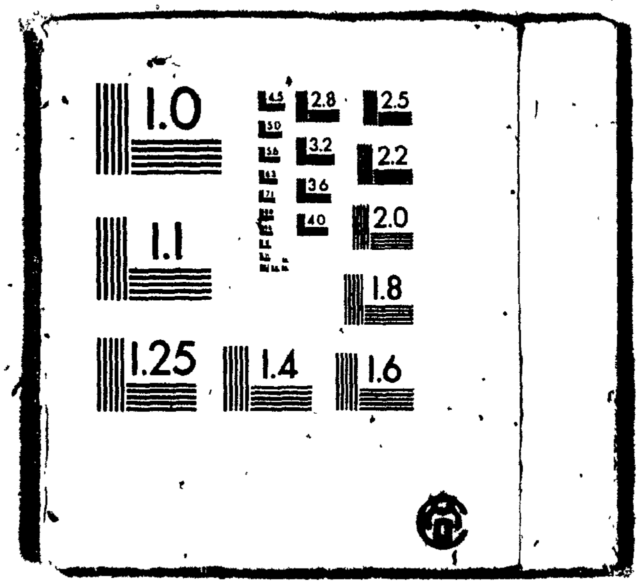
CANADIAN THESES

52652

THESES CANADIENNES

1



0-315-4158-5

52652

NAME OF AUTHOR/*NOM DE L'AUTEUR* ERIKA JENNIFER FARKAS

TITLE OF THESIS/*TITRE DE LA THÈSE* NONSTANDARD ASPECTS OF ANALYSIS

UNIVERSITY/*UNIVERSITÉ* CONCORDIA

DEGREE FOR WHICH THESIS WAS PRESENTED/ *GRADE POUR LEQUEL CETTE THÈSE FUT PRÉSENTÉE* MASTER OF SCIENCE

YEAR THIS DEGREE CONFERRED/*ANNÉE D'OBTENTION DE CE DEGRÉ* 1981

NAME OF SUPERVISOR/*NOM DU DIRECTEUR DE THÈSE* DR. M.E. SZABO

DATED/*DATÉ* June 5, 1981 SIGNED/*SIGNÉ* E. J. Farkas

PERMANENT ADDRESS/*RÉSIDENCE FIXÉ* 6260 NORTHCREST PL. #207
MONTREAL, QUE.
H3S 2T4

NL-91 (11-73)

NONSTANDARD ASPECTS OF ANALYSIS

by

ERIKA JENNIFER FARKAS

A Thesis

in

The Department

of

Mathematics

Presented in Partial Fulfilment of the Requirements
for the degree of Master of Science at
Concordia University
Montreal, Quebec, Canada

June 1981

ABSTRACT

NONSTANDARD ASPECTS OF ANALYSIS

Erika Jennifer Farkas

This thesis has been written in support of the claim that nonstandard proofs in analysis are simpler than standard proofs. We give a careful description of the nonstandard characterizations of a variety of key concepts from calculus, topology, functional analysis and integration theory and demonstrate the logical simplifications achieved in nonstandard proofs by presenting detailed comparisons of well-known proofs from these areas of mathematics.

ACKNOWLEDGEMENTS

I am deeply indebted to my advisor, Dr. M.E. Szabo, for the considerable time he devoted to the preparation of this thesis, and for the advice, encouragement and guidance which I received.

I would also like to thank the Natural Sciences and Engineering Research Council of Canada for the support of this research in the form of a Postgraduate Scholarship.

Look, stranger, at this island now
The leaping light for your delight discovers,
Stand stable here
And silent be,
That through the channels of the ear
May wander like a river
The swaying sound of the sea.

W. H. Auden

CONTENTS

INTRODUCTION

In this thesis we present a variety of nonstandard characterizations of standard concepts of analysis and describe the conceptual and computational simplifications of the proofs of several well-known theorems in terms of these characterizations. It is remarkable how higher order ad hoc set-theoretical constructions in standard proofs can be replaced in a unique way by lower level concepts inspired by analogies with finitist reasoning. An inspection of many of the nonstandard proofs given below shows that they are obtained by simply writing down what ought to be true and then verifying that it is true. In this way, nonstandard proofs are usually direct and avoid arguments involving negations, contradictions, and quantification over iterated power set objects.

A good example of the uniform nature of nonstandard criteria is that for convergent sequences. In the standard case, different sequences may satisfy a given ε condition from different points of their domain onwards and to prove convergence, these points must be identified. In the nonstandard case, all convergent sequences are inspected on the same part of their domain, viz., at their values at the infinite integers.

A closer examination of the nonstandard phenomenon reveals that much of the proof-theoretical burden of analysis has, in a global and uniform way, been shifted to the model theory of analysis, where all higher order auxiliary notions required in proofs have been incorporated once and for all into the base structure. The fact that this is possible strikes the writer as quite extraordinary. The existential richness of the nonstandard world requires not unexpectedly a sophisticated description. Thus the introduction of such a seemingly strange object like a non-principal ultrafilter on a given infinite set cannot be avoided, unless it is replaced by something no less immediate. However, the properties required for the nonstandard proofs of theorems in elementary calculus, for example, can and have been isolated and an axiomatic description has been given (cf. KEISLER [1976] and RICHTER [1981]). Except for the questions of content and consistency, the subject is therefore independent of its set-theoretical description. The result is a richer proof theory for elementary analysis. A higher order syntax has also been developed in NELSON [1977].

The directness and simplicity of nonstandard proofs has led to a variety of new theorems in standard functional analysis, differential equations, probability theory, physics, and economics. For details we refer the reader to ANDERSON [1976-77], BERNSTEIN-WATTENBERG [1969], BERNSTEIN [1973], HENSON [1972], LIGHTSTONE-ROBINSON [1975], LOEB [1972, 1974, 1975, 1979], RICHTER [1981], ROBINSON [1966], among others. The first major example of a new standard theorem was the Bernstein-Robinson nonstandard proof of the fact that a polynomially compact operator on ℓ^2 has a closed proper invariant subspace.

In our presentation, we have emphasized the rôle of nonstandard analysis as a tool for the discovery and proof of theorems in standard analysis. Our reason for this approach is partly historical and partly conceptual. When asked "What is analysis?", the standard analyst can point to a unique (up to isomorphism) second order structure $\mathbb{R}$ on which his subject is based. This gives a certain canonical nature to his subject. The nonstandard enrichments $^*\mathbb{R}$ of $\mathbb{R}$ have no such uniqueness property since they depend on the ultrafilters used in the construction of the nonstandard world. Only extremely powerful set-theoretical "saturation" assumptions induce a certain uniqueness on $^*\mathbb{R}$. Thus the subject matter of nonstandard analysis, as a field of study in its own right, is context dependent. On the other hand, the complacency of standard analysts about the definiteness of the world they live in must surely be shaken by the *proven* inability of set theory to settle even the cardinality of their universe, and by the fact that this uniqueness depends crucially on such classical principles of *propositional* logic as the law of the excluded middle and the law of contraposition. In the context of the weaker intuitionistic logic, non-isomorphic real number systems based on Dedekind cuts in one case, and on Cauchy sequences in the other, have recently been constructed (cf. GOLDBLATT [1979]).

The thesis is essentially self-contained and is divided into four chapters according to the set-theoretical complexity of the nonstandard concepts employed: The first chapter contains only proofs involving the use of infinitely small and infinitely large numbers. The second chapter is devoted to the particularly beautiful and geometrically appealing nonstandard descriptions of the topological separation properties and of compactness. We illustrate with a variety of applications how the replacement of neighbourhood *systems* by canonical *neighbourhoods* leads to direct and straightforward proofs of sometimes fairly complicated mathematical facts. The third chapter exploits the existence of global elements of concurrent relations in the nonstandard universes to replace various limit arguments, i.e., the use of "approximations from below" by arguments involving a single hyperfinite "approximation from above". The method yields an extremely transparent proof of the spectral theorem for compact Hermitian operators on Hilbert spaces, for example. The fourth chapter deals with the general use of elements of the nonstandard universe in several mathematical settings. We give an elementary nonstandard proof of the Arzelà-Ascoli Theorem and develop the elementary theory of Riemann and Lebesque integration.

Throughout the thesis we have attempted to juxtapose standard and nonstandard methods in order to contrast the logical complexity of the proof theory employed in each case. It is particularly instructive to survey the comparisons of standard and nonstandard proofs given in each chapter and to note the quantificational simplification achieved by working over the existentially richer nonstandard universe. Much of the metatheory has in this way become part of the base theory. In this way existential assertions can often also be proved directly without the use of arguments by contradiction.

Most of the notations used in this thesis are self-expanatory or are defined explicitly in one of the Appendices. We have simplified the presentation by using the letter ρ as a metavariable for metrics. This ambiguity makes many of our proofs more readable. Other similar simplifications have been made and are clear from the context. We have also included the general description of standard and nonstandard *universes* in terms of "superstructures" and "ultraproducts". Further details may be found in DAVIS [1977]. The symbol $\square$ denotes either the end of a proof or the end of a discussion.

CHAPTER 1

HYPERREAL NUMBERS

Continuous functions

1.1. *Definition* (Standard). The function f is *continuous* at $x_0 \in X$ if and only if $(\forall \varepsilon>0)(\exists \delta>0)(\forall x \in X)(\rho(x,x_0) < \delta \Rightarrow \rho(f(x),f(x_0)) < \varepsilon)$.

1.2. *Definition* (Nonstandard). The function f is *continuous* at $x_0 \in X$, if and only if $(\forall x \in {}^*X)(x \approx x_0 \Rightarrow f(x) \approx f(x_0))$.

1.3. *Theorem* (Characterization theorem). Definitions (1.1) and (1.2) are equivalent.

Proof. Let f be continuous in the sense of (1.1). Then for any $\varepsilon \in \mathbb{R}^+$ there exists a $\delta \in \mathbb{R}^+$ such that

$$\hat{S} \models (\forall x \in X)(\rho(x,x_0) < \delta \Rightarrow \rho(f(x),f(x_0)) < \varepsilon) .$$

By the Transfer Principle we have that

$$\tilde{W} \models (\forall x \in {}^*X)(\rho(x,x_0) < \delta \Rightarrow \rho(f(x),f(x_0)) < \varepsilon) .$$

If $x \approx x_0$, then $\rho(x,x_0) < \delta$ since δ is standard, and so it follows that $\rho(f(x),f(x_0)) < \varepsilon$. Since ε was an arbitrary positive real number, $\rho(f(x),f(x_0))$ is less than any positive real and so must be infinitesimal, i.e., $f(x) \approx f(x_0)$.

Conversely, suppose that $x \approx x_0 \Rightarrow f(x) \approx f(x_0)$ for $x \in {}^*X$. Let $\varepsilon \in \mathbb{R}^+$ and let δ be any positive infinitesimal. Then

$$\tilde{W} \models (\exists \delta \in {}^*\mathbb{R}^+)(\forall x \in {}^*X)(\rho(x,x_0) < \delta \Rightarrow \rho(f(x),f(x_0)) < \varepsilon) .$$

By the Transfer Principle, f satisfies Definition (1.1). □

Uniformly continuous functions

1.4. *Definition* (Standard). The function f is *uniformly continuous* on X if and only if $(\forall\varepsilon>0)(\exists\delta>0)(\forall x,x'\in X)(\rho(x,x') < \delta \Rightarrow \rho(f(x),f(x')) \leq \varepsilon)$.

1.5. *Definition* (Nonstandard). The function f is *uniformly continuous* on X if and only if $(\forall x,x'\in {}^*X)(x \approx x' \Rightarrow f(x) \approx f(x'))$.

1.6. *Theorem* (Characterization theorem). Definitions (1.4) and (1.5) are equivalent.

Proof. Let f be uniformly continuous in the sense of (1.4). Then for each real $\varepsilon > 0$ there exists a real $\delta > 0$ such that

$$\hat{S} \models (\forall x\in X)(\forall x'\in X)(\rho(x,x') < \delta \Rightarrow \rho(f(x),f(x')) < \varepsilon) .$$

By the Transfer Principle, we have that

$$\tilde{W} \models (\forall x\in {}^*X)(\forall x'\in {}^*X)(\rho(x,x') < \delta \Rightarrow \rho(f(x),f(x')) < \varepsilon) .$$

Thus $x \approx x'$ implies that $\rho(f(x),f(x'))$ is less than any real $\varepsilon > 0$, so that $f(x) \approx f(x')$.

Conversely, suppose that f is uniformly continuous in the sense of (1.5) and let $\varepsilon \in \mathbb{R}^+$. Then $\rho(f(x),f(x')) < \varepsilon$ if δ is any positive infinitesimal and $x,x' \in {}^*X$ and $\rho(x,x') < \delta$. Hence

$$\tilde{W} \models (\exists\delta\in {}^*\mathbb{R}^+)(\forall x\in {}^*X)(\forall x'\in {}^*X)(\rho(x,x') < \delta \Rightarrow \rho(f(x),f(x')) < \varepsilon) .$$

The result therefore follows by an application of the Transfer Principle. □

Equicontinuous families of functions

1.7. *Definition* (Standard). The family F of functions is *equicontinuous* on X if and only if

$$(\forall \varepsilon > 0)(\exists \delta > 0)(\forall f \in F)(\forall x, x' \in X)(\rho(x, x') < \delta \Rightarrow \rho(f(x), f(x')) < \varepsilon) .$$

1.8. *Definition* (Nonstandard). The family F of functions is *equicontinuous* on X if and only if $(\forall f \in {}^*F)(\forall x, x' \in {}^*X)(x \approx x' \Rightarrow f(x) \approx f(x'))$.

1.9. *Theorem* (Characterization theorem). Definitions (1.7) and (1.8) are equivalent.

Proof. Let F be equicontinuous in the sense of (1.7) and let ε be a positive real number. Then there exists a real $\delta > 0$ such that

$$\hat{S} \models (\forall f \in F)(\forall x, x' \in X)(\rho(x, x') < \delta \Rightarrow \rho(f(x), f(x')) < \varepsilon) .$$

By the Transfer Principle, we have that

$$\tilde{W} \models (\forall f \in {}^*F)(\forall x, x' \in {}^*X)(\rho(x, x') < \delta \Rightarrow \rho(f(x), f(x')) < \varepsilon)$$

Since ε is arbitrary, it follows that $x \approx x' \Rightarrow f(x) \approx f(x')$.

Conversely, let F be equicontinuous in the sense of (1.8) and let $\varepsilon \in \mathbb{R}^+$. Then

$$\tilde{W} \models (\exists \delta \in {}^*\mathbb{R}^+)(\forall f \in {}^*F)(\forall x, x' \in {}^*X)(\rho(x, x') < \delta \Rightarrow \rho(f(x), f(x')) < \varepsilon) .$$

The result therefore follows by an application of the Transfer Principle. ▯

Cauchy sequences

1.10. *Definition* (Standard). A sequence $\langle x_n \rangle_{n \in \mathbb{N}}$ is a *Cauchy sequence* if and only if $(\forall \varepsilon > 0)(\exists n_0 \in \mathbb{N})(m,n > n_0 \Rightarrow \rho(x_n, x_m) < \varepsilon)$.

1.11. *Definition* (Nonstandard). A sequence $\langle x_n \rangle_{n \in \mathbb{N}}$ is a *Cauchy sequence* if and only if $(\forall m,n \in {}^*\mathbb{N} - \mathbb{N})(x_n \approx x_m)$.

1.12. *Theorem* (Characterization theorem). Definitions (1.10) and (1.11) are equivalent.

Proof. Let $\langle x_n \rangle_{n \in \mathbb{N}}$ be a Cauchy sequence in the sense of (1.10) and let ε be a positive real number. Then there is some $n_0 \in \mathbb{N}$ such that

$$\tilde{S} \models (\forall n,m \in \mathbb{N})(n,m > n_0 \Rightarrow \rho(x_n, x_m) < \varepsilon).$$

By the Transfer Principle, we have that

$$\tilde{W} \models (\forall n,m \in {}^*\mathbb{N})(n,m > n_0 \Rightarrow \rho(x_n, x_m) < \varepsilon).$$

A fortiori, $\rho(x_n, x_m) < \varepsilon$ for all $n,m \in {}^*\mathbb{N} - \mathbb{N}$, and since ε was arbitrary it follows that $x_n \approx x_m$.

Conversely, let $x_n \approx x_m$ for all $n,m \in {}^*\mathbb{N} - \mathbb{N}$ and suppose that ε is a positive real number. By choosing for n_0 an infinite integer, we obtain that

$$\tilde{W} \models (\exists n_0 \in {}^*\mathbb{N})(\forall n,m \in {}^*\mathbb{N})(n,m > n_0 \Rightarrow \rho(x_n, x_m) < \varepsilon).$$

The result follows by the Transfer Principle. □

Convergent sequences of constants

1.13. *Definition* (Standard). The sequence $\langle x_n \rangle_{n\in\mathbb{N}}$ *converges* to x if and only if $(\forall\varepsilon>0)(\exists n_0\in\mathbb{N})(n > n_0 \Rightarrow \rho(x_n,x) < \varepsilon)$.

1.14. *Definition* (Nonstandard). The sequence $\langle x_n \rangle_{n\in\mathbb{N}}$ *converges* to x if and only if $(\forall n\in{}^*\mathbb{N}-\mathbb{N})(\rho(x_n,x) \approx 0)$.

1.15. *Theorem* (Characterization theorem). Definitions (1.13) and (1.14) are equivalent.

Proof. Suppose that $\langle x_n \rangle_{n\in\mathbb{N}}$ converges to x in the sense of (1.13). Then for all positive real numbers ε there exists a $n_0(\varepsilon) \in \mathbb{N}$ such that

$$\hat{S} \models (\forall n\in\mathbb{N})(n > n_0 \Rightarrow \rho(x_n,x) < \varepsilon) .$$

By the Transfer Principle, we have that

$$\tilde{W} \models (\forall n\in{}^*\mathbb{N})(n > n_0 \Rightarrow \rho(x_n,x) < \varepsilon) .$$

Since $n_0(\varepsilon)$ is always finite it follows that $\rho(x_n,x) < \varepsilon$ holds for all $n \in {}^*\mathbb{N}-\mathbb{N}$. This statement is true for all $\varepsilon > 0$. We therefore have $\rho(x_n,x) \approx 0$ for all $n \in {}^*\mathbb{N}-\mathbb{N}$.

Conversely, let $\rho(x_n,x) \approx 0$ for all $n \in {}^*\mathbb{N}-\mathbb{N}$ and choose a positive real number ε . Then, a fortiori, $\rho(x_n,x) < \varepsilon$ for all $n \in {}^*\mathbb{N}-\mathbb{N}$. In particular, if n_0 is some fixed infinite integer, we have $\rho(x_n,x) < \varepsilon$ for all $n > n_0$. Hence

$$\tilde{W} \models (\exists n_0\in{}^*\mathbb{N})(\forall n\in{}^*\mathbb{N})(n > n_0 \Rightarrow \rho(x_n,x) < \varepsilon) .$$

The result follows by an application of the Transfer Principle. □

Convergent sequences of functions

1.16. *Definition* (Standard). The sequence $\langle f_n \rangle_{n\in\mathbb{N}}$ of functions *converges* to the function $f : X \to \mathbb{R}$ if and only if

$$(\forall\varepsilon>0)(\forall x\in X)(\exists n_0\in\mathbb{N})(n > n_0 \Rightarrow \rho(f_n(x),f(x)) < \varepsilon) .$$

1.17. *Definition* (Nonstandard). The sequence $\langle f_n \rangle_{n\in\mathbb{N}}$ of functions *converges* to the function $f : X \to \mathbb{R}$ if and only if

$$(\forall x\in X)(\forall n\in{}^*\mathbb{N}-\mathbb{N})(f_n(x) \approx f(x)) .$$

1.18. *Theorem* (Characterization theorem). Definitions (1.16) and (1.17) are equivalent.

Proof. Let $f_n \to f$ on X in the sense of (1.16) and let ε be a positive real number. Then there exists a $n_0 \in \mathbb{N}$ such that

$$\hat{S} \models (\forall n\in\mathbb{N})(n > n_0 \Rightarrow \rho(f_n(x),f(x)) < \varepsilon) .$$

By the Transfer Principle, we have that

$$\tilde{W} \models (\forall n\in{}^*\mathbb{N})(n > n_0 \Rightarrow \rho(f_n(x),f(x)) < \varepsilon) .$$

A fortiori, $\rho(f_n(x),f(x)) < \varepsilon$ holds for all $\varepsilon > 0$, hence $f_n(x) \approx f(x)$ for all $n > n_0$, in particular for all $n \in {}^*\mathbb{N}-\mathbb{N}$.

Conversely, assume that for all $x \in X$ and all $n \in {}^*\mathbb{N}-\mathbb{N}$ it holds that $f_n(x) \approx f(x)$, and let ε be any fixed positive real number. Then

$$\tilde{W} \models (\exists n_0\in{}^*\mathbb{N})(\forall n\in{}^*\mathbb{N})(n > n_0 \Rightarrow \rho(f_n(x),f(x)) < \varepsilon) .$$

The result follows by an application of the Transfer Principle. □

Uniformly convergent sequences of functions

1.19. *Definition* (Standard). The sequence $\langle f_n \rangle_{n\in\mathbb{N}}$ of functions *converges uniformly* to the function $f : X \to \mathbb{R}$ if and only if

$$(\forall\varepsilon>0)(\exists n \in\mathbb{N})(\forall x\in X)(n > n_0 \Rightarrow \rho(f_n(x),f(x)) < \varepsilon) .$$

1.20. *Definition* (Nonstandard). The sequence $\langle f_n \rangle_{n\in\mathbb{N}}$ of functions *converges uniformly* to the function $f : X \to \mathbb{R}$ if and only if

$$(\forall x\in {}^*X)(\forall n \in {}^*\mathbb{N} - \mathbb{N})(f_n(x) \approx f(x)) .$$

1.21. *Theorem* (Characterization theorem). Definitions (1.19) and (1.20) are equivalent.

Proof. Let the sequence $\langle f_n \rangle_{n\in\mathbb{N}}$ converge uniformly to f in the sense of (1.19) and let ε be a positive real number. Then there exists a $n_0 \in \mathbb{N}$ such that

$$\hat{S} \models (\forall n \in\mathbb{N})(\forall x\in X)(n > n_0 \Rightarrow \rho(f_n(x),f(x)) < \varepsilon) .$$

By the Transfer Principle, we have that

$$\tilde{W} \models (\forall n\in {}^*\mathbb{N})(\forall x\in {}^*X)(n > n_0 \Rightarrow \rho(f_n(x),f(x)) < \varepsilon) .$$

Hence we have $f_n(x) \approx f(x)$ for all $n \in {}^*\mathbb{N} - \mathbb{N}$ since for any $\varepsilon > 0$ we always have a finite n_0 such that $n > n_0 \Rightarrow \rho(f_n(x),f(x)) < \varepsilon$.

Conversely, assume that for all $x \in {}^*X$, and for all $n \in {}^*\mathbb{N} - \mathbb{N}$, $f_n(x) \approx f(x)$, and let ε be any fixed positive real number. Then

$$\tilde{W} \models (\exists n_0\in {}^*\mathbb{N})(\forall n\in {}^*\mathbb{N})(\forall x\in {}^*X)(n > n_0 \Rightarrow \rho(f_n(x),f(x) < \varepsilon) .$$

The result follows by an application of the Transfer Principle. □

Complete metric spaces

1.22. *Definition* (Standard). The metric space X is *complete* if and only if every Cauchy sequence in X converges.

1.23. *Definition* (Nonstandard). The metric space X is *complete* if and only if for every remote point $p \in {}^*X$ there exists a real number $r > 0$ such that $\rho(p,q) \geq r$ for all $q \in X$.

1.24. *Theorem* (Characterization theorem). Definition (1.22) and (1.23) are equivalent.

Proof. Suppose that X is complete in the sense of (1.22) and that the conclusion does not hold. Then for some remote $p \in {}^*X$ there is a standard sequence $\langle q_n \rangle_{n \in \mathbb{N}}$ such that $\rho(p,q_n) < 1/n$ for $n \in \mathbb{N}$. Therefore

$$\rho(q_m,q_n) \leq \rho(q_m,p) + \rho(q_n,p) < 1/n + 1/m ,$$

so that the sequence $\langle q_n \rangle_{n \in \mathbb{N}}$ is a Cauchy sequence. By the completeness of X we have that $q \in X$. But then

$$\rho(p,q) \leq \rho(p,q_n) + \rho(q_n,q) < 1/n + \rho(q_n,q)$$

for all $n \in \mathbb{N}$, so that $\rho(p,q) \approx 0$, which contradicts the remoteness of p.

Conversely, let $\langle x_n \rangle_{n \in \mathbb{N}}$ be a Cauchy sequence and let ν be an infinite integer. By elementary properties of Cauchy sequences (cf. DAVIS [1977]) it suffices to show that x_ν is near standard. Suppose

x_ν is remote. By hypothesis, there exists a real $r > 0$ such that $\rho(x_\nu, q) \geq r$ for all $q \in X$. In particular, $\rho(x_\nu, x_n) \geq r$ for all $n \in \mathbb{N}$. Since $< x_n >_{n \in \mathbb{N}}$ is a Cauchy sequence, there is a $n \in \mathbb{N}$ such that

$$\hat{S} \models (\forall n,m \in \mathbb{N})(n,m > n_0 \Rightarrow \rho(x_n, x_m) < r) .$$

By the Transfer Principle, we have that

$$\tilde{W} \models (\forall n,m \in {}^*\mathbb{N})(n,m > n_0 \Rightarrow \rho(x_n, x_m) < r) .$$

Choosing $m = n$ and n any natural number greater than n_0, we get a contradiction. □

Some standard theorems (A comparison of standard and nonstandard proofs)

1.25. *Theorem* (Intermediate value theorem). Let f be a continuous real-valued function on $[a,b]$ and let $f(a) < 0 < f(b)$. Then there exists a $c \in (a,b)$ such that $f(c) = 0$.

Proof (Nonstandard). For each $n \in \mathbb{N}$, we define a sequence $< t(n,i) >_{n \in \mathbb{N}}$ of partitions of the interval $[a,b]$ into n equal parts as follows:

$$(*) \qquad t(n,i) = \begin{cases} a + \frac{b-a}{n} i & \text{if } n \in \mathbb{N} \text{ and } 0 \leq i \leq n \\ 0 & \text{otherwise} \end{cases}$$

This yields a double sequence $t : \mathbb{N} \times \mathbb{N} \to \mathbb{R}$ with a star extension $^*t : {}^*\mathbb{N} \times {}^*\mathbb{N} \to {}^*\mathbb{R}$ whose values are defined by the terms $^*t\lceil <n,i>$ for $n,i \in {}^*\mathbb{N}$. By the Transfer Principle, (*) holds for all $n,i \in {}^*\mathbb{N}$.

Next we choose a $\nu \in {}^*\mathbb{N} - \mathbb{N}$ and consider the set

$$L = \{ i \in {}^*\mathbb{N} \mid f(t(\nu,i)) > 0 , \; i \leq \nu \} .$$

By the Internal Subset Theorem, L is an internal subset of $^*\mathbb{N}$ since it is defined by the formula $(i \in {}^*\mathbb{N}) \wedge (f\lceil(t\lceil<\nu,i>)) > 0)$. Furthermore, L is nonempty since we have assumed that $f(b) > 0$ and therefore $\nu \in L$. By the Nonstandard Induction Principle, L has a least element j. It follows that $f(t(\nu,j)) > 0$. By hypothesis, $f(a) < 0$, so that $j \neq 0$, and therefore $f(t(\nu,j-1)) \leq 0$. Furthermore, $t(\nu,j)$ is finite since $a \leq t(\nu,j) \leq b$.

Let $c = st(t(\nu,j))$, so that $c \approx t(\nu,j)$. By definition, the difference between $t(\nu,j)$ and $t(\nu,j-1)$ is $\frac{b-a}{\nu}$, i.e., infinitesimal, and therefore $t(\nu,j) \approx t(\nu,j-1)$.

By Theorem 1.3, we therefore have $f(t(\nu,j)) \approx f(c)$ and also $f(t(\nu,j-1)) \approx f(c)$. Since $f(c)$ is standard, we can therefore conclude that $f(c) = st(f(t(\nu,j))) = st(f(t(\nu,j-1)))$. By virtue of the chosen inequalities and the Standard Part Theorem, it therefore follows that $f(c) = st(f(t(\nu,j))) = st(f(t(\nu,j-1)))$. Hence we have $f(c) = 0$. □

Comparison with a standard proof (RUDIN [1976]).

Structural difference.

Basic steps of the *nonstandard proof*.

(1) The construction of a sequence of partititons of $[a,b]$.

(2) The construction of an internal subset L of $^*\mathbb{N}$ containing an infinite integer ν.

(3) The verification that the standard part of the value of $^*f$ involving ν is the required point c.

Basic steps of the *standard proof*.

(1) The proof that $[a,b]$ is connected.

(2) The proof that $f([a,b])$ is connected.

(3) The proof that any point between two points of $f([a,b])$ belongs to $f([a,b])$.

Logical difference.

The *nonstandard proof* is direct, contains no negations, involves only the first order quantifier $(\exists j \in \mathbb{N})$, and uses the Nonstandard Induction Principle and the Transfer Principle. The *standard proof*, on the other hand, involves three nested proofs by contradiction, and contains several negations of the second order quantifier $(\exists U \in p(\mathbb{R}))$.

Mathematical difference.

The *nonstandard proof* uses the Internality and Standard Part Theorems, whereas the *standard proof* uses the fact that the inverse image of an open set under a continuous function is open. □

1.26. *Theorem* (Maximum value theorem). Let f be a continuous real-valued function on $[a,b]$. Then f has a maximum value.

Proof (Nonstandard). Let $t(n,i)$ be defined as in Theorem 1.25. Then

$\hat{S} \models (\forall n \in \mathbb{N})(\exists i \in \mathbb{N})\Phi$, where Φ is the formula

$$(0 \le i \le n) \wedge (\forall j \in \mathbb{N})((0 \le j \le n) \Rightarrow (f\ulcorner(t\ulcorner\langle n,i\rangle) \ge f\ulcorner(t\ulcorner\langle n,j\rangle))),$$

because the finite set $\{ f(t(n,0)), \ldots, f(t(n,n)) \}$ of real numbers has a maximum.

From the Transfer Principle it follows that for any $\nu \in {}^*\mathbb{N} - \mathbb{N}$, there is a i, $0 \le i \le \nu$, such that for all j, with $0 \le j \le \nu$, $f(t(\nu,i)) \ge f(t(\nu,j))$.

Let $c = st(t(\nu,i))$, and let x be any standard element of $[a,b]$. By the Transfer Principle, there is a $j < \nu$ such that $t(\nu,j) \le x \le t(\nu,j+1)$. Since $\frac{b-a}{\nu} \approx 0$, it follows that $x = st(t(\nu,j))$. By the continuity of f,

$$f(c) = st(f(t(\nu,i))) \ge st(f(t(\nu,j))) = f(x) .$$

Therefore, $f(c)$ is a maximum. □

Comparison with a standard proof (SPIVAK [1967]).

Structural difference.

Basic steps of the *nonstandard proof*.

(1) The proof that a hyperfinite set of finite hyperreal numbers has a maximum.

(2) The verfication that the standard part of this maximum is the required value.

Basic steps of the *standard proof*.

(1) The proof that the range of f has a supremum.

(2) The definition of a special continuous function g on [a,b] .

(3) The proof that g is unbounded if the supremum of f is not a value of f ..

Logical difference.

The *nonstandard proof* is direct, contains no negations, involves the first order quantifiers $(\forall n \in \mathbb{N})$, $(\exists n \in \mathbb{N})$, $(\forall n \in {}^*\mathbb{N})$, $(\exists n \in {}^*\mathbb{N})$, and $(\forall n \in {}^*\mathbb{N} - \mathbb{N})$, and uses the Transfer Principle. The *standard proof*, on the other hand, contains the negative statements that the supremum of f is not a functional value, that the function g is not bounded, and involves both the first order quantifiers $(\forall x \in \mathbb{R})$, $(\exists x \in \mathbb{R})$, and the second order quantifier $(\exists U \in p(\mathbb{R}))$ (asserting the existence of a *least* upper bound) . It uses both the law of the excluded middle and the contrapositive law.

Mathematical difference.

The *nonstandard proof* uses the Intermediate Value and the Standard Part Theorems, whereas the *standard proof* uses the completeness property of $\mathbb{R}$, the fact that a continuous function on [a,b] is bounded, together with the Intermediate Value Theorem. □

1.27. *Theorem* (Uniform limit theorem). Let $\langle f_n \rangle_{n\in\mathbb{N}}$ be a sequence of continuous functions converging uniformly to the function f on X. Then f is continuous on X.

Proof (Nonstandard). Let $x' \in X$ and $x \in {}^*X$, where $x \approx x'$. We need to show that $f(x) \approx f(x')$. For n finite, each f_n is continuous, so that the internal sequence $\langle \rho(f_n(x), f_n(x')) \rangle_{n\in{}^*\mathbb{N}}$ of hyperreal numbers is infinitesimal for finite n. Hence by the Infinitesimal Prolongation Theorem, there exists a $\nu \in {}^*\mathbb{N} - \mathbb{N}$ such that $\rho(f_n(x), f_n(x')) \approx 0$ for all $n < \nu$. Let $\mu \in {}^*\mathbb{N} - \mathbb{N}$ and $\mu < \nu$. Then $f_\mu(x) \approx f_\mu(x')$. By hypothesis, $f(x) \approx f_\mu(x)$ and $f_\mu(x') \approx f(x')$, hence $f(x) \approx f(x')$. □

Comparison with a standard proof (RUDIN [1976]).

Structural difference.

Basic steps of the *nonstandard proof.*

(1) The construction of an internal sequence of real numbers.

(2) The proof that for all infinite integers, the values of this sequence are infinitesimal.

(3) The verification of the nonstandard continuity criterion.

Basic steps of the *standard proof.*

(1) The construction of a convergent Cauchy sequence.

(2) The verification of an ε estimate.

(3) The introduction of a suitable neighbourhood of a given point satisfying an ε condition.

(4) The verification that all points in this neighbourhood satis-

fy the required ε condition.

Logical difference.

Both proofs are direct and involve no negations. The *nonstandard proof* contains only the first order quantifiers $(\forall n \in \mathbb{N})$, $(\exists n \in \mathbb{N})$, and $(\exists n \in {}^*\mathbb{N} - \mathbb{N})$, whereas the *standard proof,* in addition to containing the first order quantifiers $(\forall n \in \mathbb{N})$ and $(\exists n \in \mathbb{N})$, also quantifies over the points of X in the form of $(\forall x \in X)$ and $(\forall x \in V \in p(X))$.

Mathematical difference.

The *nonstandard proof* uses the Internal Function and Infinitesimal Prolongation Theorems and the nonstandard criterion for continuity, whereas the *standard proof* uses the fact that Cauchy sequences converge. □

1.28. *Theorem* (Bolzano-Weierstrass theorem). Every sequence in a compact metric space has a convergent subsequence.

Proof (Nonstandard). Let $< x_n >_{n \in \mathbb{N}}$ be a sequence of points of a space X. It is clearly sufficient to show that this sequence has a limit point in X. For this purpose we let ν be a infinite integer. By the compactness of X, x_ν is near standard, i.e., $x_\nu \approx x$ for some $x \in X$ and so x is a limit point of the given sequence. □

Comparison with a standard proof (RUDIN [1976]).

Structural difference.

Basic steps of the *nonstandard proof.*

(1) The proof that for any infinite integer ν the values x_ν of a given sequence are near standard.

(2) The verification that the standard part of any such x_ν is a limit point of the given sequence.

Basic steps of the *standard proof.*

(1) The observation that a sequence with finite range must have a value that is repeated infinitely often.

(2) The verification that a sequence with infinite range has a limit point.

(3) The construction of a subsequence converging to that limit point.

Logical difference.

The *nonstandard proof* contains no negations and the single first order quantifier $(\exists x \in X)$. The *standard proof*, on the other hand, involves four negations: (1) The denial of the existence of a limit point; (2) The denial that there exists a finite subcover of some cover of X ; (3) The denial that neighbourhoods of limit points meet X infinitely often; (4) The denial that a given point is a limit point. The proof involves the following first, second, and third order quantifiers: $(\exists n \in \mathbb{N})$, $(\forall x \in X)$, $(\exists x \in X)$, $(\exists U \in p(X))$, $(\exists C \in p(p(X)))$ and uses both the law of the excluded middle and the law of contraposition.

Mathematical difference.

The *nonstandard proof* uses merely the nonstandard criterion of compactness, whereas the *standard proof* uses the additional facts that any infinite subset of a compact set has a limit point in the set and that every neighbourhood of a limit point of a set contains infinitely many points of that set. □

1.29. *Theorem* (Dini's theorem). Let $\langle f_n \rangle_{n \in \mathbb{N}}$ be a sequence of continuous real-valued functions on a compact space X, let $f_{n+1}(p) \leq f_n(p)$ for all $n \in \mathbb{N}$ and all $p \in X$, and let $f_n(p) \to f(p)$ for all $p \in X$, with f continuous on X. Then the sequence $\langle f_n \rangle_{n \in \mathbb{N}}$ converges uniformly to f on X.

Proof (Nonstandard). Without loss of generality we may assume that $f(p) = 0$. By hypothesis, if $n \geq m$, with $n,m \in \mathbb{N}$, then $0 \leq f_n(p) \leq f_m(p)$ for all $p \in X$. Let $\nu \in {}^*\mathbb{N} - \mathbb{N}$ and $p \in {}^*X$. It suffices to establish that $f_\nu(p) \approx 0$. By the Transfer Principle, we have that $0 \leq f_\nu(p) \leq f_n(p)$ for all $n \in \mathbb{N}$.

Since X is compact, $p \approx q$ for some $q \in X$. Furthermore, since each f_n is continuous by hypothesis, it follows that $f_n(p) \approx f_n(q)$ for each $n \in \mathbb{N}$, and so, since $f_n(q) \in \mathbb{R}$, we have that $f_n(q) = st(f_n(p))$. Since $f_n(p)$ is finite, so is $f_\nu(p)$ and hence near standard. Therefore $0 \leq st(f_\nu(p)) \leq st(f_n(p)) = f_n(q)$ for all $n \in \mathbb{N}$. Since $f_n(q) \to 0$, we therefore have that $st(f_\nu(p)) = 0$, i.e., $f_\nu(p) \approx 0$. □

Comparison with standard proofs (FULKS [1961], GOLDBERG [1964], ROYDEN [1968]).

Structural differences.

Basic steps of the *nonstandard proof*.

(1) The observation that any point $p \in {}^*X$ is infinitely close to some point $q \in X$.

(2) The proof that for any infinite ν the values $f_\nu(p)$ are finite for all points $p \in {}^*X$.

(3) The verification that the standard part of $f_\nu(p)$ is 0 .

Basic steps of *the standard proof by Fulks.*

(1) The definition of a sequence of real numbers made up of the expressions $\sup |f_n(x) - f(x)|$ (taken over $[a,b]$) .

(2) The extraction of a subsequence with a positive lower bound.

(3) The introduction of a sequence of points of $[a,b]$ which allows the expression of the terms of the sequence in (2) in terms of the values of f_n and f .

(4) The verification that the sequence of points in (3) converges.

(5) The use of (2) and (4) to construct a sequence of real numbers that does and does not converge to 0 .

Basic steps of *the standard proof by Goldberg.*

(1) The definition of the functions $g_n = f - f_n$.

(2) The association with each $x \in [a,b]$ of an index $n \in \mathbb{N}$ such that g_n satisfies an ε condition at x .

(3) The association with each $x \in [a,b]$ of an open ball whose elements satisfy an ε condition.

(4) The selection of a finite subcover of the cover constructed in (3).

(5) The verification of the appropriate ε condition.

Basic steps of *the standard proof by Royden.*

(1) The construction of an open cover of X .

(2) The extraction of a finite subcover of the cover constructed in (1).

(3) The verification of an ε condition.

We remark that in FULKS [1961] and GOLDBERG [1964], the Theorem is stated for functions defined on [a,b] in place of a compact metric space X , and that in ROYDEN [1968] the Theorem is stated for upper semi-continuous functions defined on a countably compact space X . For our purposes, these variations are inessential.

Logical differences.

The *nonstandard proof* involves no negations and contains only the first order quantifiers $(\forall n \in \mathbb{N})$, $(\forall x \in X)$, and $(\exists x \in X)$ and uses the Transfer Principle. The *standard proof by Fulks* is indirect and involves three denials: (1) The denial that the given sequence converges uniformly; (2) The denial that the sequence constructed in Step (1) converges to 0 ; (3) The denial that the sequence constructed in Step (5) converges to 0 . The proof contains the first order quantifiers $(\forall n \in \mathbb{N})$, $(\exists n \in \mathbb{N})$, and $(\forall x \in X)$, and the second order quantifiers $(\exists s \in p(\mathbb{N} \times \mathbb{N}))$ and $(\exists s \in p(\mathbb{N} \times \mathbb{R}))$. It uses the law of the excluded middle and the contrapositive law. The *standard proof by Goldberg* is direct and involves no negations and only the first order quantifiers $(\forall n \in \mathbb{N})$, $(\exists n \in \mathbb{N})$, $(\forall x \in X)$, and $(\exists x \in X)$. The *standard proof by Royden* is also direct. It involves no negations, but contains both first and third order quantifiers: $(\forall n \in \mathbb{N})$, $(\exists n \in \mathbb{N})$, $(\forall x \in X)$, and $(\forall C \in p(p(X)))$.

Mathematical differences.

The *nonstandard proof* uses the nonstandard criterion for compactness and the Standard Part Theorem. The *standard proof by Fulks* uses the Maximum Value Theorem and the fact that a bounded sequence of real numbers has a limit point and therefore a convergent subsequence. The *standard proof by Goldberg* uses the fact that open balls form a topological base and the fact that the inverse images of open sets under a continuous function are open. To make the comparison with the *standard proof by Royden* complete we require the additional fact that a metric space is compact if and only if it is countably compact, the fact that a function is continuous if and only if it is both upper and lower semi-continuous, and the fact that the inverse images of open sets under a continuous function are open. □

CHAPTER 2

MONADIC NEIGHBOURHOODS

In this chapter, Ω denotes, ambiguously, *the class of open sets* of a given topological space X, and Ω_p denotes *the open neighbourhood system* of the point p in X.

Hausdorff spaces

2.1. *Definition* (Standard). A topological space is *Hausdorff* if and only if for any two distinct points p and q there exist open neighbourhoods $U(p)$ and $V(q)$ such that $U(p) \cap V(q) = \emptyset$.

2.2. *Definition* (Nonstandard). A topological space is *Hausdorff* if and only if for any two distinct points p and q $\mu(p) \cap \mu(q) = \emptyset$.

2.3. *Theorem* (Characterization theorem). Definitions (2.1) and (2.2) are equivalent.

Proof. Suppose the space is Hausdorff in the sense of (2.1). Then the given points p and q have disjoint open neighbourhoods U and V. Thus $\emptyset = {}^*\emptyset = {}^*(U \cap V) = {}^*U \cap {}^*V$. Since $\mu(p) \subset {}^*U$ and $\mu(q) \subset {}^*V$, we have $\mu(p) \cap \mu(q) = \emptyset$.

Conversely, if $\mu(p) \cap \mu(q) = \emptyset$, then it follows from the Monadic Neighbourhood Theorem that "there exist internal open sets U and V such that $p \in U$, $q \in V$, and $U \cap V = \emptyset$". Thus

$$\tilde{W} \models (\exists U \in {}^*\Omega)(\exists V \in {}^*\Omega)(p \in U \wedge q \in V \wedge U \cap V = \emptyset).$$

The result therefore follows by an application of the Transfer Principle. □

Regular spaces

2.4. *Definition* (Standard). A topological space is *regular* if and only if for any point p and any closed set S not containing p there exists open neighbourhoods $U(p)$ and $V(S)$ such that $U(p) \cap V(S) = \emptyset$.

2.5. *Definition* (Nonstandard). A topological space is *regular* if and only if for any point p and any closed set S not containing p, $\mu(p) \cap \mu(S) = \emptyset$.

2.6. *Theorem* (Characterization theorem). Definitions (2.4) and (2.5) are equivalent.

Proof. Suppose the space is regular in the sense of (2.4). Then for any point p and any closed set S not containing p there exist disjoint open neighbourhoods U and V . Thus $\emptyset = {}^*\emptyset = {}^*(U \cap V) = {}^*U \cap {}^*V$. Since $S \subset V$, it follows that ${}^*S \subset {}^*V$ (cf. Appendix 3, Property 7). By the definition of monads we have

$$\mu(p) = \bigcap_{p \in U \subset \Omega_p} {}^*U \quad \text{and} \quad \mu(S) = \bigcap_{S \subset V \subset \Omega} {}^*V ,$$

hence $\mu(S) \cap {}^*U = \emptyset$ and the result follows.

Conversely, if $\mu(p) \cap \mu(S) = \emptyset$, "there exist internal open sets U and V such that $p \in U$, ${}^*S \subset V$, and $U \cap V = \emptyset$ " . Thus

$$\tilde{W} \models (\exists U \in {}^*\Omega)(\exists V \in {}^*\Omega)(p \in U \wedge {}^*S \subset V \wedge U \cap V = \emptyset) .$$

The result therefore follows by an application of the Transfer Principle. □

Normal spaces

2.7. *Definition* (Standard). A topological space is *normal* if and only if for any two disjoint closed sets S and T there exist open neighbourhoods $U(S)$ and $V(T)$ such that $U(S) \cap V(T) = \emptyset$.

2.8. *Definition* (Nonstandard). A topological space is *normal* if and only if for any two disjoint closed sets S and T, $\mu(S) \cap \mu(T) = \emptyset$.

2.9. *Theorem* (Characterization theorem). Definitions (2.7) and (2.8) are equivalent.

Proof. Suppose the space is normal in the sense of (2.7). Then for the given disjoint closed sets S and T there exist open neighbourhoods U and V. Thus $\emptyset = {}^*\emptyset = {}^*(U \cap V) = {}^*U \cap {}^*V$. Since

$$\mu(S) = \bigcap_{S \subset U \subset \Omega} {}^*U \quad \text{and} \quad \mu(T) = \bigcap_{T \subset V \subset \Omega} {}^*V ,$$

it follows that $\mu(S) \cap \mu(T) = \emptyset$.

Conversely, if $\mu(S) \cap \mu(T) = \emptyset$, then "there exist internal open sets U and V such that ${}^*S \subset U$, ${}^*T \subset V$, and $U \cap V = \emptyset$". Thus

$$\tilde{W} \models (\exists U \in {}^*\Omega)(\exists V \in {}^*\Omega)({}^*S \subset U \wedge {}^*T \subset V \wedge U \cap V = \emptyset) .$$

The result therefore follows by an application of the Transfer Principle. □

Compact spaces

2.10. *Definition* (Standard). A topological space is *compact* if and only if every open cover has a finite subcover.

2.11. *Definition* (Nonstandard). A topological space X is *compact* if and only if for every $p \in {}^*X$ there exists a $q \in X$ such that $p \in \mu(q)$.

2.12. *Theorem* (Characterization theorem). Definitions (2.10) and (2.11) are equivalent.

Proof. Suppose X is compact in the sense of (2.10) and let $p \in {}^*X$, but $p \notin \mu(q)$ for any $q \in X$, so that for each $q \in X$, $p \notin \bigcap_{V \in \Omega_q} {}^*V$. It follows that for each $q \in X$, there is some $V_q \in \Omega_q$ such that $p \notin {}^*V_q$. Since $q \in V_q$, $X \subset \bigcup_{q \in X} V_q$, and so by the compactness of X there exist $q_1, \ldots, q_n \in X$ such that $X \subset V_{q_1} \cup \ldots \cup V_{q_n}$, i.e.,

$$\hat{S} \models (\forall x \in X)(x \in V_{q_1} \cup \ldots \cup V_{q_n}) .$$

But $p \in {}^*X$, and by the Transfer Principle we therefore have that

$$\tilde{W} \models (\forall x \in {}^*X)(x \in {}^*V_{q_1} \cup \ldots \cup {}^*V_{q_n}) ,$$

so that p belongs to one of the ${}^*V_{q_1}, \ldots, {}^*V_{q_n}$, which is a contradiction.

Conversely, suppose that for every $p \in {}^*X$, $p \in \mu(q)$ for some $q \in X$, but that X is not compact. Then there is an open cover G of X which has no finite subcover. Let

$$r = \{ \langle A,a \rangle \mid A \in G ,\ a \in X ,\ a \notin A \} .$$

The relation r is concurrent: For each $A \in G$, there is a $p \in X$

such that $\langle A,p \rangle \in r$, otherwise $\{ A \}$ would cover X . Hence the domain of r is G . Let $A_1, \ldots, A_m \in G$. By hypothesis, the set $\{ A , \ldots, A_m \}$ does not cover X , hence there is a $p \in X$ such that $p \notin A_1 \cup \ldots \cup A_m$, i.e., $\langle A_1,p \rangle, \ldots , \langle A_m,p \rangle \in r$, so r is concurrent.

By applying the Concurrence Theorem, we have the existence of a $p \in \tilde{W}$ such that for every $A \in G$, $\langle {}^*A,p \rangle \in r$. Furthermore, since ${}^*r \subset {}^*G \times {}^*X$, it follows that $p \in {}^*X$. Now for each $A \in G$,

$$\hat{S} \models (\forall a \in X)(\langle A,a \rangle \in r \Rightarrow a \notin A) .$$

By applying the Transfer Principle, we have that

$$\tilde{W} \models (\forall a \in {}^*X)(\langle {}^*A,a \rangle \in {}^*r \Rightarrow a \notin {}^*A) .$$

By hypothesis, there exists a $q \in X$ such that $p \in \mu(q)$, and so $q \in A$ for some $A \in G$. Hence we have that $p \in \mu(q) \in {}^*A$, since $\mu(q)$ is the intersection of all ${}^*A$ with $A \in \Omega_q$ and therefore both $p \in {}^*A$ and $p \notin {}^*A$, which is a contradiction. □

Some standard theorems (A comparison of standard and nonstandard proofs)

2.13. *Theorem* (Uniform continuity theorem). If f is continuous on the compact space X, then f is uniformly continuous on X.

Proof (Nonstandard). By Theorem 1.6, it suffices to show that f satisfies the nonstandard definition of continuity on $^*X$. Let $x, x' \in {}^*X$, where $x \approx x'$. We have to show that $f(x) \approx f(x')$. Since X is compact, $x \approx t$ for some $t \in X$, and hence $x' \approx t$. Thus $f(x) \approx f(t)$ and $f(x') \approx f(t)$, since f is continuous on X. Consequently, we have that $f(x) \approx f(x')$. □

Comparison with a standard proof (RUDIN [1976]).

Structural difference.

Basic steps of the *nonstandard proof.*

(1) The verification that every point of $^*X$ is near standard.

(2) The verification of the nonstandard criterion for uniform continuity.

Basic steps of the *standard proof.*

(1) The construction of a function $\phi : X \to \mathbb{R}$ whose values satisfy some ε condition.

(2) The construction of an open cover involving the function ϕ.

(3) The selection of a finite subcover.

(4) The verification of an ε condition.

Logical difference.

Both proofs are direct and do not involve negation. The *nonstandard proof* contains the first order quantifiers $(\exists x \in X)$ and $(\forall x \in {}^*X)$ and the *standard proof* contains the first order quantifiers $(\forall n \in \mathbb{N})$, $(\forall x \in X)$ and $(\exists x \in X)$. Thus the extraneous quantifier $(\forall n \in \mathbb{N})$ in the standard proof is absorbed into the canonical nonstandard quantifier $(\forall x \in {}^*X)$ in the nonstandard proof.

Mathematical difference.

Whereas the *nonstandard proof* uses no additional mathematical properties, the *standard proof* relies on the fact that the minimum of a finite set of positive numbers is positive. □

2.14. *Theorem* (Equicontinuity theorem). If $\langle f_n \rangle_{n \in \mathbb{N}}$ is a sequence of continuous functions on a compact space X and if $f_n(x) \to f(x) \in X$ for each x, then the sequence $\langle f_n \rangle_{n \in \mathbb{N}}$ converges uniformly to f if and only if it is equicontinuous.

Proof (Nonstandard). Suppose the convergence is uniform. Then it follows from the Uniform Limit Theorem that f is continuous on X. Furthermore, since X is compact, f and all f_n are uniformly continuous on X, and hence they satisfy the nonstandard definition of continuity on $^*X$ (cf. Theorem 1.3). Let $\nu \in {}^*\mathbb{N} - \mathbb{N}$. Then, if $x, x' \in {}^*X$ and $x \approx x'$, we have $f_\nu(x) \approx f(x) \approx f(x') \approx f_\nu(x')$. Therefore f_n satisfies the nonstandard definition of continuity for $n \in {}^*\mathbb{N} - \mathbb{N}$. Since, by hypothesis, this holds for all $n \in \mathbb{N}$, it follows that $\langle f_n \rangle_{n \in \mathbb{N}}$ is equicontinuous.

Conversely, suppose that the sequence is equicontinuous. Then the compactness of X guarantees that for any $x' \in {}^*X$ there exists a $x \in X$ with $x' \approx x$, and by continuity, $f_n(x') \approx f_n(x)$ for all $n \in {}^*\mathbb{N}$. Moreover, for $\nu \in {}^*\mathbb{N} - \mathbb{N}$, $f_\nu(x) \approx f(x)$. By the Transfer Principle, f is (uniformly) continuous on X, so that $f(x') \approx f(x)$. Thus we have that $f_\nu(x') \approx f_\nu(x) \approx f(x) \approx f(x')$ for all $\nu \in {}^*\mathbb{N} - \mathbb{N}$ and so the convergence is uniform. □

Comparison with standard proofs (ROYDEN [1968], RUDIN [1976]).

Structural difference.

Basic steps of the *nonstandard proof*.

(1) The proof that f is continuous on X.

(2) The proof that f and f_n are uniformly continuous on X.

(3) The proof that f and f_n are continuous on $*X$.

(4) The proof that f_ν is continuous on $*X$ for $\nu \in *\mathbb{N} - \mathbb{N}$.

(5) The proof that every point of $*X$ is near standard.

(6) The verification of an infinitesimal estimate.

Basic steps of the *standard proof*.

(1) The verification of an ε-δ-n estimate.

(2) The construction of an open cover satisfying an ε condition.

(3) The selection of a finite subcover.

(4) The verification of an ε condition involving an $n \in \mathbb{N}$.

(5) The verification of an ε condition involving the points of X.

Logical difference.

Both proofs are direct and involve no negations. The *nonstandard proof* contains only first order quantifiers. They are $(\forall n \in \mathbb{N})$, $(\forall n \in *\mathbb{N} - \mathbb{N})$, $(\forall x \in *X)$, $(\exists x \in *X)$, and $(\exists x \in X)$. The proof uses the Transfer Principle. The *standard proof* contains the first order quantifiers $(\forall n \in \mathbb{N})$, $(\forall x \in X)$, $(\exists x \in X)$ and the second order quantifiers $(\exists U \in p(X))$ and $(\forall x \in U \in p(X))$. It also contains the third order quantifier $(\exists C \in p(p(X)))$.

Mathematical difference.

The *nonstandard proof* uses the Uniform Limit and the Uniform Continuity Theorems, and the nonstandard criteria for continuity, uniform continuity, and compactness. The *standard proof* uses the Uniform Continuity Theorem, the fact that the inverse image of an open set under a continuous function is open, the fact that the open balls form a basis for the metric topology, and the fact that every point of an open set is an interior point. □

2.15. *Theorem* (Compact image theorem). If f is continuous at each point of a compact space X, then $f[X]$ is compact.

For the proof of Theorem 2.15, we use the following nonstandard characterization of continuous functions:

2.16. *Theorem* (Nonstandard continuity theorem). Let $f : X \to Y$ be a function between topological spaces and let $p \in X$. Then f is continuous at p if and only if $q \approx p \Rightarrow {}^*f(q) \approx f(p)$.

Proof. Suppose that f is continuous at p and G is any open neighbourhood of $f(p)$. Then there exists an open neighbourhood $H \in \Omega_p$ such that $f[H] \subset G$ and so ${}^*f[{}^*H] \subset {}^*G$ (cf. Appendix 3, Properties 7 and 9). Since $\mu(p) \subset {}^*H$, it is immediately clear that ${}^*f[\mu(p)] \subset {}^*f[{}^*H] \subset {}^*G$. Hence ${}^*f[\mu(p)] \subset \bigcap_{G \in \Omega_{f(p)}} {}^*G$, i.e., ${}^*f[\mu(p)] \subset \mu(f(p))$.

Conversely, by the Monadic Neighbourhood Theorem there exists an internal set D such that $D \in {}^*\Omega_p$ and $D \subset \mu(p)$. By hypothesis, we have that $q \in D$ implies ${}^*f(q) \in \mu(f(p))$, so that if H is any open neighbourhood of $f(p)$, then $q \in D$ implies ${}^*f(q) \in {}^*H$. Hence for any $H \in \Omega_{f(p)}$,

$$\tilde{W} \models (\exists D \in {}^*\Omega_p)(\forall x \in D)(f(x) \in {}^*H) .$$

By the Transfer Principle, we have that

$$\hat{S} \models (\exists D \in \Omega_p)(\forall x \in D)(f(x) \in H) .$$

Thus we have an open neighbourhood D of p such that $f[D] \subset H$, and so f is continuous at p. □

Proof of Theorem 2.15 (Nonstandard). Let $q \in {}^*(f[X])$. By Property 9 of Appendix 3, ${}^*(f[X]) = {}^*f[{}^*X]$, so that $q = {}^*f(r)$ for some point $r \in {}^*X$. Since X is compact, $r \approx r_0$ for some $r_0 \in X$, and by the Nonstandard Continuity Theorem we have $q = {}^*f(r) \approx f(r_0) \in f[X]$, so that $f[X]$ is compact (cf. Theorem 2.12). □

Comparison of the proof of Theorem 2.16 with a standard proof (RUDIN [1976]).

Structural difference.

Basic steps of the *nonstandard proof*.

(1) The observation that the star of the range of f is the range of the star of f.

(2) The observation that every point of ${}^*X$ is near standard.

(3) The verification of an infinitesimal estimate.

Basic steps of the *standard proof*.

(1) The construction of an open cover of X from an open cover of Y.

(2) The selection of a finite subcover.

(3) The verification that $f[X]$ is contained in the image of that finite subcover.

Logical difference.

Both proofs are direct and involve no negations. The *nonstandard proof* contains only the first order quantifiers $(\exists x \in X)$, $(\exists x \in {}^*X)$, and $(\forall x \in {}^*(f[X]))$ and uses the Transfer Principle. The *standard proof* contains the first order quantifier $(\exists i \in I)$ referring to an extraneous index set I.

Mathematical difference.

The *nonstandard proof* uses the Concurrence Theorem, the Monadic Neighbourhood Theorem, the nonstandard criteria for compactness and continuity, and Properties 7 and 9 of Appendix 3. The *standard proof* uses the fact that the inverse images of open sets under continuous functions are open, that the open balls form a basis for the metric topology, and that all points of the inverses of open sets under continuous functions are interior. □

2.17. *Theorem* (Tychonoff's theorem): The product space $\Pi_{i\in I} X_i$ of a set of compact spaces $\{ X_i \mid i \in I \}$ is compact.

Proof (Nonstandard). Let $X = \Pi_{i\in I} X_i$ and let $f \in {}^*X$. We have to show that $g \approx f$ for some $g \in X$. Assume that $f(\nu)$ is near standard for all $\nu \in I$. By the Axiom of Choice, there exists a $g \in X$ such that $f(\nu) \in \mu(g(\nu))$ for all $\nu \in I$. Let

$$U = \{ g \mid g(\nu_i) \in U_i , \; i = 1,\dots,k , \; U_i \in \Omega(X_{\nu_i}) \}$$

be a basic open set. Since $f(\nu_i) \in \mu(g(\nu_i))$ for $i = 1,\dots,k$, and since $\mu(g(\nu_i)) \subset {}^*U_i$, we have by the Transfer Principle that $g \in {}^*U$. Since U was an arbitrary basic open set, this calculation holds for all basic open sets containing g and therefore $f \in \mu(g)$. □

Comparison with a standard proof (KELLEY [1955]).

Structural difference.

Basic steps of the *nonstandard proof*.

(1) The selection of a point $g \in X$.

(2) The introduction of a basic open neighbourhood U of g.

(3) The proof that $g \in {}^*U$.

(4) The verification that $f \in \mu(g)$.

Basic steps of the *standard proof*.

(1) The construction of a set of basic open sets in the product topology.

(2) The selection of a subfamily which fails to cover X.

(3) The formation of the set of components of the members of the subfamily constructed in (2).

(4) The construction of a point of X which does not belong to the union of the family constructed in (2).

Logical difference.

The *nonstandard proof* is direct and involves no negations. It uses the Axiom of Choice and the Transfer Principle. It contains the first order quantifiers $(\forall i \in I)$, $(\exists x \in X)$, and $(\forall x \in {}^*X)$. The *standard proof* involves three negations: (1) The denial that the family constructed in Step 2 is a cover; (2) The denial that the point constructed in Step 4 belongs to the union of the family constructed in Step 2; and (3) the denial that there exists a finite subcover. The proof uses Tukey's Lemma, i.e., a variant of the Axiom of Choice, and contains first, second, and third order quantifiers: $(\forall i \in I)$, $(\forall x \in U \in p(X))$, and the quantifier $(\exists C \in p(p(X)))$.

Mathematical difference.

The *nonstandard proof* is based on the verification of the nonstandard criterion for compactness, whereas the *standard proof* uses Alexander's Theorem. □

2.18. *Theorem* (Heine-Borel theorem). Any closed and bounded subset of $\mathbb{R}^n$ is compact.

Proof (Nonstandard). Let B be a closed and bounded subset of $\mathbb{R}^n$. We have to show that for each $p \in {}^*B$ there exists a $q \in B$ such that $p \in \mu(q)$. Let $p = \langle p_1, \ldots, p_n \rangle$. Then Theorem 4.3.1 of ROBINSON [1966], for example, shows that p is finite since B is bounded. Hence $p_1, \ldots, p_n$ are finite and $q = \langle st(p_1), \ldots, st(p_n) \rangle$ exists. Moreover, $\rho(p,q) = ((p_1 - st(p_1))^2 + \ldots + (p_n - st(p_n))^2)^{\frac{1}{2}} \approx 0$. Thus $p \in \mu(q)$ and since B is closed it follows that $q \in B$. □

Comparison with a standard proof (KELLEY [1955]).

Structural difference.

Basic steps of the *nonstandard proof.*

(1) The proof that each point in the star of a closed and bounded subset of $\mathbb{R}^n$ is finite.

(2) The verification that each such point is near standard (with respect to the given closed and bounded subset).

Basic steps of the *standard proof.*

(1) The embedding of the given set B in a product of closed and bounded intervals.

(2) The proof that the theorem holds for closed and bounded intervals.

(3) The proof that the product in (1) is compact.

(4) The observation that a closed subset of a compact set is compact.

Logical difference.

The proof that p is finite in the *nonstandard proof* is based on Theorem 4.3.1 of ROBINSON [1966]. The proof of that theorem involved the denial (hence a negation) that p is finite. It also contains the quantifiers $(\exists x \in {}^*\mathbb{R})$, $(\forall x \in \mathbb{R})$, and $(\exists x \in \mathbb{R}^n)$. It uses the Transfer Principle, the law of the excluded middle, and the law of contraposition. The *standard proof* involves an nonequality $c \neq b$ and the second order quantifier $(\exists C \in p(p(\mathbb{R})))$.

Mathematical difference.

The *nonstandard proof* uses the fact that a metric space X is bounded if and only if every point of $^*X$ is finite. The *standard proof*, on the other hand, uses the following further facts: (1) The fact that the product of compact sets is compact; (2) The fact that closed subsets of compact sets are compact; (3) The fact that projections preserve boundedness; and (4) the completeness of $\mathbb{R}$. □

CHAPTER 3

CONCURRENT RELATIONS

Existence theorem

Let r be a concurrent relation in $\hat{S}$ and I the set of all functions with the property that

(1) α is defined on the set of concurrent relations $r \in \hat{S}$;

(2) For each such r , $\alpha(r)$ is a finite subset of dom(r) ;

and let F be an ultrafilter on I .

We require the following facts and notations:

(1) $\alpha < \beta$ if $\alpha , \beta \in I$ and $\alpha(r) \subset \beta(r)$ for all concurrent $r \in \hat{S}$;

(2) $\gamma = \alpha \vee \beta$ if $\gamma \in I$ and $\gamma(r) = \alpha(r) \cup \beta(r)$ for all concurrent $r \in \hat{S}$;

(3) $\Gamma_\alpha = \{ \beta \in I \mid \alpha < \beta \}$ for each $\alpha \in I$, with the obvious property that $\Gamma_\alpha \cap \Gamma_\beta = \Gamma_{\alpha\vee\beta}$ for all $\alpha , \beta \in I$;

(4) F is an ultrafilter containing the sets Γ_α for all α ;

(5) Let $r \in S_k$ be a concurrent relation and $f : I \to \hat{S}$ a function such that $< a,f(\alpha) > \in r$ for each $a \in \alpha(r)$ and each $\alpha \in I$. Then an easy calculation shows that for each $a \in \mathrm{dom}(r)$, $< a,f(\alpha) > \in r$ a.e. (cf. DAVIS [1977]). By the transitivity of S_k , $f(\alpha) \in S_k$, and therefore $f \in Z$.

(6) Let $b = \bar{f} \in \tilde{W}$.

3.1. *Theorem* (Concurrence theorem). If r is a concurrent relation in $\hat{S}$, then there exists an element $b \in \tilde{W}$ such that $< {}^*a,b > \in {}^*r$ for all $a \in dom(r)$.

Proof. Let $b = \bar{\bar{f}}$ as defined in (6) above, and let $h(\alpha) = < a,f(\alpha) >$ for all $\alpha \in I$. Then $\bar{\bar{h}} = < \bar{\bar{a}}, \bar{\bar{f}} > = < {}^*a,b >$ (cf. Appendix 3, Property 3), and by (5) above, $h(\alpha) \in r$ a.e., so that $\bar{\bar{h}} \in \bar{\bar{r}} = {}^*r$ (cf. Appendix 3, Property 1). Hence $< {}^*a,b > \in {}^*r$. □

Some nonstandard applications of the Concurrence Theorem

3.2. *Theorem* (Infinite integer theorem). There exists an element $b \in {}^*\mathbb{N}$ with the property that $n < b$ for all $n \in \mathbb{N}$.

Proof. Let $L = \{ \langle x,y \rangle \mid x, y \in \mathbb{N}$ and $x < y \}$. The relation L is clearly concurrent since $\mathrm{dom}(L) = \mathbb{N}$, and if $a_1, \ldots, a_n \in \mathbb{N}$ and b is the largest of $a_1, \ldots, a_n$, then $\langle a_1,b \rangle \in L, \ldots, \langle a_n,b \rangle \in L$. By the Concurrence theorem there is therefore an element $b \in \tilde{W}$ such that $\langle a,b \rangle \in {}^*L$ for all $a \in \mathbb{N}$ (in this case, ${}^*a = a$ because $a \in \mathbb{N} \subset S$). Since $L \subset \mathbb{N} \times \mathbb{N}$, we have that ${}^*L \subset {}^*\mathbb{N} \times {}^*\mathbb{N}$, and thus $b \in {}^*\mathbb{N}$.

Is $b \in \mathbb{N}$? If so, we can write ${}^*b = b$ and conclude from

$$\tilde{W} \models \langle {}^*a,b \rangle \in {}^*L$$

for all $a \in \mathbb{N}$ (by the Transfer Principle) that

$$\hat{S} \models \langle a,b \rangle \in L .$$

Thus we have $a < b$ for all $a \in \mathbb{N}$, i.e., if $b \in \mathbb{N}$, then b is a largest integer. Since no such number exists, we conclude that $b \in {}^*\mathbb{N} - \mathbb{N}$. Hence ${}^*\mathbb{N} - \mathbb{N} \neq \emptyset$. □

In the next theorem, we assume that D is an Archimedian order field, that $D \subset S \subset \hat{S}$, and define F, I, and F/I as follows:

(1) $F = \{ x \in D \mid |x| < n \text{ for some } n \in \mathbb{N} \}$;

(2) $I = \{ x \in D \mid x = 0 \text{ or } 1/x \in D - F \}$;

(3) $F/I = \{ [x] \mid x \in D \}$;

(4) $[x] = \{ y \in D \mid x - y \in I \}$.

3.3. *Theorem* (Dedekind's theorem). Let A, B be nonempty subsets of D such that $a \in A$ and $b \in B$ implies that $a < b$. Then there exists an element $c \in F/I$ such that for all $a \in A$ and $b \in B$, $a \leq c \leq b$.

Proof. Let $r = \{ \langle a,b \rangle \mid a \in A,\ b \in D,\ a \leq b < c \text{ for all } c \in B \}$. We shall show that r is a concurrent relation. Clearly, $\mathrm{dom}(r) = A$ since for any $a \in A$, $\langle a,a \rangle \in r$. Let $a_1, \ldots, a_n \in A$ and let b be the largest of $a_1, \ldots, a_n$. Then $b \in D$ and each $a_i \leq b$. Moreover, since $b \in A$, we have that b is less than any element of B. Hence $\langle a_i, b \rangle \in r$ for $i = 1, 2, \ldots, k$, and so r is concurrent. By the Concurrence Theorem there is an element $t \in \tilde{W}$ such that $\langle a,t \rangle \in {}^*r$ for all $a \in A$.

Since $r \subset A \times D$, we have that ${}^*r \subset {}^*A \times {}^*D$, and hence that $t \in {}^*D$. By the definition of r we obtain

$$\hat{S} \models (\forall x \in A)(\forall y \in D)(\langle x,y \rangle \in r \Rightarrow x \leq y)$$

$$\hat{S} \models (\forall x \in A)(\forall y \in D)(\forall u \in B)(\langle x,y \rangle \in r \Rightarrow y < u) .$$

By applying the Transfer Principle, we therefore have

$\tilde{W} \models (\forall x \in {}^*A)(\forall y \in {}^*D)(\langle x,y \rangle \in {}^*r \rightarrow x \leq y)$

$\tilde{W} \models (\forall x \in {}^*A)(\forall y \in {}^*D)(\forall u \in {}^*B)(\langle x,y \rangle \in {}^*r \rightarrow y < u)$.

Thus $a \leq t < b$ for all $a \in {}^*A$ and $b \in {}^*B$, and since $B \neq \emptyset$, t is therefore finite. Hence, by the Standard Part Theorem, we have that for $a \in A$ and $b \in B$, $a = st(a) \leq st(t) \leq st(b) = b$. The conclusion therefore follows, with $c = st(t)$. □

3.4. *Theorem* (Monadic neighbourhood theorem). For each point p of a topological space X there exists an internal set $D \in {}^*\Omega_p$ such that $D \subset \mu(p)$.

Proof. Let $r = \{ \langle x,y \rangle \mid x, y \in \Omega_p \text{ and } y \subset x \}$. We shall show that r is a concurrent relation. The domain of r is the set of open neighbourhoods of p and the concurrence of r follows from the fact that if $U_1, \ldots, U_n$ are open neighbourhoods of p, then the set $V = U_1 \cap \ldots \cap U_n$ is an open neighbourhood of p included in the sets $U_1, \ldots, U_n$. Thus $\langle U_i, V \rangle \in r$ for $i = 1, 2, \ldots, n$.

By the Concurrence Theorem, there exists a $D \in \tilde{W}$ such that for all $U \in \Omega_p$, $\langle {}^*U, D \rangle \in {}^*r$. We therefore have

$$\hat{S} \models (\forall x,y \in \Omega_p)(\langle x,y \rangle \in r \Rightarrow y \subset x) .$$

By applying the Transfer Principle, we have

$$\tilde{W} \models (\forall x,y \in {}^*\Omega_p)(\langle {}^*x,y \rangle \in {}^*r \Rightarrow y \subset {}^*x) .$$

Thus $D \subset {}^*U$ and therefore $D \subset \bigcap_{U \in \Omega_p} {}^*U = \mu(p)$. □

The spectral theorem for compact Hermitian operators

The spectral theorem for compact Hermitian operators

3.5. *Theorem* (Spectral theorem). If $T : H \to H$ is a compact Hermitian operator on a Hilbert space H and has only *finitely many* distinct eigenvalues $\pi_1, \ldots, \pi_k$, then

$$H = H_0 \oplus H_1 \oplus \ldots \oplus H_k$$
$$I = P_0 + P_1 + \ldots + P_k$$
$$T = \pi_1 P_1 + \ldots + \pi_k P_k$$

and if T has *infinitely many* distinct eigenvalues $\pi_1, \ldots, \pi_k, \ldots,$ then

$$H = H_0 \oplus H_1 \oplus \ldots \oplus H_k \oplus \ldots$$
$$I = P_0 + P_1 + \ldots + P_k + \ldots$$
$$T(a) = \pi_1 P_1(a) + \ldots + \pi_k P_k(a) + \ldots,$$

Where H_i is the eigenspace of π_i; H_0 is the orthogonal complement of $H_1 \oplus \ldots$; and $P_i : H \to H_i$ is the usual projection operator.

Proof (Nonstandard). The proof involves five distinct steps:

(1) The construction of an $\omega \in {}^*\mathbb{N} - \mathbb{N}$ - dimensional space H_ω and of an operator $T' : H_\omega \to H_\omega$;

(2) The decomposition of H_ω, T', and of the identity operator $I' : H_\omega \to H_\omega$ into

$$H_\omega = H'_1 \oplus \ldots \oplus H'_\mu$$

$$I' = P'_1 + \ldots + P'_\mu$$

$$T' = \kappa_1 P'_1 + \ldots + \kappa_\mu P'_\mu$$

by an application of the Transfer Principle to the Spectral Theorem for Hermitian operators on finite-dimensional unitary spaces ;

(3) The proof that T is the standard part of T', that the eigenvalues π_j of T are the standard parts of the eigenvalues λ_j of T', and that the eigenvectors s_j corresponding to π_j are the standard parts of the eigenvectors r_j corresponding to λ_j ;

(4) The proof that at all infinite integers $i \leq \omega$, the eigenvalues ν_i of T' are infinitesimal ;

(5) The distinction between the case of T having only finitely many distinct eigenvalues and the case of T having infinitely many distinct eigenvalues.

The construction of $T' : H_\omega \to H_\omega$. Let $\mathcal{S}$ be the set of all finite-dimensional subspaces of H and form $r = \{ \langle x,S \rangle \mid x \in S \in \mathcal{S} \}$. Since for any $x \in H$, we have $\langle x, \mathrm{span}(x) \rangle \in r$, it follows that the domain of r is H. Moreover, if $x_1, \ldots, x_n \in H$, then it is clear that $\langle x_i, \mathrm{span}(x_1, \ldots, x_n) \rangle \in r$ for $i = 1, 2, \ldots, n$. Hence the relation r is concurrent. By the Concurrence Theorem there therefore exists an E such that $\langle x, E \rangle \in {}^*r$ for all $x \in H$, and by the Transfer Principle , we have that $E \in {}^*\mathcal{E}$, so that for each $x \in H$, the point x is also in E , i.e., $H \subset E$. Moreover, the dimension function $\dim : \mathcal{E} \to \mathbb{N}$ has a unique extension ${}^*\dim : {}^*\mathcal{E} \to {}^*\mathbb{N}$ with ${}^*\dim(E) = \omega \in {}^*\mathbb{N} - \mathbb{N}$, since the space H is infinite-dimensional. Hence we denote E by H_ω .

Next we use the star map and lift the operator $T : H \to H$ to an operator ${}^*T : {}^*H \to {}^*H$. By composing ${}^*T$ with the projection $P_{H_\omega} : H \to H_\omega$ we obtain an operator $P_{H_\omega} \circ {}^*T : {}^*H \to H_\omega$ whose restriction to H_ω yields a compact Hermitian operator $T' : H_\omega \to H_\omega$.

The decomposition of H_ω , T' , *and* I' . By virtue of the Transfer Principle, we can apply the finite-dimensional theory of eigenvalues of Hermitian operators to T' and assume the non-zero eigenvalues of T' to be $\lambda_1 , \ldots , \lambda_\nu$ (permitting repetitions), ordered by their absolute values in the form $|\lambda_1| \geq |\lambda_2| \geq \ldots \geq |\lambda_\nu|$, with corresponding orthonormal eigenvectors $r_1 , \ldots , r_\nu$. By deleting repetitions, we can rewrite the sequence $\lambda_1 , \ldots , \lambda_\nu$ as $\kappa_1 , \ldots , \kappa_\mu$ $(\mu \leq \nu)$, where each λ_i is equal to exactly one κ_j . The eigenvectors of κ_j will be denoted by $r_{1j} , \ldots , r_{nj}$.

We now form $H_j = \text{span}(r_{1j} , \ldots , r_{nj})$ which, by the Transfer Principle, is the eigenspace of κ_j . By the spectral theorem for finite-dimensional spaces and the Transfer Principle, we obtain the desired decompositions, with $P_i = P_{H_i}$.

The standard part argument. For the routine calculations required to establish these facts, we refer to ROBINSON [1966], pp. 186-189.

The infinitesimal eigenvalues of T' . We recall that the eigenvalues $\lambda_1, \ldots, \lambda_\nu$ of T' were arranged so that

$$|\lambda_1| \geq \ldots |\lambda_k| \geq \ldots \geq |\lambda_\nu|$$

and an easy calculation (cf. ROBINSON [1966], p. 187) shows that if any positive real number ε is a lower bound for $|\lambda_k|$, then it follows that $k \in \mathbb{N}$, so that the set

$$W = \{ j \in {}^*\mathbb{N} \mid \kappa_j \text{ is not infinitesimal} \}$$

is either $\mathbb{N}$ or a finite subset of it, where the numbers κ_j are obtained from the λ_j by deleting repetitions. It is therefore clear that for any infinite integer j , the eigenvalue κ_j is infinitesimal.

The decomposition of T *if* $W = \{1, 2, \ldots, k\}$. We show that the eigenvalues of T are $0, \pi_1, \ldots, \pi_k$ (where $\pi_i = st(\kappa_i)$) and $T = \pi_1 P_1 + \ldots + \pi_k P_k$, using the fact that on H , the operators T and T' coincide. Let $x \in H$. Then

$$T(x) = \sum_{j \leq \mu} \kappa_j P_j(x) = \sum_{j=1}^{k} \kappa_j P_j(x) + \sum_{j=k+1}^{\mu} \kappa_j P_j(x)$$

$$\approx \sum_{j=1}^{k} \kappa_j P_j(x) .$$

By taking standard parts, we get the decomposition of T . Moreover, for any $x \in H_0$, we have $T(x) = \Sigma\, \pi_j P_j(x) = 0$ and therefore 0 is an eigenvalue of T . It remains to show that all eigenvalues of T

are among the π_j . This follows by an easy calculation from the uniqueness of direct sum compositions.

The decomposition of T *if* W = $\mathbb{N}$. We show that for all $x \in H$, $T(x) = \sum_{j=1}^{\infty} \pi_j P_j(x)$. It suffices to show that $s_n \to 0$, where

$$s_n = \| T(x) - \sum_{j=1}^{n} \pi_j P_j(x) \|$$

Now for all $n \in \mathbb{N}$, we have that

$$\sum_{j=1}^{n} \pi_j P_j(x) \approx \sum_{j=1}^{n} \kappa_j P_j(x) ,$$

and since $T(x) = T'(x)$ on H , we have that

$$T(x) = \sum_{j=1}^{\mu} \kappa_j P_j(x) ,$$

so that

$$(s_n)^2 \approx \| \sum_{j=n+1}^{\mu} \kappa_j P_j(x) \|^2$$

$$= \sum_{j=n+1}^{\mu} |\kappa_j|^2 \cdot \| P_j(x) \|^2$$

$$\leq |\kappa_{n+1}|^2 \cdot \sum_{j=1}^{\mu} \| P_j(x) \|^2$$

$$= |\kappa_{n+1}| \cdot \| x \|^2 .$$

Thus $s_n \leq | \pi_{n+1} | \cdot \| x \|$ for all $n \in \mathbb{N}$. But since $|\pi_j| \approx |\kappa_j| \approx 0$, for all $j \in {}^*\mathbb{N} - \mathbb{N}$, the result follows from Theorem 1.15.

For the remaining routine calculations required to verify the decomposition equations for the space H and for the identity operator I we refer to ROBINSON [1966], pp. 191-194. □

Comparison with a standard proof (HELMBERG [1969]).

The basic steps of the *nonstandard proof* are described at the beginning of the proof. The proof is direct and therefore contains no negations. It involves the first and third order quantifiers $(\forall n \in \mathbb{N})$, $(\forall n \in {}^*\mathbb{N} - \mathbb{N})$, $(\forall x \in H)$, $(\exists x \in H)$, $(\exists x \in \mathbb{R} \times \mathbb{R})$, and $(\exists E \in {}^*(p(p(H)))$, and uses the Transfer Principle, the Standard Part Theorem, and the Spectral Theorem for Hermitian operators on finite-dimensional unitary spaces. The *standard proof*, on the other hand, involves the idea of approximating the space by families of mutually orthogonal finite-dimensional spaces from below, and the verification that in the limit the spectral properties of Hermitian operators on these spaces are stable. We omit the details of this proof since the length and complexity of the argument makes it virtually impossible to juxtapose the two types of proof in a meaningful way. A crucial existence part of the standard argument is the proof that a compact Hermitian operator T on a Hilbert space H has an eigenvalue λ with the property that $|\lambda| = \|T\|$. For details we refer to HELMBERG [1969]. □

CHAPTER 4

INTERNAL SETS

Some standard applications of internality

4.1. *Theorem* (Cauchy sequence theorem). Every Cauchy sequence is bounded.

Proof. Let $\langle x_n \rangle_{n \in \mathbb{N}}$ be a Cauchy sequence and let ν be an arbitrary infinite integer. It suffices to show that x_ν is finite. Let

$$A = \{ n \in {}^*\mathbb{N} \mid \tilde{W} \models \rho(x_\nu, x_n) < 1 \} .$$

Then A is a definable subset of the internal set ${}^*\mathbb{N}$ and hence is internal, and since $\langle x_n \rangle_{n \in \mathbb{N}}$ is a Cauchy sequence, we have that ${}^*\mathbb{N} - \mathbb{N} \subset A$. Furthermore, $A \cap \mathbb{N} \neq \emptyset$ since ${}^*\mathbb{N} - \mathbb{N}$ is external and A is internal, so that there is a finite n such that $\tilde{W} \models \rho(x_\nu, x_n) < 1$. Hence x_ν is finite. □

4.2. *Theorem* (Convergence theorem for real sequences). If $\langle x_n \rangle_{n \in \mathbb{N}}$ is a sequence of real numbers and if $x_n \approx x_m$ *for all infinite* n and m, then $x_n \to x$ for some $x \in \mathbb{R}$.

Proof. Suppose x_n is finite for all infinite n. Then x_n has a standard part x, and since $x_\mu \approx x_\nu$ for all infinite μ and ν, we have $x = \mathrm{st}(x_\mu)$ for all $\mu \in {}^*\mathbb{N} - \mathbb{N}$. By Theorem 1.15, it suffices to show that x_n is finite for all $n \in {}^*\mathbb{N} - \mathbb{N}$. Thus let ν be an infinite integer. We want to show that x_ν is finite. For this purpose, we consider the set

$$A = \{ n \in {}^*\mathbb{N} \mid | x_n - x_\nu | < 1 \} .$$

The set A is internal since $A \subset {}^*\mathbb{N}$ and since A is defined by the formula ${}^*\mathrm{abs}\lceil {}^*d \lceil \langle {}^*x\lceil n, {}^*x\lceil \nu \rangle < 1$, with abs and d being the absolute value and difference functions, respectively (cf. the Internal Subset Theorem). For $n \in {}^*\mathbb{N} - \mathbb{N}$, we have $x_n \approx x_\nu$ and therefore $|x_n - x_\nu| < 1$. Thus ${}^*\mathbb{N} - \mathbb{N} \subset A$. But since A is internal and ${}^*\mathbb{N} - \mathbb{N}$ is external, ${}^*\mathbb{N} - \mathbb{N} \neq A$. Hence there is an element $n_0 \in A \cap \mathbb{N}$. Thus we have that $|x_\nu| = |x_{n_0} - (x_{n_0} - x_\nu)| \leq |x_{n_0}| + |x_{n_0} - x_\nu| < |x_{n_0}| + 1$. So x_ν is finite. □

4.3. *Theorem* (Open set theorem). A set G is open if and only if $\mu(p) \subset {}^*G$ for all $p \in G$.

Proof. Let G be open and let $p \in G$. Since $G \in \Omega_p$ and since $\mu(p) = \bigcap_{U \in \Omega_p} {}^*U$, it follows that $\mu(p) \subset {}^*G$.

Now let $p \in G$ be such that $\mu(p) \subset {}^*G$. Then there exists an internal set $D \in {}^*\Omega_p$ such that $D \subset \mu(p) \subset {}^*G$. Hence

$$\tilde{W} \models (\exists D \in {}^*\Omega_p)(D \subset {}^*G) .$$

By the Transfer Principle, we have that

$$\hat{S} \models (\exists D \in \Omega_p)(D \subset G) .$$

Thus we have an open neighbourhood $U_p \subset G$. Hence if $\mu(p) \subset {}^*G$ for all $p \in G$, there is an open neighbourhood $U_p \subset G$ for each $p \in G$ and it follows that $G = \bigcup_{p \in G} U_p$ is open. □

4.4. *Theorem* (Bounded sequence theorem). If $\langle x_n \rangle_{n \in {}^*\mathbb{N}}$ is an internal sequence of hyperreal numbers and if $| x_n | \leq M \in {}^*\mathbb{R}$ for all $n \in \mathbb{N}$, then there exists a $\nu \in {}^*\mathbb{N} - \mathbb{N}$ such that $| x_n | \leq M$ for all $n < \nu$.

Proof. Let $A = \{ n \in {}^*\mathbb{N} \mid | x_n | > M \}$. Then A is defined by the formula $| x_n | > M$ and hence internal, since any definable subset of an internal set is itself internal (by the Internal Subset Theorem).

If A is empty, then the theorem holds for all infinite integers. Suppose therefore that the set A is not empty. By the Nonstandard Induction Principle, A has a least element $\nu \in {}^*\mathbb{N}$. Thus we have for any $n < \nu$ that $| x_n | \leq M$. Since, by definition, A is disjoint from $\mathbb{N}$, we have $\nu \in {}^*\mathbb{N}$. □

4.5. *Theorem* (Infinitesimal prolongation theorem). If $s = \langle x_n \rangle_{n \in {}^*\mathbb{N}}$ is an internal sequence of hyperreal numbers and if $x_n \approx 0$ for all $n \in \mathbb{N}$, then there exists a $\nu \in {}^*\mathbb{N} - \mathbb{N}$ such that $x_n \approx 0$ for all $n < \nu$.

Proof. From the sequence $s = \{ \langle n, x_n \rangle \mid n \in {}^*\mathbb{N} \}$ we construct a new sequence

$$t = \langle t_n \rangle_{n \in {}^*\mathbb{N}} = \{ \langle x, n \cdot x_n \rangle \mid n \in {}^*\mathbb{N} \} .$$

By the Internal Function Theorem, this sequence is internal since s is internal and since $t_n \doteq (n \cdot (x\lceil n))$ is therefore a term of $*L$ for each n . Since $x_n \approx 0$ for all $n \in \mathbb{N}$, it follows that $t_n \approx 0$ for all $n \in \mathbb{N}$, and by Theorem 4.4, there therefore exists a $\nu \in {*\mathbb{N}} - \mathbb{N}$ with the property that $| t_n | \leq 1$ for all $n < \nu$. Hence we have that $| x_n | \leq | t_n/n | = 1/n \approx 0$ for all $n < \nu$. □

4.6. *Theorem* (ε-δ-continuity theorem). If $f : {*X} \to {*T}$ is internal and if $p \approx q \Rightarrow f(p) \approx f(q)$ for all $q \in {*X}$, then f is ε-δ-*continuous* in the sense that for all *real* $\varepsilon > 0$ there exists a *real* $\delta > 0$ such that $\rho(p,q) < \delta \Rightarrow \rho(f(p),f(q)) < \varepsilon$.

Proof. Let ε be an positive real number and

$$A = \{ n \in {*\mathbb{N}} \mid \tilde{W}. \models (\forall p \in {*X})(\rho(p,q) < 1/n \Rightarrow \rho(f(p),f(q)) < \varepsilon) \} .$$

The set A is clearly a definable subset of $*\mathbb{N}$ and is therefore internal. By the continuity of f at q on $*X$, we have ${*\mathbb{N}} - \mathbb{N} \subset A$, and since A is internal and ${*\mathbb{N}} - \mathbb{N}$ is external it follows that $A \cap \mathbb{N} \neq \emptyset$. Let $\delta = 1/k$ for some $k \in A \cap \mathbb{N}$. Then for any $p \in {*X}$, we have that $\rho(p,q) < \delta \Rightarrow \rho(f(q),f(q)) < \varepsilon$. □

4.7. *Theorem* (Approximation theorem). Let X be a compact space and f a continuous internal function on $*X$ and $f(x)$ near standard for all $x \in X$, and let F be defined by $F(x) = st(f(x))$. Then F is continuous on X and $F(x) \approx f(x)$ for all $x \in {*X}$.

Proof. Let $x \in X$. Since f is continuous and internal, f is

ε-δ-continuous at x by Theorem 4.6. Thus for any real $\varepsilon > 0$ there exists a real $\delta > 0$ such that for all $x' \in {}^*X$, $\rho(x',x) < \delta$ implies $\rho(f(x'),f(x)) < \varepsilon/2$. So let $x' \in X$ be such that $\rho(x',x) < \delta$. By the definition of F, we have that $f(x') \approx F(x')$ and $f(x) \approx F(x)$. Hence $\rho(F(x'),F(x)) \leq \rho(F(x'),f(x')) + \rho(f(x'),f(x)) + \rho(f(x),F(x)) < \varepsilon$. Therefore F is continuous at x, and since x was an arbitrary point of X, it follows that F is continuous on X.

Now let $x \in {}^*X$. Since X is compact, $x \approx x' \in X$ and by the continuity of f we have that $f(x) \approx f(x')$. By the definition of F, $f(x') \approx F(x')$ and by continuity it therefore follows that $F(x) \approx F(x')$, so that $F(x) \approx f(x)$. □

The Arzelà-Ascoli theorem

4.8. *Theorem* (Arzelà-Ascoli theorem). Any equicontinuous sequence of functions from a compact space X to a compact space T has a uniformly convergent subsequence.

Proof (Nonstandard). Let $\langle f_n \rangle_{n \in \mathbb{N}}$ be an equicontinuous family of functions and let ν be a fixed infinite integer. From the definition of equicontinuity it follows that for any positive real number ε there exists a positive real number δ such that

$$\hat{S} \models (\forall n \in \mathbb{N})(\forall p,q \in X)(\rho(p,q) < \delta \Rightarrow \rho(f_n(p), f_n(q)) < \varepsilon) .$$

By an application of the Transfer Principle we therefore have that

$$\tilde{W} \models (\forall n \in {}^*\mathbb{N})(\forall p,q \in {}^*X)(\rho(p,q) < \delta \Rightarrow \rho(f_n(p), f_n(q)) < \varepsilon) .$$

Since ε was arbitrary, it follows that $p \approx q$ implies $f_n(p) \approx f_n(q)$. Thus f_ν is continuous on $^*X$ and by the transitivity of $\tilde{W}$ it is an internal function. Thus Theorem 4.7 applies to f_ν, i.e., if we let $F(p) = st(f_\nu(p))$ for any $p \in X$, we can conclude that $f_\nu(p) \approx F(p)$ for all $p \in {}^*X$.

Thus for any fixed $k, m \in \mathbb{N}$, we have

$$\tilde{W} \models (\exists q \in {}^*\mathbb{N})((q > m) \wedge (\forall x \in {}^*X)(\rho(F(x), f_q(x)) < 1/k))) .$$

By applying the Transfer Principle once again, we have that

$$\hat{S} \models (\exists q \in \mathbb{N})((q > m) \wedge (\forall x \in X)(\rho(F(x), f_q(x)) < 1/k))) .$$

Hence there exist numbers $n_1 = 1$ and $n_{i+1} = q = q(n_i, i+1)$ determining a uniformly convergent subsequence $\langle f_{n_i} \rangle_{i \in \mathbb{N}}$. □

Comparison with a standard proof (MUNROE [1965]).

Structural difference.

Basic steps of the *nonstandard proof*.

(1) The proof that for any infinite ν the function ${}^{\bullet}f_{\nu}$ is continuous on $^*X$.

(2) The proof that $f_{\nu}(p)$ is finite for all $p \in X$.

(3) The definition of the standard part function F .

(4) The verification that the convergence of f_{n_i} to F is uniform.

Basic steps of the *standard proof*.

(1) The introduction of a countably dense subset of X .

(2) The construction of a Cauchy sequence of functions on X .

(3) The construction of a cover of open balls of X .

(4) The extraction of a finite subcover.

(5) The verification of an ε estimate.

Logical difference.

Neither proof contains a negation. The *nonstandard proof* uses the Transfer Principle and contains the first order quantifiers $(\forall n \in \mathbb{N})$, $(\exists n \in \mathbb{N})$, $(\forall n \in {}^*\mathbb{N})$, $(\exists n \in {}^*\mathbb{N})$, $(\forall x \in \mathbb{R})$, $(\exists x \in \mathbb{R})$, $(\forall x \in X)$, and $(\forall x \in {}^*X)$. The *standard proof*, on the other hand, involves first, second, and third order quantifiers: $(\forall n \in \mathbb{N})$, $(\exists n \in \mathbb{N})$, $(\forall x \in \mathbb{R})$, $(\forall x \in X)$, $(\exists \mathcal{U} \in p(X))$, $(\exists C \in p(p(X)))$, and $(\forall f \in \mathcal{F} \in p(p(p(X \times T))))$.

Mathematical difference.

The *nonstandard proof* uses the transitivity of $\tilde{W}$, the Standard Part Theorem, and the nonstandard criterion for compactness, whereas the *standard proof* uses the fact that compact metric spaces are separable, that bounded sequences of reals have convergent subsequences, and that in the topology of uniform convergence a sequence converges if and only if it converges uniformly. □

The Riemann-Lebesque integration theorem

In this section, we compare the standard and nonstandard proofs of the fact that Riemann integrable functions are measurable and that the Riemann and Lebesque integrals of such functions coincide. For this purpose we first present the rudiments of the standard and non-standard Riemann and Lebesque integration theories and prove the required characterization theorems.

4.9. *Definition* (Standard). A *partition* of an interval $[a,b]$ is a sequence $\langle x_0, \dots, x_n \rangle_{n \in \mathbb{N}}$ of points of the interval $[a,b]$ in which $a = x_0 < \dots < x_n = b$.

4.10. *Definition* (Nonstandard). A *fine partition* of an interval $[a,b]$ is a sequence $\langle x_0, \dots, x_n \rangle_{n \in {}^*\mathbb{N}}$ of points of the interval ${}^*[a,b]$ in which $a = x_0 < \dots < x_n = b$ and in which $x_i \approx x_{i+1}$ for all i, $0 \leq i \leq n-1$.

Let $f : [a,b] \to [\alpha,\beta]$ be a given function and $\langle x_0, \dots, x_n \rangle$ either a partition or a fine partition P' of $[a,b]$ and let the sequence $\langle y_0, \dots, y_n \rangle$ be either a partition or a fine partition Q of the interval $[\alpha,\beta]$. Then we put

$$M_i = \sup f(x) \quad \text{on } [x_{i-1}, x_i]$$
$$m_i = \inf f(x) \quad \text{on } [x_{i-1}, x_i]$$
$$E_i = \{ x \mid y_{i-1} \leq f(x) < y_i \}$$

and let

$$U(f,P) = \sum_{i=1}^{n} M_i \cdot (x_i - x_{i-1})$$

$$L(f,P) = \sum_{i=1}^{n} m_i \cdot (x_i - x_{i-1})$$

$$U(f,Q) = \sum_{i=1}^{n} y_i \cdot m(E_i)$$

$$L(f,Q) = \sum_{i=1}^{n} y_{i-1} \cdot m(E_i) ,$$

where $m(E_i)$ denotes the Lebesque measure of E_i. Then it is easy to see that if P is an *internal* fine partition, the sums $U(f,P)$ and $L(f,P)$ are finite and for any two distinct internal fine partitions P and P' we have $U(f,P) \approx U(f,P')$ and $L(f,P) \approx L(f,P')$. The sums $U(f,Q)$ and $L(f,Q)$ have the same property and make sense provided that the values $m(E_i)$ exist. For details, we refer to ROBINSON [1966].

Finally, we write $\int_a^b f\,dx$ for the *Riemann integral* and $\int_a^b f\,dm$ for the *Lebesque integral* of f on $[a,b]$.

4.11. *Definition* (Standard). $\int_a^b f\,dx = \inf U(f,P) = \sup L(f,P)$ over all partitions P of $[a,b]$, whenever this value exists.

4.12. *Definition* (Nonstandard). $\int_a^b f\,dx = st(U(f,P)) = st(L(f,P))$, for any internal fine partition P of $[a,b]$, whenever this value exists.

4.13. *Theorem* (Characterization theorem). Definitions (4.11) and (4.12) are equivalent.

Proof. Suppose that the standard Riemann integral $\int_a^b f\,dx$ exists. Then for any positive real number ε there exists a positive real number δ such that

$$\widehat{S} \models (\forall P \in \mathrm{Part}([a,b]))(\|P\| < \delta \Rightarrow U(f,P) - L(f,P) < \varepsilon),$$

where Part([a,b]) denotes the set of partitions of $[a,b]$ and $\|P\|$ denotes the length of the longest subinterval of $[a,b]$ determined by P. By the Transfer Principle, we have that

$$\widetilde{W} \models (\forall P \in {}^*\mathrm{Part}([a,b]))(\|P\| < \delta \Rightarrow U(f,P) - L(f,P) < \varepsilon).$$

Let P be any internal fine partition of $[a,b]$. Then $\|P\| \approx 0 < \delta$ for all δ and hence for all positive real ε, i.e., $U(f,P) \approx L(f,P)$, and therefore $\mathrm{st}(U(f,P)) = \mathrm{st}(L(f,P))$.

The proof of the converse is obtained by reversing the steps in the previous argument. □

4.14. *Definition* (Standard). $\int_a^b f\,dm = \inf(U(f,Q)) = \sup(L(f,Q))$ over all partitions Q of $[a,b]$, whenever this value exists.

4.15. *Definition* (Nonstandard). $\int_a^b f\,dm = \mathrm{st}(U(f,Q)) = \mathrm{st}(L(f,Q))$, for any internal fine partition Q of $[\alpha,\beta]$, whenever this value exists.

4.16. *Theorem* (Characterization theorem). Definitions (4.14) and (4.15) are equivalent.

Proof. Analogous to that of Theorem 4.13. □

4.17 *Theorem* (Riemann-Lebesque integration theorem). Let $f : [a,b] \to [\alpha,\beta]$ be a Riemann integrable function. Then f is continuous except on a set of Lebesque measure zero, hence measurable, and $\int_a^b f\,dx = \int_a^b f\,dm$.

Proof (Nonstandard). We require the following data:

(1) An internal fine partition $P = < x_0, \ldots, x_\omega >_{\omega \in {}^*\mathbb{N}}$ of $[a,b]$;

(2) An internal fine partition $Q = < u_0, \ldots, u_\mu >_{\mu \in {}^*\mathbb{N}}$ of $[\alpha,\beta]$, refining the sequences $< v_0, \ldots, v_\nu >$ and $< w_0, \ldots, w_\eta >$, where $v_\nu = \beta$, and $v_0 < v_1 < \ldots < v_{\nu-1}$ are the distinct infima m_i of f determined by the partition P , and where $w_0 = \alpha$, and $w_1 < w_2 < \ldots < w_\eta$ are the distinct suprema M_i of f determined by P ;

(3) $A_n = \{ x \mid (\forall\delta>0)(\exists y)(|x-y| < \delta \Rightarrow |f(x)-f(y)| > 1/n \}$;

(4) $T = \{ i \in {}^*\mathbb{N} \mid (\exists n \in N)(|M_i - m_i| > 1/n \}$;

(5) $E = \{ x \mid$ the function f is discontinuous at $x \}$;

(6) The function $g : [a,b] \to [\alpha,\beta]$ defined by $g(x) = m_i$ on $[x_{i-1},x_i)$;

(7) The function $h : [a,b] \to [\alpha,\beta]$ defined by $h(x) = M_i$ on $[x_{i-1},x_i)$.

We first show that $\sum_{i \in T} (x_i - x_{i-1}) \approx 0$, that therefore $m(A_n) = 0$, and that therefore $m(E) = 0$, and that the function f is therefore measurable.

If T is not empty, there exists an $m \in \mathbb{N}$ such that

$$0 \le \frac{1}{m} \cdot \sum_{i \in T} (x_i - x_{i-1}) = \sum_{i \in T} \frac{1}{m} \cdot (x_i - x_{i-1})$$

$$< \sum_{i \in T} (M_i - m_i)(x_i - x_{i-1})$$

$$\le \sum_{i=1}^{\omega} (M_i - m_i)(x_i - x_{i-1})$$

$$\approx 0 ,$$

provided that f is Riemann integrable, and since m is finite, we have $\sum_{i \in T} (x_i - x_{i-1}) \approx 0$. If T is empty, the result holds trivially.

Now let $x \in {}^*A_n$, with $x_i < x < x_{i+1}$. Then it follows from the definition of A_n by an application of the Transfer Principle that there is a $y \in [x_i, x_{i+1}]$ such that $|f(x) - f(y)| > 1/n$. Hence $i \in T$. Similarly for $x = x_i$ or $x = x_{i+1}$. Hence it is clear that ${}^*A_n \subset \bigcup_{i \in T} [x_i, x_{i+1}]$. For the Lebesque outer measure $m^☆$ we therefore have

$$0 \le m^☆(A_n) = m^☆({}^*A_n) \le m\left(\bigcup_{i \in T} [x_i, x_{i+1}]\right) = \sum_{i \in T} (x_i - x_{i-1}) \approx 0 .$$

Hence $m^☆(A_n) = 0$ since $m^☆(A_n)$ is standard. The set A_n is therefore measurable and has Lebesque measure zero.

Since $E = \bigcup_{n \in \mathbb{N}} A_n$ is a countable union of sets of measure zero, it itself has Lebesque measure zero.

It remains to show that f is measurable, i.e., we must show that the inverse of every open subset of $[\alpha, \beta]$ is a measurable subset of $[a,b]$. But this follows at once from the fact that open sets are measurable, that subsets of Lebesque measure zero are Lebesque measurable and the fact that for every open subinterval (y,y') of $[\alpha,\beta]$ we have

$$f^{-1}[(y,y')] = f^{-1}[(y,y')] \cap ([a,b]-E) \cup f^{-1}[(y,y')] \cap E ,$$

with $f^{-1}[(y,y')]$ open by the continuity of f.

We complete the proof of the theorem by showing that the Riemann and Lebesque integrals f coincide. It suffices to show that

$$\int_a^b g\,dm \approx \int_a^b f\,dx \approx \int_a^b h\,dm$$

since the inequality $g \le f \le h$ then entails that

$$\int_a^b f\,dx \approx \int_a^b g\,dm \le \int_a^b f\,dm \le \int_a^b h\,dm \approx \int_a^b f\,dx .$$

Since the functions g and h are clearly bounded and measurable, their Lebesque integrals exist and we have that

$$\begin{aligned}
\int_a^b g\,dm &= st(U(g,Q)) \\
&= st(L(g,Q)) \\
&\approx \sum_{i=1}^{\mu} u_{i-1}\cdot m(\{\, x \mid u_{i-1} \le g(x) < u_i \,\}) \\
&\approx \sum_{i=1}^{\mu} u_i\cdot m(\{\, x \mid u_{i-1} \le g(x) < u_i \,\}) \\
&= \sum_{k=0}^{\eta-1} w_k\cdot m(\bigcup \{\, [x_{i-1},x_i) \mid M_i = w_k \,\}) \\
&= \sum_{k=0}^{\eta-1} w_k\cdot \sum_{i\in I} (x_i - x_{i-1}) \\
&= \sum_{k=0}^{\eta-1} \sum_{i\in I} M_i\cdot(x_i - x_{i-1}) \\
&= \sum_{k=1}^{\omega} M_i\cdot(x_i - x_{i-1}) \approx st(U(f,P)) = \int_a^b f\,dx ,
\end{aligned}$$

where I consists of all i $(0 \leq i \leq \omega)$ with the property that $M_i = w_i$, and where it is assumed that f is Riemann integrable. □

Comparison with a standard proof (RUDIN [1976]).

Structural difference.

Basic steps of the *nonstandard proof.*

(1) The choice of a fixed internal fine partition of $[a,b]$.

(2) The choice of a fixed internal fine partition of $[\alpha,\beta]$.

(3) The definition of a subset of $^*\mathbb{N}$ indexing the jumps of f .

(4) The verification that the sum of the lengths of the subintervals of $^*[a,b]$ involving jumps of f is infinitesimal.

(5) The definition of a sequence of sets of points of $[a,b]$ at which f violates the continuity criterion by more than $1/n$.

(6) The verification that the terms of this sequence have Lebesque measure zero.

(7) The verification of the measurability condition for f .

(8) The definition of hyperfinitely piecewise constant functions g and h bounding f .

(9) The proof that the Lebesque integrals of g and h are infinitely close to the Riemann integral of f .

Basic steps of the *standard proof.*

(1) The introduction of a sequence of refining partitions.

(2) The proof that the lower and upper Riemann integrals of f are limits over this sequence of partitions.

(3) The definition of approximating step functions g and h analogous to those in the nonstandard proof.

(4) The proof that the Lebesque integral of g equals the lower Riemann integral of f and that the Lebesque integral of h equals the upper Riemann integral of f.

(5) The verification that the Lebesque integrals of g and h coincide if and only if $g = h$ a.e..

(6) The proof that the equality of g and h almost everywhere entails the measurability of f.

(7) The verification that the Riemann integrability of f entails the continuity of f almost everywhere.

Logical difference.

Both proofs are direct and contain no negations. The *nonstandard proof* uses the Transfer Principle and involves the first order quantifiers $(\exists n \in \mathbb{N})$ and $(\exists x \in {}^*\mathbb{R})$. The *standard proof*, on the other hand requires the second order quantifier $(\exists f \in p([a,b] \times \mathbb{R}))$ and the third order quantifier $(\exists s \in p(\mathbb{N} \times p([a,b])))$.

Mathematical difference.

Both proofs use the fact that the countable union of sets of measure zero is a set of measure zero. The *nonstandard proof* also uses the fact that bounded measurable functions are Lebesque integrable and that sets whose outer measure is zero have measure zero. The *standard proof* also uses the effect of refining partitions on the ordering of upper and lower Riemann sums and the Monotone Convergence Theorem. □

The Loeb representation of integrals

In this section we sketch a construction due to LOEB [1979] of the integrals of bounded measurable functions with respect to finitely additive bounded measures in terms of the standard parts of hyperfinite sums. This approach generalizes the nonstandard approach to Riemann and Lebesque integration described in the previous section. We omit all proofs and refer the reader to LOEB [1979] and [1972].

Let X be a fixed set, $\mathcal{A}$ a fixed σ-algebra of subsets of X, with both X and $\mathcal{A}$ infinite. Let $\mathcal{P}$ denote the collection of all finite $\mathcal{A}$-partitions of X, and let Q, M, and MB denote the following sets of functions:

$$Q = \{ f \mid f : X \to \mathbb{R}^{\#} \}$$

$$M = \{ f \in Q \mid f \text{ is } \mathcal{A}\text{-measurable} \}$$

$$MB = \{ f \in M \mid f \text{ is bounded} \} .$$

Next we choose a fixed initial segment $I = \{ i \in {}^*\mathbb{N} \mid 1 \leq i \leq \omega \in {}^*\mathbb{N} - \mathbb{N} \}$ and let E be the set of all *internal* functions $f : I \to {}^*\mathbb{R}$. We invoke the Concurrence Theorem to construct a hyperfinite partition of ${}^*X$, based on the family $\mathcal{P}$ above, and assume as given a choice function $c_{\mathcal{P}} : I \to {}^*X$. Using this function, we define a function

$$T : MB \to E$$

by the equation $T(f)(i) = {}^*f(c_{\mathcal{P}}(i))$. Finally we let $\Phi(X,\mathcal{A})$ be the set of finitely additive bounded set functions $\mu : \mathcal{A} \to \mathbb{R}$ and define $U : \Phi(X,\mathcal{A}) \to E$ by $U(\mu)(i) = {}^*\mu(A_i)$, with A_i belonging to the hyperfinite partition constructed above. Then we have the following theorem:

4.18. *Theorem* (Loeb representation theorem). If $\mu \in \Phi(X,A)$, $f \in MB$, and $B \in A$, then

$$\int_B f\, d\mu = \mathrm{st}(\Sigma\, {}^*f(c_p(i)) \cdot {}^*\mu(A_i))$$

with the summation taken over the set $I_B = \{ i \in I \mid A_i \subset {}^*B \}$. In particular we have that

$$\int_X f\, d\mu = T(f) \diamond U(\mu) \ ,$$

where $\diamond$ denotes the standard inner product on E . □

APPENDIX 1

THE STANDARD UNIVERSES $\hat{S}$

For any given set S, we define a new inductive set $\hat{S}$, called the *standard universe* generated by S. The set S varies with the context in which we are working. It is usually the standard set of real numbers, but it may also be the set of vectors of the standard Hilbert space ℓ^2, or some other standard set whose standard proof theory we wish to explore by nonstandard means.

With the help of the power set operator p we define the following chain $S_0 \subset S_1 \subset \ldots \subset S_n \subset \ldots$ of sets:

$$S_0 = S \; ; \quad S_{n+1} = S_n \cup p(S_n) \; ,$$

and we let

$$\hat{S} = \bigcup_{n \in \mathbb{N}} S_n \; .$$

We call the elements of the generating set S *individuals* and refer to the elements of $\hat{S} - S$ as *sets*. Among the basic properties of $\hat{S}$ are the facts that $\emptyset \in \hat{S}$, that $S \subset \hat{S}$, that $\hat{S}$ is transitive, and that if $x, y \in \hat{S}$, then $\{x, y\} \in \hat{S}$. From these facts we can prove that $\hat{S}$ is closed under the formation of subsets, power sets, ordered pairs, cartesian products, functional application, and function formation, and under infinitary unions and products over index sets belonging to $\hat{S}$. For details, we refer the reader to DAVIS [1977].

APPENDIX 2

THE NONSTANDARD UNIVERSES $\tilde{W}$

In this appendix, we construct the *nonstandard universes* $\tilde{W}$ determined by the standard universes $\hat{S}$. For this purpose, we let I be some fixed infinite index set and let F be a nonprincipal ultrafilter on I. We say that a property of elements indexed by I *holds almost everywhere* (a.e.) or *holds for almost all i*, if the set of indices for which the property holds belongs to the filter F. The construction of $\tilde{W}$ proceeds in three stages: First we construct a new set of individuals W, then we generate the standard universe $\hat{W}$, and finally we single out the subset $\tilde{W}$ of $\hat{W}$.

We begin by defining the following subset Z of the function space $\hat{S}^I$:

$$Z = \bigcup_{n \in \mathbb{N}} Z_n ,$$

where for each $n \in \mathbb{N}$, $Z_n = \{ f : I \to \hat{S} \mid f(i) \in S_n \text{ a.e.} \}$.

There is a natural embedding of the standard universe $\hat{S}$ into the function space Z: We identify the element $x \in \hat{S}$ with the constant function $\mathrm{const}(x)$, so that $\mathrm{const}(x)(i) = x$ for all $i \in I$.

For each $f \in Z_0$, we let $\bar{f} = \{ g \in Z_0 \mid g \equiv f \}$, where $g \equiv f$ if and only if $g(i) = f(i)$ a.e., so that Z_0 is divided into disjoint equivalence classes $\bar{f}$ by the equivalence relation $\equiv$. We now let $W = \{ \bar{f} \mid f \in Z_0 \}$ and form the standard universe $\hat{W}$.

We continue by induction. With each $f \in Z_n$ we associate a corresponding element $\bar{f} \in W_n$ and we then define the nonstandard universe $\tilde{W}$ determined by the set S, the index set I and the nonprincipal ultrafilter F by

$$\tilde{W} = \{ \bar{f} \mid f \in Z \} .$$

If $f \in Z_0$, then the element $\bar{f} \in W_0$ has already been defined. Thus suppose that $f \in Z_{n+1}$ and that $\bar{f}$ has not yet been defined. Then we let $\bar{f} = \{ \bar{g} \mid g \in Z_n \text{ and } g(i) \in f(i) \text{ a.e.} \}$.

The members of $\tilde{W}$ will be called *internal* and those of $\hat{W} - \tilde{W}$ will be called *external*. In particular, the elements of $\tilde{W} - W$ will be called *internal sets* and, as in the standard case, the elements of W will be called *individuals*. This terminology is consistent with the standard terminology of Appendix 1 since the standard universe $\hat{S}$ is faithfully embedded in the nonstandard universe $\tilde{W}$ by the mapping $x \in \hat{S} \to \overline{\text{const }(x)} \in \tilde{W}$.

The following properties of $\tilde{W}$ are basic: $\emptyset \in \tilde{W}$, $W \subset \tilde{W}$, $\tilde{W}$ is transitive, and if $x, y \in \tilde{W}$, then $\{x, y\} \in \tilde{W}$. Hence $\tilde{W}$ is closed under a variety of set-theoretical constructions. For details, we refer to DAVIS [1977].

For the purposes of this thesis, the following are the most frequently required set-theoretical properties of $\tilde{W}$:

(1) $\bar{f} \in \bar{g}$ if and only if $f(i) \in g(i)$ a.e.

(2) $\bar{f} = \bar{g}$ if and only if $f(i) = g(i)$ a.e.

(3) $\bar{h} = < \bar{f}, \bar{g} >$ if and only if $h(i) = < f(i), g(i) >$ a.e.

(4) $\bar{h} = (\bar{f} \lceil \bar{g})$ if and only if $h(i) = (f \lceil g)$ a.e.

(5) If $x \in y \in \tilde{W}$, then $x \in \tilde{W}$.

APPENDIX 3

THE EMBEDDING $* : \hat{S} \to \tilde{W}$

By virtue of the existential richness of the languages L and $*L$, each individual $x \in S$, each set $A \in \hat{S} - S$, and the standard universe $\hat{S}$ itself, can be faithfully embedded in the nonstandard universe $\tilde{W}$. For this purpose we appeal to model theory and call a set $A \in \hat{S} - S$ *definable* if there exists a monadic formula $\alpha(x) \in L$ with the property that the set $A = \{ b \in \hat{S} \mid \hat{S} \models \alpha(b) \}$. We note at once that every set $A \in \hat{S} - S$ is definable since it suffices to take $\alpha(x)$ to be the formula $(x \in A)$. Moreover, the standard universe is definable since we can take $\alpha(x)$ to be the formula $(x = x)$.

Thus we put $*x = \bar{x}$ for all $x \in S$,

$$*A = \{ b \in \tilde{W} \mid \tilde{W} \models *\alpha(b) \}$$

$$*(\hat{S}) = \tilde{W} .$$

We note that for any $A \in \hat{S} - S$,

$$\begin{aligned} *A &= \{ \bar{g} \in \tilde{W} \mid \tilde{W} \models \bar{g} \in \bar{A} \} \\ &= \{ \bar{g} \in \tilde{W} \mid \bar{g} \in \bar{A} \} \\ &= \bar{A} . \end{aligned}$$

For the purposes of this thesis, the following are the most frequently required properties of the star map $* : \hat{S} \to \tilde{W}$:

(1) If $A \subset S$, then $A \subset *A$ and $*A \cap S = A$.

(2) For all $x, y \in \hat{S}$, $x = y$ if and only if $*x = *y$.

(3) For all $x, y \in \hat{S}$, $x \in y$ if and only if $*x \in *y$.

(4) For all $x, y \in \hat{S}$, $*\langle x, y \rangle = \langle *x, *y \rangle$.

(5) For all $x, y \in \hat{S}$, $*(x \lceil y) = (*x \lceil *y)$.

(6) For all definable *subsets* of $\hat{S}$,

(a) $*(A \cup B) = *A \cup *B$.

(b) $*(A \cap B) = *A \cap *B$.

(c) $*(A - B) = *A - *B$.

(d) $*(A \times B) = *A \times *B$.

(e) $*\emptyset = \emptyset$.

(7) If $A, B \in \hat{S}$ and $A \subset B$, then $*A \subset *B$.

(8) If $B \in \hat{S}$ and $A \in \tilde{W}$ and $A \subset *B$, then $A \in *p(B)$,
where p denotes the power set operator.

(9) If $f \in \hat{S}$ is a function and $A \subset \text{dom}(f)$, then $*(f[A]) = *f[*A]$.

APPENDIX 4

THE LANGUAGES L AND $*L$

Here we specify the language L corresponding to a given standard universe $\hat{S}$ and the language $*L$ corresponding to the nonstandard universe $\tilde{W}$ determined by $\hat{S}$.

The language L consists of an alphabet, terms, formulas, and sentences. We first describe the alphabet of L.

(1) The *basic symbols* of L are $=, \in, \neg, \wedge, \exists, (,), <, >$, and $\lceil$.

(2) The *variables* of L are $x_1, x_2, \ldots, x_n, \ldots$ (countably many).

(3) The *constants* of L are the elements of $\hat{S}$.

A4.1. *Definition.* The *terms* of L are described inductively as follows:

(1) Every variable of L is a term.

(2) Every constant of L is a term.

(3) If s and t are terms, then so is $< s, t >$.

(4) If s and t are terms, then so is $(s \lceil t)$.

A4.2. *Definition.* The *formulas* of L are described inductively as follows:

(1) If s and t are terms of L, then $(s = t)$ is a formula.

(2) If s and t are terms of L, then $(s \in t)$ is a formula.

(3) If A is a formula, then $\neg A$ is a formula.

(4) If A and B are formulas, then $(A \wedge B)$ is a formula.

(5) If t is a term of L not containing the variable x and A is a formula of L, then $(\exists x \in t)A$ is a formula.

A4.3. *Definition.* A *sentence* of L is a formula of L containing no free variable.

A4.4. *Definition.* A *closed term* of *L* is a term of *L* containing no variables.

For closed terms, we define the following semantics in the standard universe $\hat{S}$ (calling the element Val (t) in $\hat{S}$ corresponding to the closed term t of *L* its *value* in $\hat{S}$):

(1) Val (b) = b for all $b \in \hat{S}$.

(2) Val (< s , t >) = < Val (s) , Val (t) > .

(3) Val ((s⌈t)) = (Val (s) ⌈ Val (t)) .

A4.5. *Definition.* We define the relation $\hat{S} \models \alpha$ by induction on the construction of sentences:

(1) $\hat{S} \models (s = t)$ if and only if Val (s) = Val (t) .

(2) $\hat{S} \models (s \in t)$ if and only if Val (s) $\in$ Val (t) .

(3) $\hat{S} \models \neg A$ if and only if it is not the case that $\hat{S} \models A$.

(4) $\hat{S} \models (A \wedge B)$ if and only if $\hat{S} \models A$ and $\hat{S} \models B$.

(5) $\hat{S} \models (\exists x \in t)\alpha(x)$ if and only if $\hat{S} \models \alpha(c)$ for some constant c.

The construction of the nonstandard language **L* and the definitions of terms, formulas, sentences, closed terms, the valuation function, and of the relation $\tilde{W} \models \alpha$ are analogous. The initial difference between the languages *L* and **L* lies in the fact that in *L the set of constants* is $\hat{S}$, whereas in **L the set of constants* is $\tilde{W}$. This changes the sets of terms, formulas, and sentences in the obvious way.

A4.5. *Definition.* If t is a term of *L*, then *t is the term of **L* obtained by replacing every constant c in t by the constant $\bar{c}$:

APPENDIX 5

SOME BASIC NONSTANDARD RESULTS

A5.1. *Theorem* (Transfer principle). If α is a sentence in the standard language L and $*\alpha$ its translation into the nonstandard language $*L$, then $\widetilde{W} \models *\alpha$ if and only if $\widehat{S} \models \alpha$. □

A5.2. *Theorem* (Nonstandard induction principle). If A is a nonempty internal subset of $*\mathbb{N}$, then A has a least element. □

A5.3. *Theorem* (Internality theorem). If A is internal and B is definable, then $A \cap B$ is internal. □

A5.4. *Corollary* (Internal subset theorem). If B is a definable subset of an internal set A, then B is internal. □

A5.5. *Theorem* (Internal product theorem). If A and B are internal sets, then their cartesian product $A \times B$ is internal. □

A5.6. *Theorem* (Internal function theorem). If A and B are internal sets, $f : A \to B$ a function in $\widehat{W}$, and $\mu(x)$ a monadic term of the nonstandard language $*L$ such that for each $a \in A$, $f(a) = Val_*(\mu(a))$, then f is internal. □

A5.7. *Theorem* (External set existence theorem). The set $\mathbb{N}$ of standard natural numbers is external. □

A5.8. *Theorem* (Nonarchimedian order theorem). If $n \in \mathbb{N}$ and $\nu \in *\mathbb{N} - \mathbb{N}$, then $n < \nu$ in the ordering of $*\mathbb{N}$. □

A5.9. *Theorem* (Characterization theorem). A hyperreal number $x \in {}^*\mathbb{R}$ is infinitesimal if and only if $|x| < \frac{1}{n}$ for all $n \in \mathbb{N}$. □

A5.10. *Theorem* (Finite arithmetic theorem).

(1) If x and y are finite, then so are $x+y$ and $x \cdot y$.

(2) If x is finite and nonzero, then $\frac{1}{x}$ is finite.

(3) If x and y are infinitesimal, then so are $x+y$ and $x \cdot y$.

(4) If x is finite and ~~y~~ ~~is~~ infinitesimal, then $x \cdot y$ is infinitesimal.

(5) If x, x', y, y' are finite and $x \approx x'$ and $y \approx y'$, then $x+y \approx x'+y'$ and $x \cdot y \approx x' \cdot y'$.

(6) If x is finite and nonzero and $x \approx y$, then $\frac{1}{x} \approx \frac{1}{y}$. □

A5.11. *Theorem* (Standard part theorem).

(1) If $x \in \mathbb{R}$, then $st(x) = x$.

(2) If x and y are finite, then

(a) $st(x+y) = st(x) + st(y)$.

(b) $st(x \cdot y) = st(x) \cdot st(y)$.

(c) $st(x) \leq st(y)$ if and only if $x \leq y$. □

APPENDIX 6

SOME NONSTANDARD DEFINITIONS

A6.1. *Definition* (General). A relation r in $\hat{S}$ or $\tilde{W}$ is *concurrent* in $\hat{S}$ or $\tilde{W}$ if whenever $\langle a_1, \ldots, a_n \rangle \in \mathrm{dom}(r)$, there exists an element b such that $\langle a_i, b \rangle \in r$ for all $i = 1, 2, \ldots, n$.

A6.2. *Definition* (Nonstandard). In any metric space $\langle X, \rho \rangle$, a point $p \in {}^*X$ is *finite* if $\rho(p,q) \in \mathbb{R}$ for some point $q \in X$.

A6.3. *Definition* (Nonstandard). In any topological space $\langle X, \Omega \rangle$, the *monad* $\mu(p)$ of a point $p \in X$ is $\bigcap_{U \in \Omega_p} {}^*U$, where Ω_p denotes the ~~set of all~~ open sets containing p.

A6.4. *Definition* (Nonstandard). Two points p and q are related by $p \approx q$ if and only if $p \in \mu(q)$, in which case we say that p *is infinitely close to* q.

A6.5. *Definition* (Nonstandard). In any topological space $\langle X, \Omega \rangle$, a point $p \in {}^*X$ is *near standard* (with respect to X), if $p \approx q$ for some point $q \in X$.

A6.6. *Definition* (Nonstandard). In any topological space a point is *remote* (from the space) if it is not a near standard point of the space.

A6.7. *Definition* (Nonstandard). In any topological space the *standard part* of a near standard point p is the unique point q such that $p \in \mu(q)$.

BIBLIOGRAPHY

ANDERSON, R. M.

[1976] *A non-standard representation for brownian motion and Itô integration*, Israel Journal of Mathematics, 25, pp. 15-46.

[1977] Star-finite probability theory, Ph.D. dissertation, Yale University.

BELL, J. L. and SLOMSON, A. B.

[1969] Models and ultraproducts, North-Holland, Amsterdam.

BERNSTEIN, A. R. and WATTENBERG, F.

[1969] *Nonstandard measure theory*, in: Applications of model theory to algebra, analysis, and probability (W. A. J. Luxemburg, editor), Holt, Rinehart and Winston; pp. 171-185.

BERNSTEIN, A. R.

[1973] *Nonstandard analysis*, in: Studies in model theory (M. D. Morley, editor), The Mathematical Association of America, pp. 35-58.

CHANG, C. C. and KEISLER, H. J.

[1973] Model theory, North-Holland, Amsterdam.

DAVIS, M.

[1977] Applied nonstandard analysis, Wiley, New York.

FULKS, W.

[1961] Advanced calculus, Wiley, New York.

GOLDBERG, R. R.

[1964] Methods of real analysis, Blaisdell, New York.

GOLDBLATT, R.

[1979] Topoi, North-Holland, Amsterdam.

HALMOS, P. R.

[1958] Finite-dimensional vector spaces, Van Nostrand, Princeton.

[1974] Measure theory, Springer, New York.

HELMBERG, G.

[1969] Introduction to spectral theory in Hilbert space, North-Holland, Amsterdam.

HENSON, C. W.

[1972] *On the nonstandard representation of measures*, Transactions of the American Mathematical Society, *172*, pp. 437-446.

KEISLER, H. J.

[1976] Elementary calculus, Prindle, Weber and Schmidt, Boston.

KELLEY, J. L.

[1955] General topology, Van Nostrand, Princeton.

KREISEL, G.

[1954] *Applications of mathematical logic to various branches of mathematics*, Colloque de la logique mathématique, Paris, pp. 37-49.

[1965] *Mathematical logic*, in: Lectures on modern mathematics (T. L. Saaty, editor), Wiley, New York, pp. 95-195.

LAUGWITZ, D. and SCHMIEDEN, C.

[1958] *Eine Erweiterung der Infinitesimalrechnung*, Mathematische Zeitschrift, *69*, pp. 1-39.

LIGHTSTONE, A. H. and ROBINSON, A.

[1975] Nonarchimedian fields and asymptotic expansions, North-Holland, Amsterdam.

LOEB, P.

[1972] *A non-standard representation of measurable spaces,* L_∞ *and* L_∞^*, in: Contributions to non-standard analysis (W. A. J. Luxemburg and A. Robinson, editors), North-Holland, Amsterdam, pp. 65-80.

[1974] *A nonstandard representation of Borel measures and σ-finite measures*, in: Victoria symposium on nonstandard analysis (A. Hurd and P. Loeb, editors), Lecture Notes in Mathematics, *369*, Springer.

[1975] *Conversion from nonstandard to standard measure spaces and applications in probability theory*, Transactions of the American Mathematical Society, *211*, pp. 113-122.

[1979] *An introduction to nonstandard analysis and hyperfinite probability theory*, in: Probabilistic analysis and related topics 2 (A. T. Bharucha-Reid, editor), Academic Press, pp. 105-142.

MUNROE, M. E.

[1965] Introductory real analysis, Addison-Wesley, Reading, Mass.

NELSON, E.

[1977] *Internal set theory: A new approach to nonstandard analysis*, Bulletin of the American Mathematical Society, *83*, pp. 1165-1198.

RICHTER, M. M.

[1981] Einführung in die Nichtstandard Analysis, Vieweg, Darmstadt.

ROBINSON, A.

[1966] Non-standard analysis, North-Holland, Amsterdam.

[1973] *Metamathematical problems*, The Journal of Symbolic Logic, *38*, pp. 500-516.

ROYDEN, H. L.

[1968] Real analysis, Macmillan, New York.

RUDIN, W.

[1966] Real and complex analysis, McGraw-Hill, New York.

[1976] Principles of mathematical analysis, McGraw-Hill, New York.

SCOTT, D. S.

[1961] *On constructing models for arithmetic*, in: Infinitistic Methods, Symposium on foundations of mathematics, Warsaw, pp. 235-255.

SIMMONS, G. F.

[1963] Topology and modern analysis, McGraw-Hill, New York.

SKOLEM, Th.

[1970] *Über die Nichtcharakterisierbarkeit der Zahlenreihe mittels endlich oder abzählbar unendlich vieler Aussagen mit ausschliesslich Zahlenvariablen*, in: Selected works in logic by Th. Skolem (J. E. Fenstad, editor), Oslo, pp. 353-366.

SPIVAK, M.

[1967] Calculus, Benjamin, New York.

SZABO, M. E.

[1969] (editor), The collected papers of Gerhard Gentzen, North-Holland, Amsterdam.

[1981] *Standard detours*, Oberwolfach Tagungsbericht.

WEBB, S. M. and CHAPIN, E. W.

[1973] *A non-standard proof in the theory of integration*, Notre Dame Journal of Formal Logic, *14*, pp. 125-128.