically linear.*

As another illustration of the use of infinitely small and infinitely large numbers we shall prove here in a nonstandard fashion the following slightly more general result.

THEOREM 7.1. *Let $\{s_n: n = 1, 2, \cdots\}$ be a standard sequence of real numbers for which there exist constants p, s such that $0 < p < 1$ and $|s_{n+m} - s_n - s_m| \leqq s(n^p + m^p)$ for all $n, m = 1, 2, \cdots$. Then there exists a constant $\sigma \in R$ such that $|s_n - n\sigma| \leqq sn^p/(1 - 2^{p-1})$ for all $n = 1, 2, \cdots$. In particular, $\{s_n\}$ is asymptotically linear.*

Proof. From the hypothesis it follows immediately that for all $k = 1, 2, \cdots$ and for all $n = 1, 2, \cdots$ we have

(7.2) $$\left|\frac{s_{2^k n}}{2^k} - s_n\right| \leqq sn^p \frac{1 - 2^{(p-1)k}}{1 - 2^{p-1}}.$$

Then it follows from the F.T. that, by passing to $*(\hat{R})$, (7.2) holds for all $k, n \in {}^*N$. In particular, if $k = \omega \in {}^*N - N$ is infinitely large, then

(7.3) $$\left|\frac{{}^*s_{2^\omega n}}{2^\omega} - {}^*s_n\right| \leqq sn^p \frac{1 - 2^{(p-1)\omega}}{1 - 2^{p-1}} \quad \text{for all } n \in {}^*N.$$

Since $0 < p < 1$, and ω is infinitely large, $2^{(p-1)\omega}$ is infinitely small, and so for all finite n we see that ${}^*s_{2^\omega n}/2^\omega$ is finite. Let

$$a_n = \frac{{}^*s_{2^\omega n}}{2^\omega}, \quad n \in {}^*N.$$

Then the internal sequence $\{a_n : n \in {}^*N\}$ satisfies, by hypothesis, the condition that $|a_{n+m} - a_n - a_m| \leqq s2^{(p-1)\omega}(n^p + m^p)$ for all $n, m \in {}^*N$. Since a_n is finite for all finite $n \in N$ we obtain by setting $t_n = st(a_n)$, $n \in N$, that $|t_{n+m} - t_n - t_m| = 0$, that is, $t_n = nt_1 = n\sigma$, $n = 1, 2, \cdots$. Finally, if we take standard parts in (7.3) keeping n finite we obtain that $|s_n - n\sigma| \leqq sn^p/(1 - 2^{p-1})$, and the proof is finished.

8. Continuity and differentiability. Let f be a real-valued function of a real variable which is defined on an open interval $a < x < b$ of R. On passing to $*(\hat{R})$ the function f extends to a function ${}^*f$ whose domain of definition is the open interval $a < x < b$, $x \in {}^*R$ and with values in ${}^*R$. Furthermore, we have to keep in mind that the F.T. implies that ${}^*f$ satisfies in $*(\hat{R})$ all the properties of f as far as they can be expressed by sentences of K_0.

For instance, if for some $a < x_0 < b$, $\lim_{x \to x_0} f(x) = l$ holds, then the following sentence belongs to K_0.

$$(\forall \varepsilon)[0 < \varepsilon \in R] \Rightarrow (\exists \delta)[0 < \delta \in R] \wedge (\forall x)[x \in R \wedge 0 < |x - x_0| < \delta]$$
$$\Rightarrow [|f(x) - l| < \varepsilon].$$

Using the same methods as in the proof of Theorem 6.1 we obtain immediately the following result.

THEOREM 8.1. *$\lim_{x \to x_0} f(x) = l$ if and only if ${}^*f(x_0 + h) =_1 l$ for all $0 \neq h \in M_1$. In particular, f is continuous at x_0 if and only if ${}^*f(x_0 + h) =_1 f(x_0)$ for all $h \in M_1$, that is, equivalently, $st({}^*f(a) = f(st(a))$ for all $a \in {}^*R$ such that $st(a) = x_0$.*

The derivative of f at x_0 exists if and only if

$$\lim_{y \to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

exists. Thus, by Theorem 8.1, f is differentiable at x_0 if and only if there exists a constant $l \in R$ such that

$$\frac{{}^*f(x_0 + h) - {}^*f(x_0)}{h} =_1 l$$

for all $0 \neq h \in M_1$. As we might have expected the derivative of a differentiable function is the standard part of the quotient of infinitesimals

$$\frac{\Delta f}{\Delta x} = \frac{{}^*f(x + \Delta x) - f(x)}{\Delta x},$$

where $\Delta x \neq 0$ denotes an infinitesimal.

If f is differentiable at x_0, then f is continuous at x_0, Indeed, from ${}^*f(x_0 + h) - f(x_0) =_1 h\, f'(x_0)$ for all $0 \neq h \in M_1$ it follows, using $hl =_1 0$, that ${}^*f(x_0 + h) - f(x_0) =_1 0$ for all $h \in M_1$.

A real function f defined on an arbitrary interval is uniformly continuous whenever for every $0 < \varepsilon \in R$ there exists a constant $0 < \delta \in R$ such that $|f(x) - f(y)| < \varepsilon$ for all $x, y \in \operatorname{dom} f$ and $|x - y| < \delta$. In passing to ${}^*(\hat{R})$ we obtain immediately the following criterion for uniform continuity.

THEOREM 8.2. *Let f be a real function of a real variable. Then f is uniformly continuous if and only if ${}^*f(a) =_1 {}^*f(b)$ for all $a, b \in \operatorname{dom} {}^*f$ and $a =_1 b$.*

From the above results the following famous theorem of Heine can now be obtained immediately.

THEOREM 8.3. (*Heine*). *Let f be a real function of a real variable defined on the bounded and closed interval $x_1 \leqq x \leqq x_2$, $x_1, x_2 \in R$. If f is continuous, then f is uniformly continuous.*

Proof. Let $a, b \in {}^*R$ satisfy $x_1 \leqq a, b \leqq x_2$ and $a =_1 b$. Then $a, b \in M_0$ and $x = st(a) = st(b)$ satisfies $x_1 \leqq x \leqq x_2$. Since f is continuous we have, by Theorem 8.1, that ${}^*f(a) =_1 f(x) =_1 {}^*f(b)$, and so ${}^*f(a) =_1 {}^*f(b)$, that is, by Theorem 8.2, f is uniformly continuous, and the proof is finished.

For a more detailed account of the theory of real functions of a real variable in non-standard analysis we refer the reader to [3] and [**8**].

9. Euler's product for the sine function. On passing from $\hat{R}$ to ${}^*(\hat{R})$, the elementary functions of the calculus such as the functions $\log x$, e^x, $\sin x$, $\cos x$, and so on, extend to functions defined in ${}^*R$ and which have the same properties as their standard counterpart as far as they can be expressed by sentences of K_0. In order to simplify the notation we shall not use the *-notation to denote the extensions of the elementary functions. Thus, for instance, in place of writing ${}^*(\sin)(x)$, $x \in {}^*R$, we simply write $\sin x$, $x \in {}^*R$. For a discussion of the elementary functions of ${}^*(\hat{R})$ we refer the reader to [3].

One of the many beautiful formulas which were discovered by Euler is the so-called product formula for the sine-function. By this we mean the following formula.

$$\sin z = z \prod_{k=1}^{\infty} \left(1 - \frac{z^2}{k^2\pi^2}\right), \; z \text{ is complex.} \tag{9.1}$$

Nowadays this representation for the sine function belongs to that part of function theory that studies the behavior of entire functions whenever its zeros are given. There one learns that the quotient of the functions on the left and right-hand side of (9.1) is a function of the form e^f, where f is entire. The whole problem is then to determine f, and, in fact, to show that $f = 0$ in the case of the sine function. There are many proofs known for this result. Some of the proofs are even elementary. But all of these proofs are somewhat artificial in the sense that they rely on some analytical trick. It is therefore not without interest to examine how Euler proved his formula. As far as the author knows, Euler's original proof is contained in his book *Introductio ad Analysin Infinitorum* which appeared in 1748. It runs as follows. The mathematical expressions such as "infinitely large" and "infinitely close" which occur in it are Euler's and not the author's.

For infinitely large values of n we have

$$2\sinh x = \left(1 + \frac{x}{n}\right)^n - \left(1 - \frac{x}{n}\right)^n . \tag{9.2}$$

We are now going to factorize the polynomial occurring on the right-hand side of (9.2), by observing that $a^n - b^n = (a-b)(a-\varepsilon_1 b)\cdots(a-\varepsilon_{n-1}b)$, where $1, \varepsilon_1, \cdots, \varepsilon_{n-1}$ are the nth roots of unity. Now combine the pairs of complex conjugate roots to obtain the real quadratic polynomials

$$\left(a - b\exp\left(\frac{2k\pi i}{n}\right)\right)\left(a - b\exp\left(-\frac{2k\pi i}{n}\right)\right) = a^2 + b^2 - 2ab\cos\frac{2k\pi}{n},$$

and so, since $a^2 + b^2 = 2 + (2x^2/n^2)$ and $2ab = 2 - (2x^2/n^2)$, we obtain

$$2\left(1 - \cos\frac{2k\pi}{n}\right) + 2\left(1 + \cos\frac{2k\pi}{n}\right)\frac{x^2}{n^2} = 4\sin^2\frac{k\pi}{n}\left(1 + \frac{x^2}{n^2\tan^2(k\pi/n)}\right).$$

It follows that the polynomial is divisible by x and for all values of $k = 1, 2\cdots$, by $1 + \{x^2/n^2\tan^2(k\pi/n)\}$. Since n is infinitely large this factor is infinitely close to $1 + (x^2/k^2\pi^2)$. Furthermore, it is easy to see that the coefficient of x is equal to 2, and so we obtain that

$$\sinh x = x \prod_{k=1}^{\infty} \left(1 + \frac{x^2}{k^2\pi^2}\right). \tag{9.3}$$

Finally, by applying it for $x = iz$, the required formula is obtained.

The reader who has read this far will agree with the author that Euler's proof is a typical example of the way infinitely large and infinitely small numbers were used with great success in the early stages of the development of the calculus. It is, however, no wonder that the inability to give the theory of infinitely large and infinitely small numbers a firm foundation led to the unacceptibility of such proofs. Of course, it is no problem at all with the methods of nonstandard analysis to make Euler's proof precise.

From Theorem 6.1 it follows that for all standard $x \in R$ and for all infinitely large natural numbers $\omega \in {}^*N - N$ we have

$$(9.4) \qquad 2\sinh x =_1 \left(1+\frac{x}{\omega}\right)^{\omega} - \left(1-\frac{x}{\omega}\right)^{\omega}.$$

Factorizing the polynomial as before leads to the formula.

$$(9.5) \quad \left(1+\frac{a}{m}\right)^{m} - \left(1-\frac{a}{m}\right)^{m} = \frac{4^{m/2}}{m}\left(\sum_{k=1}^{[(m-1)/2]} \sin^2\frac{k\pi}{m}\right) a \sum_{k=1}^{[(m-1)/2]} \left(1+\frac{a^2}{m^2\tan^2\frac{k\pi}{m}}\right),$$

for all $a \in {}^*R$ and for all $m \in {}^*N$, and where $[(m-1)/2]$ as in Section 4 denotes the largest natural number $\leqq (m-1)/2$. Dividing by $a \neq 0$ and letting $a = 0$ shows that

$$(9.6) \qquad \frac{4^{m/2}}{m} \prod_{k=1}^{[(m-1)/2]} \sin^2\frac{k\pi}{m} = 2 \quad \text{for all } m \in {}^*N.$$

Thus we obtain finally that

$$(9.7) \qquad \left(1+\frac{x}{\omega}\right)^{\omega} - \left(1-\frac{x}{\omega}\right)^{\omega} = 2x \prod_{k=1}^{[(\omega-1)/2]} \left(1+\frac{x^2}{\omega^2\tan^2(k\pi/\omega)}\right),$$

or all $x \in R$ and for all $\omega \in {}^*N - N$.

We shall now prove the following lemma.

LEMMA 9.8. *If $x \in R$ is standard, then for all infinitely large $\omega \in {}^*N - N$ we have*

$$st\left(\prod_{k=1}^{[(\omega-1)/2]} \left(1+\frac{x^2}{\omega^2\tan^2(k\pi/\omega)}\right)\right) = \prod_{k=1}^{\infty}\left(1+\frac{x^2}{k^2\pi^2}\right).$$

Proof. Since for all $x \in R$, the infinite product $\prod_{k=1}^{\infty}(1+x^2/k^2\pi^2)$ is convergent it follows from Theorem 6.1 that

$$(9.9) \qquad \prod_{k=1}^{\infty}\left(1+\frac{x^2}{k^2\pi^2}\right) =_1 \prod_{k=1}^{[(\omega-1)/2]}\left(1+\frac{x^2}{k^2\pi^2}\right)$$

for all $x \in R$ and for all $\omega \in {}^*N - N$.

Since $\omega^2 \tan^2(k\pi/\omega) \geqq k^2\pi^2$ for all $1 \leqq k \leqq [(\omega-1)/2]$ we obtain that

$$\sum_{k=1}^{[(\omega-1)/2]}\left(\log\left(1+\frac{x^2}{k^2\pi^2}\right)-\log\left(1+\frac{x^2}{\omega^2\tan^2(k\pi/\omega)}\right)\right) \geqq 0, \quad \text{for all } x\in R.$$

From Theorem 3.10 it follows that the following sequence is internal

$$(9.10)\quad \eta_n = \sum_{k=1}^{n}\left(\log\left(1+\frac{x^2}{k^2\pi^2}\right)-\log\left(1+\frac{x^2}{\omega^2\tan^2(k\pi/\omega)}\right)\right),\ n\in {}^*N \text{ and } x\in R.$$

If n is finite, then, by Theorem 8.1, the continuity of the log-function and $n/\omega =_1 0$, it follows that

$$\log\left(1+\frac{x^2}{\omega^2\tan^2(k\pi/\omega)}\right) =_1 \log\left(1+\frac{x^2}{k^2\pi^2}\right),\ x\in R,$$

and so $\eta_n =_1 0$ for all finite $n\in N$. Then it follows from Theorem 6.5 that there exists an infinitely large natural number $\nu \leqq [(\omega-1)/2]$ such that $\eta_n =_1 0$ for all $n \leqq \nu$.

Observing that

$$\log\left(1+\frac{x^2}{\omega^2\tan^2(k\pi/\omega)}\right) \geqq 0 \quad \text{for all } 1\leqq k\leqq[(\omega-1)/2],$$

we obtain that

$$0 \leqq \eta_{[(\omega-1)/2]} \leqq \eta_\nu + \sum_{k=\nu+1}^{[(\omega-1)/2]}\log\left(1+\frac{x^2}{k^2\pi^2}\right) =_1 \sum_{k=\nu+1}^{[(\omega-1)/2]}\log\left(1+\frac{x^2}{k^2\pi^2}\right).$$

From Cauchy's criterion in the form of Theorem 6.4 it follows, however, that $\sum_{k=\nu+1}^{[(\omega-1)/2]}\log(1+(x^2/k^2\pi^2)) =_1 0$ for all $x\in R$, and so we obtain that $\eta_{[(\omega-1)/2]} =_1 0$. Finally, the lemma follows from the continuity of the log-function.

In order to complete the proof observe that from (9.4) and (9.7) it follows that for all standard $x\in R$ we have

$$\sinh x =_1 x \prod_{k=1}^{[(\omega-1)/2]}\left(1+\frac{x^2}{\omega^2\tan^2(k\pi/\omega)}\right),\quad \omega\in {}^*N - N,$$

and so by taking standard parts using Lemma (9.9) we obtain finally that

$$\sinh x = x\prod_{k=1}^{\infty}\left(1+\frac{x^2}{k^2\pi^2}\right) \text{ for all } x\in R.$$

From the latter formula the product formula can be obtained by using the uniqueness theorem for analytic functions. In this connection it is not without interest to remark that a slight extension of the argument presented above will give the result for all complex $z \neq \pm k\pi$, $k = 0, 1, 2, \cdots$. We shall leave it to the reader to verify this.

10. Nonmeasurable functions. In this final section of the present paper we shall present a simple example of a function which is not measurable in the sense of Lebesgue. The construction or rather the definition of the example will be based on the theory of infinitely large and infinitely small numbers.

In a previous paper [5], the present author already defined such a function. It involved some nontrivial properties of the sine-function in $^*R$. We shall follow here another idea.

Let $\omega \in {}^*N - N$ be an infinitely large natural number. Then by Theorem 3.10 the following function is internal.

$$\phi(x) = [2^{\omega}x] - 2[2^{\omega-1}x], \quad x \in {}^*R. \tag{10.1}$$

The internal function ϕ is obviously periodic modulo one. It can also be defined as the ωth coefficient of the dyadic expansion of $x - [x]$ $(x \in {}^*R)$, and so it takes on only the values 0 and 1.

By f we shall denote the restriction of ϕ to the set of standard real numbers R of $^*R$. Then the following result holds.

THEOREM 10.2. *The real function* $f(x) = [2^{\omega}x] - 2[2^{\omega-1}x]$, $x \in R$, *is not measurable in the sense of Lebesgue.*

Proof. Observe that f has the following properties. (i) For every (standard) dyadic number d, $0 \leqq d \leqq 1$, $f(d) = 0$. (ii) Every dyadic number d, $0 \leqq d \leqq 1$, is a period of f, that is, $f(x + d) = f(x)$ for all $x \in R$. (iii) For all x, $0 \leqq x \leqq 1$, we have $f(1 - x) = 1 - f(x)$ provided x is not dyadic. (i) and (ii) follow immediately from the fact that since ω is infinitely large, $2^{\omega}d$ and $2^{\omega-1}d$ are natural numbers for all standard dyadic numbers d, $0 \leqq d \leqq 1$. (iii) follows from the fact that $f(x)$ is the ωth coefficient of the dyadic expansion of x in $^*R$. We shall now assume that f is a measurable function. We shall have to use the following well-known result.

(10.3) *A measurable function which has arbitrarily small periods is equal to a constant almost everywhere.*

From the assumption that f is measurable, property (ii) of f, (10.3), and the fact that f takes on only the values 0 and 1 it follows that $f = 0$ a.e. or $f = 1$ a.e.

Let $A = \{x: 0 \leqq x \leqq 1 \text{ and } f(x) < 1/2\}$. Then A is a measurable set, and the characteristic function of A has all the dyadic number as periods, and so, by (10.3), $m(A) = 0$ or $m(A) = 1$, where m denotes the Lebesgue measure. Consider now also the set $B = \{x: 0 \leqq x \leqq 1 \text{ and } f(x) > \frac{1}{2}\}$. Then properties (ii) and (iii) of f imply that if x is not dyadic $0 < x < 1$, then $x \in A$ if and only if $1 - x \in B$. Thus the set A_0 of non-dyadic points of A and the set B_0 of nondyadic points of B are symmetric with respect to the point $1/2$. Since the set of dyadic points is countable its Lebesgue measure is zero and so $m(A) = m(A_0) = m(B_0) = m(B)$. Then $A_0 \cap B_0 = \emptyset$, $m(A_0 \cup B_0) \leqq 1$, $m(A_0) = m(B_0)$ and $m(A_0) = 0$ or $m(A_0) = 1$ imply that $m(A_0) = m(B_0) = 0$. Hence, $f(x) = 1/2$ a.e., which contradicts the fact that f does not take on the value $1/2$. We conclude that f is not measurable in the sense of Lebesgue and the proof is finished.

Work on this paper was also supported in part by Grant No. GP-7691 from the National Science Foundation.

References

1. A. R. Bernstein and A. Robinson, Solution of an invariant subspace problem of K. T. Smith and P. R. Halmos, Pacific J. Math., 16 (1966) 421–431.

2. A. R. Bernstein, Invariant subspaces of polynomially compact operators on a Banach space, Pacific J. Math., 21 (1967) 445–464.

3. W. A. J. Luxemburg, Non-Standard Analysis. Lectures on A. Robinson's theory of infinitesimals and infinitely large numbers, Pasadena (1962) and revised edition (1964).

4. ———, Two applications of the method of construction by ultrapowers, Bull. Amer. Math. Soc., (2) 68 (1962) 416–419.

5. ———, Addendum to, "On the measurability of a function which occurs in a paper by A. C. Zaanen", Proc. Roy. Acad. Sci., Amsterdam, A66 (1963) 587–590.

6. ———, Applications of Model Theory to Algebra, Analysis and Probability Theory. Proceedings of an International Symposium on Nonstandard Analysis, Holt-Rinehart and Winston, New York, 1969.

7. A. Robinson, Non-standard analysis, Proc. Roy. Acad. Sci., Amsterdam, A64 (1961) 432–440.

8. ———, Non-standard Analysis, Studies in Logic and the Foundations of Mathematics, North-Holland, Amsterdam, 1966.

9. ———, On generalized limits and linear functionals, Pacific J. Math., 14 (1964) 269–283.

10. ———, On the theory of normal families, Acta Philos. Fenn., 18 (R. Nevanlinna anniversary volume) (1965) 159–184.

11. ———, On some applications of model theory to algebra and analysis, Rend. Mat. e Appl., 25 (1966) 1–31.

12. ———, A new approach to the theory of algebraic numbers, Atti Accad. Naz. Lincei Rend., (8) 40 (1966) 222–225, 770–774.

13. ———, Nonstandard theory of Dedekind rings, Proc. Roy. Acad. Sci., Amsterdam, A70 (1967) 444–452.

14. ———, Nonstandard arithmetic, Bull. Amer. Math. Soc., 73 (1967) 818–843.

15. R. T. Taylor, Invariant Subspaces in Hilbert and Normed Spaces, Ph. D. Thesis, Calif. Institute of Technology, Pasadena, 1968.