BIRKHÄUSER

Martin Väth

Nonstandard Analysis

Birkhäuser Verlag
Basel · Boston · Berlin

Author:

Martin Väth
Freie Universität Berlin
Fachbereich Mathematik und Informatik (WE1)
Arnimallee 2-6
14195 Berlin
Germany
e-mail: vaeth@mathematik.uni-wuerzburg.de

2000 Mathematical Subject Classification 03H05, 26E35, 46S20 (primary); 03C20, 03C62, 03C98, 26A03, 46A99, 46B08, 46M07, 54A99, 54E15 (secondary)

Library of Congress Control Number: 2006934216

Bibliographic information published by Die Deutsche Bibliothek
Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the Internet at <http://dnb.ddb.de>.

ISBN 978-3-7643-7773-1 Birkhäuser Verlag, Basel – Boston – Berlin

Part of Springer Science+Business Media
Printed on acid-free paper produced from chlorine-free pulp. TCF ∞
Printed in Germany
ISBN-10: 3-7643-7773-9 e-ISBN-10: 3-7643-7774-7
ISBN-13: 978-3-7643-7773-1 e-ISBN-13: 978-3-7643-7774-8

9 8 7 6 5 4 3 2 1 www.birkhauser.ch

Contents

Preface

Historically, the idea of nonstandard analysis was to rigorously justify calculations with infinitesimal numbers. For example, formally, the chain rule of Leibniz' calculus for the function $F = f(g(x))$ can be written as

$$\frac{dF}{dx} = \frac{dF}{dg}\frac{dg}{dx},$$

and for a formal proof, one may just divide numerator and denominator by the "infinitesimal small number" dg.

Nowadays, nonstandard analysis has gone far beyond the realm of infinitesimals. In fact, it provides a machinery which enables one to describe "explicitly" mathematical concepts which by standard methods can only be described "implicitly" and in a cumbersome way. In the above example the "standard" notion of a limit is in a certain sense replaced by the "nonstandard" notion of an infinitesimal. If one applies a similar approach to other objects than the real numbers (like topological spaces or Banach spaces etc.), one has a tool which provides "explicit" definitions for objects which can in principle not be described explicitly by standard methods. Examples of such objects are sets which are not Lebesgue measurable, or functionals with certain properties like so-called Hahn–Banach limits. Since it is possible in nonstandard analysis to simply "calculate" with such objects, one can obtain results about them which are extremely hard to obtain by standard methods.

This book is an introduction to nonstandard analysis. In contrast to some other textbooks on this topic, it is not meant as an introduction to basic calculus by nonstandard analysis. Instead, the above mentioned applications in analysis (which are not easily accessible by standard methods) are our main motivation. The infinitesimals are only described as an elementary example for the provided machinery.

Consequently, the reader is supposed to be already familiar with (standard) basic calculus. For deeper understanding, also experience with (basic) topology and

Chapter 1

Preliminaries

§1 Introduction

1.1 General Remarks

Historically, the idea of nonstandard analysis was to find a rigorous justification for calculations with infinitesimal numbers. However, in the author's opinion, this is not the most important property of nonstandard analysis. Instead, it appears to the author that it is more essential that so-called concurrent relations are satisfied. We will make this more precise later, but we already mention that this means, roughly speaking, the following.

If there is a statement which holds for any finite subset of a given set, then it holds for the whole set in nonstandard analysis.

Consider, for example, for any set M of positive real numbers the statement "there is some $c > 0$ with $c < \varepsilon$ for all $\varepsilon \in M$". Clearly, this statement is true for any finite set M of positive real numbers (in our later terminology, we denote such a fact by "concurrency"). This implies that the statement is also true for the set of all positive numbers in nonstandard analysis and so there indeed exists an infinitesimal $c > 0$ which is less than any positive real number. In other words, nonstandard analysis allows us to conclude that "true for each finite number" implies "true for all".

The formulation of the above considerations in precise mathematical terms is rather involved. For this reason (and to have a further motivation up to this point), we will first concentrate on the "classical" topic of nonstandard analysis: This is Leibniz' idea which may be described as follows. Leibniz' program is to join "infinitesimals" to the system $\mathbb{R}$ of real numbers such that the enlarged system obeys the same "rules" as $\mathbb{R}$. As we shall see, this program cannot be carried out directly, because the system of real numbers is uniquely determined by these rules (up to an isomorphism).

The solution proposed by A. Robinson and W. A. J. Luxemburg out of this dilemma is the following: Consider together with $\mathbb{R}$ a nonstandard real line ${}^*\mathbb{R}$ which contains $\mathbb{R}$ and also infinitesimal numbers as elements and which satisfies the following: Any so-called transitively bounded sentence about $\mathbb{R}$ can be transferred into an analogous sentence about ${}^*\mathbb{R}$, and the latter sentence is true if and only if the sentence about $\mathbb{R}$ was true. The crucial point in this concept is that more sentences can be formulated about ${}^*\mathbb{R}$ than those transferred from a sentence about $\mathbb{R}$ (and many of these additional sentences are true). Later, these additional sentences will be called sentences about nonstandard objects. A fundamental point is that true sentences about nonstandard objects can be combined to give a true sentence ${}^*\alpha$ about ${}^*\mathbb{R}$ which can be obtained by transferring some sentence α about $\mathbb{R}$. This allows us to conclude that α is true.

To make this approach precise, one has of course to define what is meant by a "sentence α about $\mathbb{R}$". Then one has to define what is meant by the transferred sentence ${}^*\alpha$. This is the first problem we shall attack.

After this is done, there arises the fundamental question: Does there actually exist an object ${}^*\mathbb{R}$ with the required properties? Or does in contrast the assumption that such an object ${}^*\mathbb{R}$ exists even lead to a contradiction?

The answer to the first question is "yes" if one assumes the axiom of choice (which we therefore do throughout). For this reason the answer to the second question is "no" (even if one rejects the axiom of choice). However, the axiom of choice really is essential.

Applying the above ideas, one can "explicitly construct" objects (in the nonstandard world) which in principle cannot be constructed in the standard world. Such objects are e.g. sets which are not Lebesgue-measurable or so-called Hahn-Banach limits: It is possible to prove the *existence* of such objects in the standard world by means of the axiom of choice, but it is not possible to give explicit formulas for them without the axiom of choice (even if one allows a weaker form of this axiom which allows countable recursive or nonrecursive choices). In fact, assuming the consistency of a so-called inaccessible cardinal, this was first proved in the famous paper [Sol70]. Since in the nonstandard world we can really "calculate" with such objects, it is easy to obtain results which cannot be obtained by standard methods, or only with very abstract applications of the axiom of choice.

Thus, in a sense, nonstandard analysis might just be considered as a machinery to simplify such abstract applications of the axiom of choice by providing objects which implicitly contain this application. Of course, nonstandard analysis means actually much more, but in the author's opinion this is the most important advantage of nonstandard analysis over standard analysis: To have convenient (almost “explicit”) representations of certain obects like Hahn–Banach limits for which by standard methods more or less only their mere existence can be proved with the axiom of choice.

Of course, the above property means that the axiom of choice must actually be involved in the definition of the nonstandard world (or ${}^*\mathbb{R}$); we will see that (a rather strong form of) this axiom comes into play by the choice of an appropriate so-called filter. Due to this crucial role of the axiom of choice in nonstandard analysis, we will assume it throughout. Somebody who rejects the general axiom of choice always has to replace phrases like “... then ... is true” by a phrase like “... then it does not lead to a contradiction to assume that ... is true”.

The study of nonstandard analysis naturally divides into two parts: One part is to define ${}^*\mathbb{R}$, and the other part is to “work” with ${}^*\mathbb{R}$ by using the above described transferring of sentences. The first part belongs to the realm of so-called *model theory* while the second part can be considered as the actual nonstandard analysis (or also just as an application of the first part). It turns out that the second part can be done to a large extent without appealing to the first part, i.e. without explicitly knowing how ${}^*\mathbb{R}$ is defined. There even is an approach to nonstandard analysis (Nelson's internal set theory [Nel77]; see also e.g. [vdB87, LG81, Ric82, Rob88]) which completely hides the definition of ${}^*\mathbb{R}$, and only uses some axioms to describe a new set theory for ${}^*\mathbb{R}$. However, we shall use Robinson's approach ([Rob70], see also [AFHKL86, Cut88, Dav77, Gol98, HL85, LR94, Lux73, Lux69b, SL76, SB86]) which is also concerned with the definition of ${}^*\mathbb{R}$.

This has not only the advantage that we work with “more concrete” objects. But this has also an important practical advantage: In the definition of ${}^*\mathbb{R}$ one has many choices, more than can be described by *any* axiomatic system (this is related with the axiom of choice which we use for the construction: Roughly speaking, we can fix any finite number of choices in a way that we like). Thus, by choosing an “appropriate” definition of ${}^*\mathbb{R}$, we can get some additional properties. This is why the Robinson/Luxemburg approach is actually more powerful than the Nelson approach to nonstandard analysis. There are some applications where this difference really plays a role. A comparison of the two approaches can be found in [DS88] (see also [CK90, LR94]). There is another more algebraic approach to nonstandard analysis via the so-called “Ω-calculus” (see [Lau86]) which, however, is in essence contained in the Robinson/Luxemburg approach. For another approach due to Hrbacek (which is similar to Nelson's approach) and several other approaches and comparisons, we refer the reader to the monograph [RK04].

The crucial point for further applications of nonstandard analysis is that the definition of a nonstandard object ${}^*X$ is not only possible for the case $X = \mathbb{R}$ but also for any other objects X, for example a topological space.

1.2 Archimedean Fields and Infinitesimals

As mentioned in Section 1.1, the idea that the set ${}^*\mathbb{R}$ (containing infinitesimals) should have the same properties as $\mathbb{R}$ soon leads to severe difficulties. We shall discuss these difficulties now in more detail.

Recall that a relation $\leq$ on a set X is called an *order*, if

1. $a \leq a$.
2. $a \leq b$ and $b \leq a$ implies $a = b$.
3. $a \leq b$ and $b \leq c$ implies $a \leq c$.

The order is called *total*, if for each two elements a, b of the set we have either $a \leq b$ or $b \leq a$. We write $a < b$ to denote that $a \leq b$ and $a \neq b$. The order is called *well-order* if each nonempty set has a smallest element. Each well-order is a total order (consider the set $\{a, b\}$ to see this).

A set X with two operations $+$ and $\cdot$ is called a *(commutative) field*, if X is a commutative group with respect to $+$ (we denote the neutral element by 0_X), and $X \setminus \{0_X\}$ is a commutative group with respect to $\cdot$ (we denote the neutral element by 1_X), $0_X \cdot a = a \cdot 0_X = 0_X$, and if furthermore the *distributive law* $a(b + c) = ab + ac$ holds. X is a *totally ordered field* if it is equipped with a total order such that the relations $a \leq b$ and $c \geq 0_X$ imply $a + c \leq b + c$ and $ac \leq bc$. Note that this implies $a^2 \geq 0_X$ for each $a \in X$: For $a \geq 0_X$, this is clear, and for $a < 0_X$, we have $0_X = a - a \leq 0_X - a = -a$, and so $-a \geq 0_X$ which implies $a^2 = (-a)^2 \geq 0_X$, as claimed.

Each totally ordered field X contains a "canonical copy" of the set $\mathbb{N}$, namely $\{1_X, 1_X + 1_X, 1_X + 1_X + 1_X, \ldots\}$ (we write $\mathbb{N}_X := \{1_X, 2_X, \ldots\}$). Note that $\mathbb{N}_X$ is indeed infinite, since $1_X < 2_X < 3_X < \cdots$ because $1_X^2 = 1_X > 0_X$, and so e.g. $2_X < 2_X + 1_X = 3_X$, and so on. Similarly, X contains a "canonical copy" of the sets $\mathbb{Z}$ and $\mathbb{Q}$ of integer and rational numbers. By "canonical copy", we mean that there is an isomorphism, i.e. a bijection $f : \mathbb{Q} \to \mathbb{Q}_X$ which preserves the order and the arithmetic operations (e.g. $f(x + y) = f(x) + f(y)$).

Theorem 1.1. *In a totally ordered field X, the following statements are equivalent:*

1. *X has the* Archimedean property*. For each $x \in X$ there is some $n \in \mathbb{N}_X$ such that $n > x$.*
2. *X has the* Eudoxos property*: For each $\varepsilon \in X$, $\varepsilon > 0_X$, there is some $n \in \mathbb{N}_X$ such that $n^{-1} < \varepsilon$.*
3. *$\mathbb{Q}_X$ is dense in X, i.e. for each $x < y$ there is some $q \in \mathbb{Q}_X$ with $x < q < y$.*

Proof. Assume that X has the Archimedean property, and $x < y$. Then we find some $n \in \mathbb{N}_X$ such that $n > (y - x)^{-1}$ and some $m \in \mathbb{N}_X$ with $m > nx$ and $m > -nx$. Let z be the smallest number of $\{-m, \ldots, m\}$ which satisfies $z > nx$, hence $z - 1_X \leq nx$. Then $q := z/n \in \mathbb{Q}_X$ satisfies $q > x$. Moreover, $q < y$, since otherwise $z \geq ny = n(y-x) + nx > 1_X + nx$. Hence, $\mathbb{Q}_X$ is dense in X.

If $\mathbb{Q}_X$ is dense in X, then $\varepsilon > 0_X$ implies that we find some $q \in \mathbb{Q}_X$ such that $\varepsilon > q > 0_X$. Then $q = z/n$ for $z, n \in \mathbb{N}_X$, and so $\varepsilon > n^{-1} > 0_X$. Hence, X has the Eudoxos property.

Let X have the Eudoxos property and $x \in X$. If $x \leq 0_X$, we have $1_X > x$. Otherwise $\varepsilon := x^{-1} > 0$, and we find some $n \in \mathbb{N}_X$ with $n^{-1} < \varepsilon$. In both cases, we find some $n \in \mathbb{N}_X$ with $n > x$, and so X has the Archimedean property. □

If a totally ordered field X has the Archimedean property, we simply call X an *Archimedean field.*

A totally ordered field X is called *(Dedekind) complete* if each nonempty subset $A \subseteq X$ which is bounded from above has a smallest upper bound (i.e. if the set B of upper bounds for A is nonempty, this set has a minimum).

Given some totally ordered field X, we define the *Dedekind completion* $\overline{X}$ of X as follows:

A nonempty set $A \subseteq \mathbb{Q}_X$ is called a *Dedekind cut*, if

1. A is bounded from above,
2. A does not possess a largest element, and
3. the relations $a \in A$, $b \in \mathbb{Q}$ and $b \leq a$ imply $b \in A$.

(We emphasize that this definition is only useful if X is a totally ordered field as in our situation; in more general situations of so-called Riesz spaces, one has to use more sophisticated definitions, see e.g. [LZ71, §32]). The set $\overline{X}$ of all Dedekind cuts is ordered by inclusion, i.e. $A \leq B$ if and only if $A \subseteq B$. Note that this is a total order, and in particular

$$A < B \iff B \setminus A \neq \varnothing.$$

We define an addition $+$ on $\overline{X}$ by

$$A + B = \{a + b : a \in A \text{ and } b \in B\}.$$

Observe that $A + B \in \overline{X}$: If A is bounded by a and B is bounded by b, then $A + B$ is bounded by $a + b$. The set $A + B$ does not contain a largest element $a + b$ with $a \in A$, $b \in B$, since we find some $a_0 \in A$ with $a_0 > a$, and so $a + b < a_0 + b \in A + B$. Finally, we have for any $a \in A$, $b \in B$, $c \in \mathbb{Q}_X$ with $c \leq a + b$ that $a_0 = c - b \in A$, and so $c = a_0 + b \in A + B$.

We define $\mathbb{Q}_{\overline{X}}$ as the set of all cuts of the form

$$q_{\overline{X}} := \{x \in \mathbb{Q}_X : x < q\} \qquad (q \in \mathbb{Q}_X).$$

Moreover, we put

$$-A = \begin{cases} \{b \in \mathbb{Q}_X : b < -a \text{ for all } a \in A\} = \{a \in \mathbb{Q}_X : -a \notin A\} & \text{if } A \notin \mathbb{Q}_{\overline{X}}, \\ (-q)_{\overline{X}} & \text{if } A = q_{\overline{X}}. \end{cases}$$

For the equality in the above definition, observe that if $b := -a \notin A$, then $b > x$ for all $x \in A$, and so $a = -b < -x$ for all $x \in A$ which implies that $a \in -A$; conversely, if $b < -a$ for all $a \in A$, then $-b \notin A$. Note that $-A$ is indeed a Dedekind cut: The only nontrivial property is that if $A \notin \mathbb{Q}_{\overline{X}}$, then $-A \notin \mathbb{Q}_{\overline{X}}$ has no largest element. But if b were such a largest element, then $A = (-b)_{\overline{X}}$, a contradiction: Indeed, if $a \in A$, then $b < -a$ by definition of $-A$, i.e. $a \in (-b)_{\overline{X}}$. Conversely, if $x \in (-b)_{\overline{X}}$, then $a := -x > b$, and we must have $x \in A$: Otherwise $-a \notin A$ which means $a \in -A$ and contradicts the assumption that $b < a$ is the largest element of $-A$.

For $A, B > 0_{\overline{X}}$, we define a multiplication

$$A \cdot B := \{a \cdot b : a \in A,\ b \in B,\ a, b > 0_X\} \cup \{x \in \mathbb{Q}_X : x \leq 0_X\}$$

which defines a Dedekind cut: If $a \in A$, $b \in B$ with $a, b > 0_X$ and $c \in \mathbb{Q}_X$ satisfy $c < ab$, then either $c \in A \cdot B$ because $c \leq 0_X$, or $a_0 := c/b < a$, and so $a_0 \in A$ and $a_0 > 0$ imply $c = a_0 b \in A \cdot B$. No number $a \cdot b$ with $a \in A$, $b \in B$, $a, b > 0_X$ can be a largest element of $A \cdot B$, since there is some $a_0 \in A$ with $a_0 > a$, and so $ab < a_0 b \in A \cdot B$.

For $A < 0_{\overline{X}} < B$, we define $A \cdot B := -((-A) \cdot B)$, for $B < 0_{\overline{X}} < A$, we define $A \cdot B := -(A \cdot (-B))$, for $A, B < 0_{\overline{X}}$, we put $A \cdot B := (-A)(-B)$, and if $A = 0_{\overline{X}}$ or $B = 0_{\overline{X}}$, we put $A \cdot B := 0_{\overline{X}}$.

The name “Dedekind completion” is indeed justified for $\overline{X}$:

Theorem 1.2 (Dedekind)**.** *The Dedekind completion $\overline{X}$ of a totally ordered field X is a complete Archimedean field with $\mathbb{Q}_{\overline{X}}$ as the canonical copy of $\mathbb{Q}_X$.*

If X is an Archimedean field, then $\overline{X}$ contains a canonical copy of X. Moreover, if X is a complete Archimedean field, then this canonical copy is X, i.e. $\overline{X}$ is isomorphic to X (i.e. there is a bijection which preserves the order and the arithmetic operations).

Proof. The fact that $\overline{X}$ is a totally ordered field is a straightforward verification of the axioms. We leave the details to the reader.

To see that $\overline{X}$ is Archimedean, we prove that $\mathbb{Q}_{\overline{X}}$ is dense in $\overline{X}$: If $A < B$, we find by definition some $b \in B \setminus A$. Since b is not a largest element of B, we find some $q \in B$ with $q > b$. Then $A \leq b_{\overline{X}} < q_{\overline{X}} < B$, and so we have found some $q_{\overline{X}} \in \mathbb{Q}_{\overline{X}}$ with $A < q_{\overline{X}} < B$, as desired.

Let us now prove that $\overline{X}$ is complete. Thus, let $\mathscr{A} \subseteq \overline{X}$ be a nonempty subset which is bounded from above by some $B \in \overline{X}$. We claim that $S = \bigcup \mathscr{A}$ is a Dedekind cut:

Indeed, S is bounded from above: Since $B + 1_{\overline{X}} > B$, we find some $b \in B + 1_{\overline{X}} \setminus B$. Then $b_{\overline{X}} > A$ for each $A \in \mathscr{A}$ which implies that b is an upper bound for S. Moreover, S has no largest element, since for each $a \in S$ we find some $A \in \mathscr{A}$ with $a \in A$ and some $a_0 > a$ with $a_0 \in A \subseteq S$. Finally, the relations $s \in S$ and $s > a \in \mathbb{Q}_X$ trivially imply $a \in \mathscr{A}$.

Hence, $S \in \overline{X}$. Since $A \subseteq S$ for any $A \in \mathscr{A}$, S is an upper bound for $\mathscr{A}$. But S is the smallest upper bound, since each upper bound for $\mathscr{A}$ must at least contain any set $A \subseteq \mathscr{A}$ as a subset.

Let now X be Archimedean. Then each of the sets

$$x_{\overline{X}} = \{q \in \mathbb{Q}_X : q < x\} \qquad (x \in X) \tag{1.1}$$

belongs to $\overline{X}$: Since X is Archimedean, each $x_{\overline{X}}$ is nonempty and bounded from above. Moreover, since $\mathbb{Q}_X$ is dense (Theorem 1.1), $x_{\overline{X}}$ has no maximal element Now the map $x \mapsto x_{\overline{X}}$ is the desired embedding of X into $\overline{X}$: Since $\mathbb{Q}_X$ is dense, we have $x_{\overline{X}} < y_{\overline{X}}$ if and only if $x < y$ (and so in particular the mapping is one-to-one). Moreover, $(x \pm y)_{\overline{X}} = x_{\overline{X}} \pm y_{\overline{Y}}$, and similarly for the multiplication.

It remains to prove that this embedding is onto, if X is complete, i.e. that any Dedekind cut A has the form (1.1) with some $x \in X$. But since X is Dedekind complete, any Dedekind cut A has a smallest upper bound $x = \sup A$ in X which implies $A = x_{\overline{X}}$. □

Theorem 1.3. *Any Archimedean field X is isomorphic to a subfield of $\mathbb{R}$. If X is Dedekind complete, it is even isomorphic to $\mathbb{R}$.*

Proof. Let $\overline{\mathbb{Q}}$ denote the Dedekind completion of $\mathbb{Q}$. We show that any Archimedean field X is isomorphic to a subfield of $\overline{\mathbb{Q}}$ and isomorphic to $\overline{\mathbb{Q}}$ if X is complete. Since then in particular $\mathbb{R}$ is isomorphic to $\overline{\mathbb{Q}}$, no matter which definition for $\mathbb{R}$ the reader wants to choose (for any possible definition, $\mathbb{R}$ is a Dedekind complete Archimedean field), the statement follows.

The isomorphism is defined as follows: Let $f : \mathbb{Q} \to \mathbb{Q}_X$ be the canonical isomorphism. Then f induces an isomorphism $F : \overline{\mathbb{Q}} \to \overline{X}$, defined by $F(A) = \{q \in \mathbb{Q}_X : f^{-1}(q) \in A\}$ (it is evident that this is a bijection, but it is also easily verified that the algebraic operations and the order structure are preserved). Hence, $\overline{\mathbb{Q}}$ and $\overline{X}$ are isomorphic. Since Theorem 1.2 implies that $\overline{X}$ contains X as a subfield (resp. is isomorphic to X if X is complete) we may conclude that X is isomorphic to a subfield of $\overline{\mathbb{Q}}$ (resp. isomorphic to $\overline{\mathbb{Q}}$ if X is complete). This is what we had to prove. □

We thus have the problem that it actually is not possible to join "infinitesimals" to the set $\mathbb{R}$ such that we get a set ${}^*\mathbb{R}$ in which the same rules hold as in $\mathbb{R}$: If ${}^*\mathbb{R}$ were in particular a Dedekind complete Archimedean field, we would have ${}^*\mathbb{R} = \mathbb{R}$ (up to an isomorphism). Thus, the main goal of introducing a set ${}^*\mathbb{R}$ containing infinitesimals cannot be achieved! Then what is the rest of this book about?

The trick is that *although* ${}^*\mathbb{R}$ *is neither Archimedean nor Dedekind complete*, it shares all properties of $\mathbb{R}$! This appears to be a contradiction, but of course depends on the definition of the term "property": Of course, if one considers e.g. Dedekind completeness as a property of $\mathbb{R}$, this is not true. However, Dedekind completeness is actually not a property of $\mathbb{R}$ or of the real numbers (i.e. of the *elements* of $\mathbb{R}$) but a property of the *subsets* of $\mathbb{R}$. This may be considered as a "higher type" property: Let us for a moment call statements concerning relations of real numbers (such as e.g. the distributive law) *properties of type* 0, while we call a statement concerning subsets of real numbers (such as Dedekind completeness) a *property of type* 1; a *property of type* 2 could be called a statement about sets of subsets of real numbers, and so on. It turns out that ${}^*\mathbb{R}$ completely shares the properties of $\mathbb{R}$ of type 0. But it does not satisfy all properties of higher type. However, the theory would not be very useful, if ${}^*\mathbb{R}$ completely violates properties of higher type. In a weak sense ${}^*\mathbb{R}$ also satisfies the properties of higher type of $\mathbb{R}$, but with some restrictions. As remarked above, ${}^*\mathbb{R}$ is not Dedekind complete, i.e. not any subset of ${}^*\mathbb{R}$ which is bounded from above must have a smallest upper bound. However, this holds for the so-called *internal* subsets of ${}^*\mathbb{R}$ which, roughly speaking, are sets which can be described by properties of type 0.

Actually, one could be more precise than to define a type as a *number*; instead one could choose a certain *set* which "represents the type" in a more detailed way—such a theory of types was Robinson's original approach. But this is a rather technical procedure. Instead, we follow the approach of Luxemburg in which one does not have to care much about types: It turns out that any *bounded sentence* which holds for $\mathbb{R}$ also holds for ${}^*\mathbb{R}$. Properties of type 0 in the sense sketched above can always be formulated as bounded sentences.

Let us also mention other ways of introducing infinitesimals by dropping some of the axioms of Dedekind complete Archimedean fields. These approaches are, however, not too useful for our purpose, since the constructed sets X have a rather different structure than $\mathbb{R}$.

Example 1.4. Let X be the set of all rational functions (i.e. $x(t) = p(t)/q(t)$ where p and q are polynomials and q is not the zero-polynomial). Addition and multiplication are defined in the evident way (pointwise), and an order is introduced as follows: We say $x \leq y$ if there is some t_0 such that $x(t) \leq y(t)$ for $t > t_0$, and $x < y$ if $x \leq y$ and $x \neq y$.

Then X is a totally ordered field which contains a copy of $\mathbb{R}$ as the subset of constant functions. Moreover, X contains nonzero infinitesimals, i.e. elements x which satisfy $0 < x < \varepsilon$ for any *real number* $\varepsilon > 0$. Indeed, the element $x(t) = 1/t$ has this property.

Exercise 1. Prove the statements of Example 1.4. Is the field X defined there Dedekind complete? Is it Archimedean?

Example 1.5. Let

$$X := \{x_{a,b} := \begin{pmatrix} a & b \\ 0 & a \end{pmatrix} : a, b \in \mathbb{R}\}.$$

Addition and multiplication are defined in the obvious way (matrix-multiplication). We define an order by saying that $x_{a_1,b_1} < x_{a_2,b_2}$ if either $a_1 < a_2$ or if simultaneously $a_1 = a_2$ and $b_1 < b_2$ (i.e. we order the pairs (a, b) lexicographically). A straightforward calculation shows that X is a totally ordered ring, i.e. it satisfies all axioms of a totally ordered field with the possible exception of the existence of inverses with respect to multiplication. The set X contains a copy of $\mathbb{R}$ (namely $\{x_{a,0} : a \in \mathbb{R}\}$). Moreover, X contains nonzero infinitesimals: The element $x_{0,1}$ satisfies $x_{0,0} < x_{0,1} < x_{\varepsilon,0}$ for each $\varepsilon > 0$.

Exercise 2. Prove the statements of Example 1.5. Is the totally ordered ring X defined there even a field? Is it Dedekind complete? Is it Archimedean?

The following example has some relation to the way in which ${}^*\mathbb{R}$ will later actually be defined. To motivate this example, recall that $\mathbb{R}$ can alternatively be defined from $\mathbb{Q}$ by Cauchy sequences: Call two rational Cauchy sequences x_n and y_n equivalent if $x_n - y_n \to 0$. The set of all such equivalence classes, equipped with the natural operations, is isomorphic to $\mathbb{R}$. One might try to introduce an infinitely large number into $\mathbb{R}$ by a similar method: For example, one may define $+\infty$ as the equivalence class of all sequences $x_n \to \infty$.

Example 1.6. Let X_0 be the set of all sequences with values in $\mathbb{R}$, i.e. of all mappings $x : \mathbb{N} \to \mathbb{R}$. We call two elements of X_0 equivalent, if they differ at most on finitely many points of $\mathbb{N}$. Let X be the set of all equivalence classes $[x]$ where $x \in X_0$. Addition and multiplication in X_0 are defined in an evident way (pointwise). The addition and multiplication in X is defined by $[x] + [y] := [x + y]$ and $[x] \cdot [y] := [x \cdot y]$, and we write $[x] \leq [y]$ if and only if x does not exceed y at all except possibly finitely many points. It is straightforward to check, that these notions are well-defined, i.e. that they actually do not depend on the choice of the representatives x and y. Moreover, with these notions and the convention $[x] < [y]$ if and only if $[x] \leq [y]$ and $[x] \neq [y]$, the set X becomes an ordered ring, i.e. it satisfies the axioms of a totally ordered ring with the exception that the order need not be total. The equivalence classes of constant sequences constitute a canonical copy of $\mathbb{R}$ in X. The equivalence class of the sequence $x : n \mapsto n^{-1}$ is a nonzero infinitesimal in X, since for any $\varepsilon \in \mathbb{R}$, $\varepsilon > 0$, we have n^{-1} for all except finitely many n, and so $0 < [x] < \varepsilon$.

Exercise 3. Prove the statements of Example 1.6. Is the ordered ring X defined there even totally ordered? Or a field? Is it Dedekind complete? Is it Archimedean?

The disappointing properties of Example 1.6 are due to the fact that we chose the "wrong" definition for the equivalence of sequences. Later, we will call two sequences equivalent, if they are equal "almost everywhere". However, the term "almost everywhere" will be defined in a very tricky manner by means of an ultrafilter. The fact that (nontrivial) ultrafilters cannot be *constructed* is a deeper reason why the model ${}^*\mathbb{R}$ defined later is not very explicit (but the axiom of choice will imply its existence).

Exercise 4. (Very difficult). Let X_0 consist of all measurable functions $x : [0,1] \to \mathbb{R}$. Call two functions equivalent, if they differ only on a (Lebesgue) null set, and let X be the set of the corresponding equivalence classes. Addition, multiplication and the order on X is introduced analogously to Example 1.6. Prove that X is an ordered ring. The reals are embedded into X as equivalence classes of constant functions. Prove that X contains also "infinitesimals" in the sense that there is some $x > 0$ in X which is not larger than any real $\varepsilon > 0$. Is X totally ordered? Or a field? Is it Dedekind complete? Is it Archimedean?

The difficulty of nonstandard analysis lies not only in the mere introduction of infinitesimals. Think, for example, of a natural infinitesimal description of the statement "The function $f : \mathbb{R} \to \mathbb{R}$ is continuous at 0". The "intuitive" description will read: Whenever x is "infinitely close" to 0, then $f(x)$ is "infinitely close" to $f(0)$. Even if we know some extension ${}^*\mathbb{R}$ of $\mathbb{R}$ with infinitesimals, there arises the problem that we have given a function $f : \mathbb{R} \to \mathbb{R}$ (and not a function $f : {}^*\mathbb{R} \to {}^*\mathbb{R}$): So how should $f(x)$ be defined if x is an infinitesimal?

For this reason, a proper theory of infinitesimals should also associate to each function $f : \mathbb{R} \to \mathbb{R}$ an extension ${}^*f : {}^*\mathbb{R} \to {}^*\mathbb{R}$. How could such an extension be defined? If $f(x) = a_0 + a_1 x + \cdots + a_n x^n$ is a polynomial, it is clear how the extension should be defined. Similarly, if f is a rational function or a power series. The reader who has some knowledge in complex analysis will find an analogy with analytic functions: All "basic" analytic real functions have a unique canonical extension into (a large part of) the complex plane. But how should functions like the Dirichlet function ($f(x) = 1$ if $x \in \mathbb{Q}$ and $f(x) = 0$ otherwise) be extended?

Of course, one would like that the extension ${}^*f$ of f has the same "formal" properties as f. For example, for $f(x) = \sin(x)$, one would like to have that $|{}^*f(x)| \leq 1$ for all $x \in {}^*\mathbb{R}$ and that ${}^*f(x+\pi) = -{}^*f(x)$ for $x \in {}^*\mathbb{R}$.

In this connection, the approach of Example 1.6 is very promising: Any $x \in X$ is an equivalence class of sequences. Of course, any function $f : \mathbb{R} \to \mathbb{R}$, no matter how complicated, maps any sequence $x : \mathbb{N} \to \mathbb{R}$ into another sequence $y : \mathbb{N} \to \mathbb{R}$ in a canonical way by the formula $y(n) = f(x(n))$. If the equivalence class of y depends only on the equivalence class of x, one may consider this as a mapping ${}^*f : X \to X$. This is the idea that we will actually use.

However, such a direct approach will lead to many technical difficulties: For example what is to be done if f is only defined on a subset $A \subseteq \mathbb{R}$? Moreover, how could we discuss functions which are constructed by means of infinitesimals like $f(x) = \sin(x/c)$ where $c > 0$ is infinitesimally small (later, such functions are called *internal*). Of course, one could try to discuss such special cases in all detail, but we will use another more axiomatic approach: We will define a mapping $* : \widehat{S} \to \widehat{{}^*S}$ on a very large set $\widehat{S}$ which contains $\mathbb{R}$, all subsets of $\mathbb{R}$, all functions on $\mathbb{R}$, etc., and which associates to each such element a a nonstandard element ${}^*a$. Moreover, this mapping will be defined in such a way that the truth of sentences is preserved in the following sense: If α is a sentence, and if we put a $*$ at each element of the sentence, then the corresponding sentence ${}^*\alpha$ is true if and only if α is a true sentence. This will be called the *transfer principle*. Observe that one may conversely also conclude that α is true if ${}^*\alpha$ is true: In this way it is possible to prove "standard" results with "nonstandard" methods.

The plan for the beginning is as follows: In §2, we describe more precisely how $\widehat{S}$ and sentences α are defined. The mapping $*$ is axiomatically introduced in §3, without discussing how such a mapping might be defined. One way to define such a mapping (similarly to Example 1.6) will be discussed in §4. We already point out that without the axiom of choice the existence of such a mapping $*$ cannot be proved, i.e. the existence proofs of $*$ must always be nonconstructive. In particular, even the simplest results of nonstandard analysis contain implicitly a nonconstructive element. This is one of the main arguments against the use of nonstandard analysis. On the other hand, physicists like to think of infinitesimals as "existing" objects, in particular in string theory. In this sense it is perhaps not false to think of nonstandard analysis as an "extreme idealization" of reality which, however, might be even "too idealized"; so one should treat the results with care.

The reader is asked to be patient in these first sections: The existence proof of $*$ and the fundamental properties are rather complicated and technical; the calculation with infinitesimals which has been briefly discussed above appears after this as a simple exercise (and is in fact not much more than a special case of general properties of $*$). However, the main power of nonstandard analysis becomes clear if one considers more complicated structures than $\mathbb{R}$, such as Banach spaces etc. These applications are the topic of the second part of this book.

§2 Superstructures, Sentences, and Interpretations

Before we can define the details of nonstandard analysis, we need some fundamental concepts of model theory which we describe in this section.

2.1 Superstructures

As we have mentioned before, we intend to define a map $*$ which maps each object (number, function, set, etc.) of the standard world into a corresponding object of the nonstandard world. If one tries to map actually *any* set into a nonstandard set, one is in the realm of category theory, and serious fundamental difficulties arise (for example, $*$ cannot be a *map* in the sense of set theory). The easiest way to overcome these difficulties is to work with a "restricted universe" which is still a set. This has the disadvantage that we have to work with *atoms* which however is not a big problem from the viewpoint of applications. We define such a universe now:

Let S be some set in which we are interested. For example, $S = \mathbb{R}$, or S is the point set of a topological space. If we speak of "statements about S", we are actually interested in statements about elements of S, subsets of S, functions and relations on such subsets, sets of such functions, functions on such sets of functions, etc. All these objects can be found in a set which is called the *superstructure* of S. This superstructure is defined in the following way:

If A is some set, we denote by $\mathcal{P}(A)$ the *powerset* of A, i.e. the system of all subsets of A. Let $S_0 := S$, and for $n = 1, 2, \ldots$ define inductively $S_n := S_0 \cup \mathcal{P}(S_{n-1})$. Then $\widehat{S} := \bigcup S_n$ is the *superstructure* of S. The elements of S are called *individuals* or *atoms*, and the elements of $\widehat{S}$ which are not atoms are called *entities*.

The notion "atom" is chosen, because the sentence $a \in s$ should always be false for an atom $s \in S$ and $a \in \widehat{S}$ (at least, we will assume this throughout: If necessary, we have to "rename" the elements of S to achieve this). For example, if we are interested in statements about $S = \mathbb{R}$, this means that e.g. the statement $a \in 1$ is always false (for $a \in \widehat{S}$). Thus, the choice $S = \mathbb{R}$ means that we do not care how the real numbers might be constructed—we just assume them as given. If we are also interested in the definition of the real numbers from the natural numbers, we could start with the set $S = \mathbb{N}$ instead. The only restriction is that we should start with an infinite set S (because otherwise things degenerate and no nonstandard analysis is possible, as we shall soon see).

To see that also functions and relations are contained in the superstructure, we briefly have to recall how functions and relations are defined in set theory:

A *pair* (a,b) is defined by the formula $(a,b) := \{\{a\},\{a,b\}\}$. By induction, an n-*tuple* $(a_1,\dots,a_n)$ is defined as the pair $((a_1,\dots,a_{n-1}),a_n)$ for $n \geq 2$ where we put $(a_1) := a_1$ for $n = 1$. The *Cartesian product* $A_1 \times \dots \times A_n$ of finitely many sets $A_1,\dots,A_n$ is the set of all n-tuples $(a_1,\dots,a_n)$ where $a_i \in A_i$ $(i = 1,\dots,n)$. An *(n-ary) relation* Φ over $A_1,\dots,A_n$ is a subset of $A_1 \times \dots \times A_n$. If Φ is a *binary relation* over A, B, we write $\mathrm{dom}(\Phi)$ for the *domain* of Φ (i.e. $\mathrm{dom}(\Phi)$ is the set of all $a \in A$ such that there is some $b \in B$ with $(a,b) \in \Phi$); similarly, $\mathrm{rng}(\Phi)$ denotes the *range* of Φ (i.e. $\mathrm{rng}(\Phi)$ is the set of all $b \in B$ such that there is some $a \in A$ with $(a,b) \in \Phi$). If for each $a \in A$ we find precisely one $b \in B$ with $(a,b) \in \Phi$, then Φ is called a *function*. In this case, the usual notation $\Phi : A \to B$ is used, and $\Phi(a)$ denotes the unique $b \in B$ with $(a,b) \in \Phi$. Note that by this convention, a function f is not only determined by its *graph* $\{(x, f(x)) : x \in \mathrm{dom}(f)\}$, it even more *is* its graph!

Now it is easy to see that all functions mentioned in the beginning are entities. More general, if A and B are entities, then any mapping $f : A \to B$ (and any relation $R \subseteq A \times B$) is also an entity. Roughly speaking, $\widehat{S}$ is closed under all natural set operations:

Theorem 2.1. *The following holds in any superstructure* $\widehat{S}$*:*

1. $S_0 \in S_1 \in S_2 \in \dots \in \widehat{S}$ *and* $S_0 \subseteq S_1 \subseteq \dots \subseteq \widehat{S}$*. In particular,* S_n *are entities. Also* $\varnothing$ *is an entity.*
2. *Each* S_n *is* transitive*, i.e. each element of* S_n *which is not an atom is a subset of* S_n*. The same holds for* $\widehat{S}$*. In other words: If* A *is an entity and* $x \in A$*, then* x *is either an entity or an atom.*
3. *If* A *is an entity and* $B \subseteq A$*, then* B *is also an entity. In particular, if* $B \subseteq S_n$ *for some* n*, then* B *is an entity.*
4. *If* A *is an entity, then* $\mathcal{P}(A)$ *is an entity.*
5. *Let* A *be a set of entities. If* $A \neq \varnothing$*, then* $\bigcap A := \bigcap_{x \in A} x$ *is an entity. If* $A \in \widehat{S}$*, then* $\bigcup A := \bigcup_{x \in A} x$ *is an entity.*
6. *If* $x_1,\dots,x_k \in \widehat{S}$*, then* $\{x_1,\dots,x_k\}$ *is an entity.*
7. *If* $A_1,\dots,A_k$ *are entities, then* $\bigcup_j A_j = A_1 \cup \dots \cup A_k$ *is an entity.*
8. *If* $A_1,\dots,A_n$ *are entities, then* $A_1 \times \dots \times A_n$ *is an entity.*
9. *All* n*-ary relations on entities are entities. In particular, all functions acting between entities are themselves entities.*

Theorem 2.1 implies that the superstructure $\widehat{S}$ is built of "levels" S_n: If we can show that a set belongs to some level (either as an element or as a subset), this set belongs to the superstructure. Conversely, each element of the superstructure is contained in some level (both as an element and as a subset). For this reason, the elements of $S_n \setminus S_{n-1}$ are said to be of *type* n; the atoms are said to be of type 0. As we shall see in the following proof, the "operations" $\mathcal{P}$ and $\times$ may increase the type, but they do not lead out of the superstructure.

Proof of Theorem 2.1. 1. The inclusion $S_{n-1} \subseteq S_n$ follows by induction: This is true for $n = 1$, and if it is true for some n, then $S_n = S_0 \cup \mathcal{P}(S_{n-1}) \subseteq S_0 \cup \mathcal{P}(S_n) = S_{n+1}$. The inclusions $S_{n-1} \in S_n$ and $S_n \subseteq \widehat{S}$ follow immediately from the definition. Finally, $\varnothing, S_n \in S_{n+1} \subseteq \widehat{S}$.
2. If $A \in S_n = S_0 \cup \mathcal{P}(S_{n-1})$ is not an atom, we must have $A \in \mathcal{P}(S_{n-1})$, i.e. $A \subseteq S_{n-1} \subseteq S_n$ by 1. If $A \in \widehat{S}$, then $A \in S_n$ for some n, i.e. by what we just proved $A \subseteq S_n$ for some n.
3. $A \in S_n$ for some n which implies $A \subseteq S_n$ by 2. Hence, any $B \subseteq A$ is a subset of S_n, i.e. $B \in \mathcal{P}(S_n) \subseteq S_{n+1} \subseteq \widehat{S}$.
4. As in 3., we have $A \subseteq S_n$ for some n. This means $\mathcal{P}(A) \subseteq \mathcal{P}(S_n) \subseteq S_{n+1}$. Hence, $\mathcal{P}(A)$ is an entity by 3.
5. If $A \neq \varnothing$, choose some $x \in A$. Then x is an entity, and so $\bigcap A \subseteq x$ is an entity by 3. If A is an entity, we have $A \in S_n$ for some $n \geq 1$, which implies $A \subseteq S_n$ by 2., i.e. $A \subseteq S_0 \cup \mathcal{P}(S_{n-1})$. Since A contains no atoms, we have $A \subseteq \mathcal{P}(S_{n-1})$, i.e. all elements of A are subsets of S_{n-1}. Consequently, $\bigcup A \subseteq S_{n-1}$. Hence, the statement follows from 3.
6. By 1., we have $x_1, \ldots, x_k \in S_n$ for some n, and so $\{x_1, \ldots, x_k\} \in S_{n+1} \subseteq \widehat{S}$.
7. By 6., the set $A := \{A_1, \ldots, A_k\}$ is an entity, and so $\bigcup_j A_j = \bigcup A$ is an entity by 5.
8. For $n = 2$, it suffices to observe that by definition of pairs $A_1 \times A_2 \subseteq \mathcal{P}(\mathcal{P}(A_1 \cup A_2))$ and to apply 4. and 5. The general case follows by a trivial induction, since by definition $A_1 \times \cdots \times A_n = (A_1 \times \cdots \times A_{n-1}) \times A_n$.
9. If Φ is an n-ary relation on entities $A_1, \ldots, A_n$, then $\Phi \subseteq A_1 \times \cdots \times A_n$. Since $A_1 \times \cdots \times A_n$ is an entity by 8., also Φ is an entity by 3. $\square$

More surprising than the above operations which are allowed in $\widehat{S}$ is that some important operations are not allowed: In general, subsets of $\widehat{S}$ are not entities. For example $\widehat{S}$ itself is no entity; more generally, the subsets of $\widehat{S}$ which are not entities are precisely those sets which contain elements of infinitely many types (i.e. which contain an infinite subset of elements of pairwise different type).

In other words: A set which is the "collection" of elements of $\widehat{S}$ is not an entity, in general. However, *finite* collections of such elements are entities, as we have seen.

For the same reason as above, the union of a set A of entities need not be an entity (a counterexample is $A := \{\{x\} : x \in \widehat{S}\}$). However, by what we have proved, this is the case if either A itself is an entity or if A is finite.

Thus, roughly speaking, the operations "union", and "collecting to some set" are in general not admissible. An exception can be made if we either consider only finitely many entities or if the whole collection forms an entity.

The reader familiar with set theory might find some analogies of these restrictions to usual set theory: Recall e.g. that there is no set containing all sets in the universe (this is analogous to the fact that $\widehat{S}$ is no entity). Moreover, it is in general not allowed to "collect" arbitrary elements to some set. In particular, the "union" of a class A of arbitrary sets need not be a set, in general. However, we have the union axiom: If A *is* a set, then also $\bigcup A$ is a set.

In this sense, we might consider the superstructure $\widehat{S}$ as a model of a set theory (with so-called "urelements" which are the atoms, i.e. the elements of S): The superstructure $\widehat{S}$ serves as a model, if we just interpret the system of entities as the class of all sets in the universe. $\widehat{S}$ satisfies all axioms of set theory with the exception of the infinity axiom. Note, that if we eliminate the urelements, i.e. if we consider $S := \varnothing$, the corresponding superstructure $\widehat{S}$ indeed contains only finite sets. The same is true if S is finite. Thus, the only "interesting" case is the one when S is infinite.

There is one axiom in set theory which in a certain sense allows us to "collect" elements to some set: This is the axiom of choice. Since we assume the axiom of choice throughout, it turns out that the axiom of choice holds also in the superstructure $\widehat{S}$:

Let $A \in \widehat{S}$ be a set whose elements are entities, and $\varnothing \notin A$. By the axiom of choice, there is a function $f : A \to \bigcup A$ such that $f(x) \in x$ for each $x \in A$. If A is an entity, then f is an entity, because A and $\bigcup A$ are entities.

The superstructure $\widehat{S}$ is large enough to represent all structures that typically occur in problems of analysis, topology, ... if S is chosen appropriately. For example, if one is interested in a continuous mapping $f : X \to Y$ of topological spaces, one may choose $S := X \cup Y$. Note that then also the topology of X and Y is contained in the superstructure $\widehat{S}$.

For *real analysis* an appropriate choice is $S := \mathbb{N}$ or $S := \mathbb{R}$ (it suffices to consider $S := \mathbb{N}$, since $\mathbb{R}$ may be constructed from the reals in one of the usual well-known manners: This construction can completely be carried out within the superstructure; in this sense, $\mathbb{N} \in \widehat{S}$ already implies $\mathbb{R} \in \widehat{S}$). Note that also the algebraic structure and the order structure of $\mathbb{R}$ is contained in $\widehat{S}$, since e.g. the addition is represented by the entity $\{(x, y, z) \in \mathbb{R}^2 : x + y = z\}$.

As another example, if one is interested in a mapping $f : X \to Y$ in normed spaces X and Y, the choice $S := X \cup Y \cup \mathbb{N}$ or $S := X \cup Y \cup \mathbb{R}$ is appropriate.

We already point out an important logical observation: If one is interested in a statement like "any mapping $f : \{0\} \to A$ with a topological space A is continuous", one cannot find a superstructure in which this statement makes sense (because the class of *all* topological spaces A is not a set). However, given *any* topological space A, we may consider the superstructure $\widehat{S}$ corresponding to

$S = A \cup \{0\}$ and can verify the statement in $\widehat{S}$. If we can do this for any given instance of A, we may conclude that the general statement is true.

Now observe that *all* statements occurring in analysis, topology, ... have such a form that it suffices to verify them for given instances, and so superstructures are actually large enough to represent the corresponding statements. Later, we will tacitly make use of the above mentioned reasoning.

Let us introduce now a notational convention that we shall use throughout: Let $i : X \to Y$ be a map. Then the value of the image of a point $x \in X$ under this map is usually denoted by $i(x)$. However, in nonstandard analysis it is sometimes more convenient to write ${}^{i}a$ for the value. We shall use both conventions. The value of a set $A \subseteq X$ is defined as $i(A) = \{i(x) : x \in A\}$. This definition may be ambiguous: If e.g. $X = \widehat{S}$ is a superstructure and $A \in \widehat{S}$ is an entity, it is not clear whether $i(A)$ means the image of the *element* A, or of the *set* A (which consists of elements from $\widehat{S}$). By ${}^{i}A$ we always mean the image of the *element* A.

2.2 Formal Language

To properly formulate the transfer principle which is the heart of nonstandard analysis, we have to work with a formal language which is defined as follows:

The formal language $\mathscr{L}$ has the following *atomic symbols* (=the "letters" from which our language is built):

1. The *variables* which are symbols from a countable set (in practice, letters from the Roman alphabet are used, but one can add more letters if required: the number of variables is not restricted). We occasionally underline variables to distinguish them from constants.
2. The *constants* which form a set of symbols sufficiently large to be put in one-to-one correspondence with the elements of whatever structures are under consideration. For example, if we consider the superstructure of $S = \mathbb{R}$, each real number is such a constant; also each set of numbers, etc. When we explicitly write down a formula, we use of course the traditional notation for the constants whenever there are some (e.g. 1, 2, ..., e, π, $\{1/2, \varnothing\}$). The set of all constants is denoted by $\operatorname{cns}(\mathscr{L})$. We already point out that in all our applications, constants will represent sets (or atoms).
3. The *basic predicates* $\in$ and $=$.
4. The *separation symbols* : (and).
5. The *logical connectives* $\wedge$, $\vee$, $\implies$, $\iff$ and $\neg$.
6. The *quantifiers* $\forall$ and $\exists$.

We note that in the general theory of formal language, one allows a general collection of basic predicates (not only $\in$ and $=$) and also allows basic functors; however, we confine ourselves throughout to languages of the above type.

Not all combinations of the above symbols are admissible in our language. Only the so-called *well-formed formulas* (*wffs*) are admissible. These are defined inductively:

The smallest wffs are the *atomic formulas* $a = b$ and $a \in b$ where a and b are variables or constants. If α and β are wffs, then also $(\alpha \vee \beta)$, $(\alpha \wedge \beta)$, $(\alpha \implies \beta)$, $(\alpha \iff \beta)$, and $(\neg\alpha)$ are wffs. Moreover, if α does not already contain one of $\forall \underline{x}$ or $\exists \underline{x}$, then $(\forall \underline{x} : \alpha)$ and $(\exists \underline{x} : \alpha)$ are wffs. In these formulas, α is called the *scope of the quantifier* $\forall \underline{x}$ resp. $\exists \underline{x}$.

For simpler notation, we sometimes add or eliminate braces in a wff, if there is no ambiguity.

An occurrence of a variable $\underline{x}$ in a wff α is called *free*, if it is not the occurrence in a quantifier $\forall \underline{x}$ or $\exists \underline{x}$ and not within the scope of such a quantifier. All other occurrences of $\underline{x}$ are called *bound*. For example, in the formula $(\forall \underline{x} : \underline{x} = \underline{x}) \wedge (\underline{x} = \underline{y})$ the first three occurrences are bound while the last one is free.

If $\underline{x}_1, \ldots, \underline{x}_n$ occurs freely in a wff α, we sometimes point out this fact by writing $\alpha(\underline{x}_1, \ldots, \underline{x}_n)$. In this case we mean by $\alpha(a_1, \ldots, a_n)$ the string where each free occurrence of $\underline{x}_i$ is replaced by a_i. We use this notation also if not all of the variables $\underline{x}_i$ actually occur (freely) in α. In any case, the above notation does *not* mean that there are no other free variables in α (unless we state this explicitly).

If all occurrences of all variables in α are bound, then α is called a *sentence*. Otherwise, α is called a *predicate*. In other words: α is a sentence if variables occur only in quantifiers or the scope of quantifiers; *all other objects of sentences are constants* (here, we do not count e.g. logical connectives as an "object" of a sentence). We point this out, since this means that although we think of a sentence that it concerns sets, it actually concerns only constants: In a sentence like $\forall \underline{x} : (\underline{x} \in \{\varnothing\} \implies \forall \underline{y} \in \underline{x} : (\neg \underline{y} = \underline{y}))$ the symbol $\{\varnothing\}$ is a *constant* (otherwise this would not be a sentence in our formal language). To solve this "paradox", one has to think of $\{\varnothing\}$ as a constant which represents the set which contains the empty set as its only element; under this *interpretation*, the above sentence should be true. However, the pure "symbol" $\{\varnothing\}$ is not this set but just a name (the "constant") which traditionally is interpreted as this set. We will make the term *interpretation* precise in Section 2.3. However, we already point out that later we will use *constants* of e.g. the form $\{\underline{x} : \alpha(\underline{x})\}$ where $\alpha(\underline{x})$ is some formula of our language; under the traditional interpretation it is clear what we mean by such constants.

For most applications, it suffices to consider so-called bounded quantifiers: The quantifiers $\forall \underline{x}$ resp. $\exists \underline{x}$ are *transitively bounded* if they occur in the form $\forall \underline{x} : (\underline{x} \in A \implies (\alpha))$ resp. $\exists \underline{x} : (\underline{x} \in A \wedge (\alpha))$ where A is either a constant or a variable of the language. In this connection, we use the shortcut $\forall \underline{x} \in A : \alpha$ resp. $\exists \underline{x} \in A : \alpha$ for the above formulas. If A is a constant, we call the quantifier *bounded*. We call a wff *(transitively) bounded*, if all its quantifiers are (transitively) bounded.

The term "bounded" is not used uniquely in literature: Sometimes bounded wffs are meant, and sometimes transitively bounded wffs are meant. The reason for our term "transitively bounded" in this connection will become clear later on.

Without further mention, we make use of the common abbreviations, e.g. $(x \notin y)$ means $(\neg x \in y)$, and $\forall \underline{x}, \underline{y} \in A, \underline{z} \in B : \alpha$ means $\forall \underline{x} \in A : \forall \underline{y} \in A : \forall \underline{z} \in B : \alpha$ (here, we already eliminated some superfluous braces).

Remark 2.2. We already point out that the symbols $\subseteq$ (subset) and $\mathcal{P}$ (powerset) are not atomic symbols. Instead, we define $A \subseteq B$ as the shortcut of the formula $\forall \underline{x} : (\underline{x} \in A \implies \underline{x} \in B)$. The symbol $\mathcal{P}$ will not even be used as some shortcut: When A is a constant (other cases will not occur), we always consider $\mathcal{P}(A)$ as a *constant* (the meaning of this constant is evident).

2.3 Interpretations

So far, we have defined only the *syntax* of a formal language $\mathscr{L}$. Now we are going to define the *semantic*, i.e. we associate a truth value to each of its sentences. The reader is advised to read this section carefully, because the interpretation of a sentence is not what it appears to be at first glance.

The reader should note that a formal language $\mathscr{L}$ is defined through its constants $\operatorname{cns}(\mathscr{L})$. An *abstract interpretation map* I is a one-to-one map of a subset $\operatorname{dom}(I) \subseteq \operatorname{cns}(\mathscr{L})$ into a set $\mathscr{S}$ which is equipped with two binary relations $\in^*$ and $=^*$. Each interpretation map gives rise to an interpretation of all sentences of $\mathscr{L}$ which have the form that all constants occurring in these sentences are in $\operatorname{dom}(I)$:

Given such a sentence α, we define the *interpreted sentence* ${}^I\alpha$ as follows: We replace all occurrences of constants $c \in \operatorname{dom}(I)$ by the interpreted value $I(c)$, and all occurrences of $\in$ and $=$ by $\in^*$ and $=^*$, respectively. The *truth value* of the sentence α under the interpretation map I is then defined as the truth value of the interpreted sentence ${}^I\alpha$ which in turn is defined in the obvious way by induction on the structure of α:

The formula $x \in^* y$ resp. $x =^* y$ is true if and only if the pairs (x, y) belong to the relations $\in^*$ resp. $=^*$; the logical connectives have their usual meaning, e.g. $\alpha \iff \beta$ is true if and only if α and β have the same truth value. Finally, the quantified expressions $\forall \underline{x} : \alpha$ resp. $\exists \underline{x} : \alpha$ are true if α is true for every resp. at least one value of $\underline{x} \in \mathscr{S}$ (i.e. if we replace the free occurrences of $\underline{x}$ in α by the corresponding value).

This is explained best by means of an example: Let M be a constant of the language, and let α be the sentence

$$\exists \underline{x} : \forall \underline{y} : (\underline{y} \in M \implies \underline{x} \in \underline{y}).$$

Then ${}^I\alpha$ is the sentence

$$\exists \underline{x} : \forall \underline{y} : (\underline{y} \in^* I(M) \implies \underline{x} \in^* \underline{y}).$$

By definition, this sentence is true (i.e. α is true under the interpretation map I) if and only if there is an element $\underline{x} \in \mathscr{S}$ such that for each element $\underline{y} \in \mathscr{S}$ for which the pair $(\underline{y}, I(M))$ belongs to the relation $\in^*$ also the pair $(\underline{x}, \underline{y})$ belongs to the relation $\in^*$.

The point which causes difficulties here is that the interpretation takes place in $\mathscr{S}$ and not only in $\operatorname{rng}(I)$, i.e. for quantified variables ($\underline{x}$ and $\underline{y}$ in the above example) we actually take *all* elements of $\mathscr{S}$ into account and not only the elements arising as the image of a constant. In the above example, this means in particular that $\underline{x}$ (and similarly $\underline{y}$) only has to exist in $\mathscr{S}$ and need not necessarily be of the form $\underline{x} = I(c)$ with some $c \in \operatorname{cns}(\mathscr{L})$ (if I is not onto). If I is not onto, this leads to strange effects, as we shall see soon.

The case most important for us is that $\mathscr{S}$ is a superstructure, and that $\in^*$ and $=^*$ are defined in the set-theoretical way, i.e. for elements $x, y \in \mathscr{S}$ the relation $x \in^* y$ resp. $x =^* y$ holds if and only if (in the set-theoretical meaning) $x \in y$ resp. $x = y$. In this special case, we say that I is an *interpretation map in set theory.*

The essential point is that if I is *not onto*, we do not get the usual interpretation of $\mathscr{L}$ as a set theoretical formula.

Let us give an example: Assume that the constants of the language $\operatorname{cns}(\mathscr{L})$ are the elements of a superstructure $\widehat{S}$. Let I be an interpretation map in set theory, say $I : \operatorname{cns}(\mathscr{L}) \to \widehat{T}$. If α denotes the formula "$A \subseteq B$" (with constants $A, B \in \mathscr{S}$) which we use as a shortcut for

$$\forall \underline{x} : (\underline{x} \in A \implies \underline{x} \in B),$$

the interpreted formula ${}^I\alpha$ becomes

$$\forall \underline{x} : (\underline{x} \in^* I(A) \implies \underline{x} \in^* I(B)),$$

and so (since I is an interpretation map in set theory, i.e. $\in^*$ has the usual meaning of the element relation in $\widehat{T}$) α becomes defined as "true" (under the interpretation map I) if and only if $I(A) \subseteq I(B)$ in the usual set-theoretical sense.

This sounds natural at first glance, but if the set $I(A)$ contains elements $\underline{x}$ which do not belong to $\operatorname{rng}(I)$, this is strange because the truth of the sentence

"$A \subseteq B$" then depends on *elements which cannot be described by the language $\mathscr{L}$, because $\mathscr{L}$ has no constants for it*: In the above example the variable $\underline{x}$ runs in ${}^{I}\alpha$ over *all* elements of $I(A)$, not only over the elements of the form $I(x)$ where $I(x) \in I(A)$.

We point this out once more: Even if some sentence $\alpha(c)$ holds for any constant c *this does not imply that the sentence* $\forall \underline{x} : \alpha(\underline{x})$ is true; similarly, if $\exists \underline{x} : \alpha(\underline{x})$ is true, *it does not follow that there is some constant* c *such that* $\alpha(c)$ *is true.* The reason is that it might happen that the language has no constant for the object in question (an exception of this rule is when I is onto).

Let us summarize: If a set theoretical interpretation map I is onto, then the interpretation has the "canonical" meaning. However, if I is not onto, then also elements may play a role which cannot be named in the language.

The above observation is the heart of nonstandard analysis: For example, let $\mathscr{L}$ denote the "natural language" over the superstructure $\widehat{S}$ of $S = \mathbb{R}$ (in particular, the constants of the language are the elements of this superstructure, i.e. $\operatorname{cns}(\mathscr{L}) = \widehat{\mathbb{R}}$). Assume that $I : \operatorname{cns}(\mathscr{L}) \to \widehat{T}$ is an interpretation map into another superstructure $\widehat{T}$ with the property that $\operatorname{rng}(I)$ does not contain all elements of $I(\mathbb{R})$ (here, $\mathbb{R}$ stands for the constant $\mathbb{R}$ in the language $\mathscr{L}$), then the expression $\forall \underline{x} \in \mathbb{R}$ becomes interpreted under I in such a way that x runs actually *through more elements than the elements of* $I(\mathbb{R})$. If the interpretation I is "canonical" in the sense that $I(\mathbb{R})$ has the same "structure" as $\mathbb{R}$ (we will make this more precise, soon), the interpretation thus adds additional ("nonstandard") elements to $\mathbb{R}$ (later we will call these elements *internal*; it turns out that the elements added in this way may be considered as the "infinitesimals").

Chapter 2

Nonstandard Models

§3 The Three Fundamental Principles

3.1 Elementary Embeddings and the Transfer Principle

We shall now make use of non-surjective interpretation maps in set theory (see Section 2.3) to add nonstandard elements to a given superstructure:

Let S be a given set of atoms, and $\widehat{S}$ be the corresponding superstructure. Consider a language $\mathscr{L}$ together with a *surjective* interpretation map $I : \mathrm{dom}(I) \to \widehat{S}$ where $\mathrm{dom}(I) \subseteq \mathrm{cns}(\mathscr{L})$. Then the set K_0 of true sentences whose constants are taken from $\mathrm{dom}(I)$ "describes" the superstructure $\widehat{S}$ in the canonical set-theoretical way.

Assume now that we have another interpretation map $I' : \mathrm{dom}(I') \to \widehat{T}$ for $\mathscr{L}$ in set theory.

Definition 3.1. We call the mapping $* = I' \circ I^{-1} : \widehat{S} \to \widehat{T}$ *elementary* (or *elementary embedding*) if the following holds:

1. Each *transitively bounded* sentence from K_0 (true under the interpretation I) is a true sentence under the interpretation I'.
2. ${}^*S = T$.

Since $T = {}^*S$, we will often simply write $* : \widehat{S} \to \widehat{{}^*S}$. In this connection, $\widehat{S}$ is called the *standard universe*. We shall see that $\widehat{{}^*S}$ contains $\widehat{S}$ in a certain sense, but moreover (for "interesting" maps $*$) it may contain also additional ("nonstandard") elements: Indeed, if $C \in \mathrm{dom}(I)$, the expression $\forall \underline{x} \in C : \ldots$ becomes interpreted under I' in such a way that $\underline{x}$ runs actually through all elements of $I'(C)$ and thus *through more elements than the elements of* $I'(\{c : c \in C\})$.

However, $\widehat{{}^*S}$ is even "too" large in a certain sense, and so we will reserve the name "nonstandard universe" for a certain subset of $\widehat{{}^*S}$ which will be defined later. The mapping $*$ is considered as the embedding of the standard world into the nonstandard world.

For this reason, we call every element in the range of $*$ a *standard element*. Hence, an element of $\widehat{{}^*S}$ is standard, if it can be written in the form ${}^*A$ with $A \in \widehat{S}$. This notion may be slightly confusing, since one would expect that "standard elements" are elements of the standard world $\widehat{S}$. However, since $*$ is injective (this follows from the definition or also from Lemma 3.5 below), the map $*$ provides a one-to-one correspondence between the elements of $\widehat{S}$ and the standard elements. We call a standard element a *standard entity*, if it is not an atom in $\widehat{{}^*S}$ (no confusion will be possible: It follows from Lemma 3.5 below that standard elements are precisely those elements ${}^*a$ where $a \in \widehat{S}$ is not an atom in $\widehat{S}$).

If $A \in \widehat{S}$ is an entity in the standard world, we are also interested in the *standard copy* of A which is denoted by

$$^{\sigma}A = \{{}^*a : a \in A\}.$$

As already pointed out at the end of Section 2.3, an interpretation may add more elements to a set. For this reason, it is not surprising that we always have ${}^{\sigma}A \subseteq {}^*A$ (Lemma 3.5), but that the inclusion may be strict. It is a good idea to think of $*$ as a "blow-up functor".

Definition 3.2. An elementary map $* : \widehat{S} \to \widehat{{}^*S}$ is called a *nonstandard map* (or *nonstandard embedding*) if ${}^{\sigma}A \neq {}^*A$ for each infinite entity $A \in \widehat{S}$.

The existence of nonstandard embeddings will be the topic of §4. The requirement that ${}^{\sigma}A \neq {}^*A$ for *each* infinite entity $A \in \widehat{S}$ appears rather restrictive. However, we will see later (Theorem 3.22) that already the existence of *some* countable infinite set A with ${}^{\sigma}A \neq {}^*A$ implies that $*$ is a nonstandard embedding.

Sometimes in literature, it is additionally required that $S \subseteq T$ and ${}^*s = s$ for each $s \in S$, see e.g. [LR94]. However, this is more or less a formal restriction, since it follows from Lemma 3.5 below that $*$ maps S into T and is one-to-one. Thus, the question whether ${}^*s = s$ ($s \in S$) is just a question of naming the atoms in T.

Mathematically, we still have to prove that the predicate "elementary" in Definition 3.1 is actually well-defined, i.e. that it depends only on the map $*$:

Proposition 3.3. *Definition* 3.1 *depends only on* $*$ *and not on the particular choice of the language* $\mathscr{L}$ *or the interpretation maps* I *and* I'.

Proof. Let $\mathscr{L}_0$ be another language, and $I_0 : \operatorname{dom}(I_0) \to \widehat{S}$ and $I_0' : \operatorname{dom}(I_0') \to \widehat{T}$ be other interpretation maps with $\operatorname{dom}(I_0), \operatorname{dom}(I_0') \subseteq \operatorname{cns}(\mathscr{L}_0)$ such that $I_0' \circ I_0^{-1} = * = I' \circ I^{-1}$ (in particular, I_0 is onto). Assuming that $*$ is elementary with respect to $\mathscr{L}, I, I'$ we shall prove now that $*$ is elementary with respect to $\mathscr{L}_0, I_0, I_0'$. Exchanging the roles of $\mathscr{L}, I, I'$ with $\mathscr{L}_0, I_0, I_0'$ in this conclusion, we find the desired equivalence.

Thus, let a sentence α_0 of the language $\mathscr{L}_0$ with constants from $\operatorname{dom}(I_0)$ be given which is true under the interpretation map I_0. This means that ${}^{I_0}\alpha_0$ is a true sentence of set theory. Then the sentence $\alpha = {}^{I^{-1}}({}^{I_0}\alpha_0)$ from $\mathscr{L}$ is true under the interpretation map I, because ${}^{I}\alpha = {}^{I_0}\alpha_0$. Since $*$ is elementary with respect to $\mathscr{L}, I, I'$, we may conclude that α is true under the interpretation map I', i.e. ${}^{I'}\alpha$ is a true sentence of set theory. But we have ${}^{I'}\alpha = {}^{I' \circ I^{-1} \circ I_0}\alpha_0 = {}^{I_0'}\alpha_0$. Hence, α_0 is true under the interpretation I_0', and the property 1. of Definition 3.1 is satisfied with respect to $\mathscr{L}_0, I_0, I_0'$. □

Let us now discuss the requirements of Definition 3.1 step by step. One requirement is apparently that $I' \circ I^{-1}$ is a function, i.e. that $\operatorname{dom}(I') \supseteq \operatorname{dom}(I)$. However, this is already a consequence of the requirement 1.: Indeed, K_0 contains all sentences of the form $c = c$ where $c \in \operatorname{dom}(I)$. Since this sentence is bounded and thus true under the interpretation map I' by 1. (in particular, it *has* an interpretation), we must have $c \in \operatorname{dom}(I')$. Hence, $\operatorname{dom}(I') \supseteq \operatorname{dom}(I)$, as claimed.

But the requirement 1. of Definition 3.1 implies much more: Actually, this is the key property of nonstandard analysis:

Proposition 3.4 (Transfer principle, First version)**.** *Let* $* = I' \circ I^{-1}$ *be an elementary map. A transitively bounded sentence whose constants are taken from* $\operatorname{dom}(I)$ *is true under the interpretation map* I *if and only if it is true under the interpretation map* I'*.*

Proof. Let α be a sentence whose constants are taken from $\operatorname{dom}(I)$. If α is true, then $\alpha \in K_0$, and by property 1., α has a true interpretation under I'. Conversely, if α is false, then $\neg\alpha$ is true, i.e. $\neg\alpha \in K_0$, and so $\neg\alpha$ has a true interpretation under I' which means that α has a false interpretation under I'. □

The transfer principle is sometimes also called *Leibniz's principle*. The reason is that this principle implies, as we will discuss later, that the hyperreal numbers (with infinitesimals) satisfy the same "formal" properties as the real numbers: This was Leibniz's demand which we mentioned in Section 1.1.

We note that in older references on nonstandard analysis like [SL76, Lux73], the first property of Definition 3.1 (i.e. the transfer principle) is required only for bounded sentences (not for *transitively* bounded sentences). In contrast, in the book [LR94], the transfer principle is assumed to hold for even more general formulas (the range of a variable in a quantifier can be a so-called term). We shall discuss this later. However, we stress that the transfer principle does not hold for *all* sentences if $*$ is a nonstandard map (if we exclude the case that S is finite which becomes rather trivial as we shall see).

The second requirement in Definition 3.1 is of minor importance: This property implies together with the transfer principle that $*$ maps entities of $\widehat{S}$ into entities of $\widehat{T}$. Such a requirement is necessary since we are formally working with a set theory with atoms: The transfer principle alone is not sufficient to e.g. determine the value ${}^*\varnothing$. Indeed, the transfer principle just implies that ${}^*\varnothing$ is "something which contains no elements"; however from this property we do not know whether this happens because ${}^*\varnothing = \varnothing$ or because ${}^*\varnothing$ is an atom. Of course, we want to have the first alternative.

Lemma 3.5. *Let $* : \widehat{S} \to \widehat{T}$ be an elementary map. Then for any $x, y \in \widehat{S}$ the following holds:*

1. *We have $x = y$, $x \in y$, resp. $x \subseteq y$ if and only if ${}^*x = {}^*y$, ${}^*x \in {}^*y$, resp. ${}^*x \subseteq {}^*y$.*
2. *x is an atom resp. an entity in $\widehat{S}$ if and only if ${}^*x$ is an atom resp. an entity in $\widehat{T} = \widehat{{}^*S}$.*
3. *${}^\sigma A \subseteq {}^*A$ for any entity $A \in \widehat{S}$.*

Proof. Let c and d be the constants from the language $\mathscr{L}$ which denote x resp. y, i.e. $c = I^{-1}(x)$ and $d = I^{-1}(y)$. Then we have $x = y$, $x \in y$, resp. $x \subseteq y$ if and only if the sentence $c = d$, $c \in d$, resp. $\forall \underline{z} \in c : \underline{z} \in d$ is true. This is the case if and only if these sentences are true under the interpretation I', i.e. if and only if (in the usual set-theoretical sense) $I'(c) = I'(d)$, $I'(c) \in I'(d)$, resp. $\forall \underline{z} \in I'(c) : \underline{z} \in I'(d)$. But this means ${}^*x = {}^*y$, ${}^*x \in {}^*y$ resp. ${}^*x \subseteq {}^*y$.

By definition, ${}^*x$ is an atom in $\widehat{T}$ if and only if ${}^*x \in T$. Since $T = {}^*S$, this is the case if and only if ${}^*x \in {}^*S$. But by what we just proved this is equivalent to $x \in S$, i.e. to the fact that x is an atom in $\widehat{S}$. Thus, ${}^*x$ is an atom if and only if x is an atom. But this means also that ${}^*x$ is an entity if and only if x is an entity.

The last statement follows from the fact that $b \in {}^\sigma A$ means $b = {}^*a$ for some $a \in A$. By what we had proved before, the latter means ${}^*a \in {}^*A$, i.e. $b \in {}^*A$. □

From now on, we will assume without loss of generality that I is the identity (by just renaming the constants in $\operatorname{cns}(\mathscr{L})$ if necessary). This means that each element of $\widehat{S}$ is simultaneously a constant in the language $\mathscr{L}$, and each formula in $\mathscr{L}$ whose constants are in $\operatorname{dom}(I) = \widehat{S}$ is simultaneously a formula in standard set theory. By this convention we will henceforth not have to distinguish between such a formula and its interpretation in set theory. Since this is confusing in some occasions, we will sometimes use the symbol I anyway.

With the above convention, it makes now sense to consider all the other convenient shortcuts commonly used in set theory as part of our language. Let us give a small list:

Proposition 3.6. *If $\varphi, X, X_1, \dots, X_n$ are constants or variables which do not represent atoms, and $x_1, \dots, x_n$ are variables or constants, then the following sentences resp. predicates are abbreviations of transitively bounded formulas:*

1. $X_1 \subseteq X_2$, $X = X_1 \setminus X_2$, $X = X_1 \cup X_2$, $X = X_1 \cap X_2$.
2. $X = \{x_1, \dots, x_n\}$.
3. $X = (x_1, \dots, x_n)$, $(x_1, \dots, x_n) \in X$.
4. $X = X_1 \times \cdots \times X_n$.
5. $\varphi \subseteq X_1 \times \cdots \times X_n$. *(i.e. φ is a relation on $X_1, \dots, X_n$).*
6. $\varphi : X_1 \to X_2$, $x_2 = \varphi(x_1)$.
7. *X is transitive.*

The mentioned variables/constants are the only free variables or constants occurring in these formulas. Moreover, if X, X_1, X_2 and $x_1, \dots, x_n$ are constants, then the formulas for 1. *and* 2. *are even bounded.*

Proof. The formula $X_1 \subseteq X_2$ is a shortcut for $(\forall \underline{x} \in X_1 : \underline{x} \in X_2)$, and $X = X_1 \setminus X_2$ is a shortcut for $((\forall \underline{x} \in X : (\underline{x} \in X_1 \wedge \underline{x} \notin X_2)) \wedge (\forall \underline{x} \in X_1 : (\underline{x} \notin X_2 \implies \underline{x} \in X)))$; the other formulas in 1. are similar, and their formulation is left to the reader. The formula $X = \{x_1, x_2\}$ is a shortcut for $(x_1 \in X \wedge x_2 \in X \wedge (\forall \underline{x} \in X : (\underline{x} = x_1 \vee \underline{x} = x_2)))$. Using this shortcut, we may treat the formula $X = (x_1, x_2)$ as a shortcut for $(\exists \underline{x} \in X : \exists \underline{y} \in X : (\underline{x} = \{x_1\} \wedge \underline{y} = \{x_1, x_2\}) \wedge X = \{\underline{x}, \underline{y}\})$. The sentence $(x_1, x_2) \in X$ can then be considered as a shortcut for $(\exists \underline{x} \in X : \underline{x} = (x_1, x_2))$. Using also this shortcut, we may treat $X = X_1 \times X_2$ as a shortcut for $((\forall \underline{x} \in X : \exists \underline{x}_1 \in X_1, \underline{x}_2 \in X_2 : \underline{x} = (\underline{x}_1, \underline{x}_2)) \wedge (\forall \underline{x}_1 \in X_1, \underline{x}_2 \in X_2 : (\underline{x}_1, \underline{x}_2) \in X))$. Now the formula $\varphi \subseteq X_1 \times X_2$ may be written with the previous shortcuts. The cases $n \geq 3$ in the previous shortcuts are similar and left to the reader (one may use an induction, if one wants to be extraordinarily precise). The sentence $\varphi : X_1 \to X_2$ is a shortcut for $((\varphi \subseteq X_1 \times X_2) \wedge (\forall \underline{x} \in X_1, \underline{y}_1, \underline{y}_2 \in X_2 : (((\underline{x}, \underline{y}_1) \in \varphi \wedge (\underline{x}, \underline{y}_2) \in \varphi) \implies \underline{y}_1 = \underline{y}_2)))$. In this case, $\varphi(x_1) = x_2$ is a shortcut for $(x_1, x_2) \in \varphi$. "X is transitive" is a shortcut for $\forall \underline{y} \in X : \forall \underline{x} \in \underline{y} : \underline{x} \in X$. □

The requirement that $X, X_1, \dots, X_n$ do not represent atoms is not essential for all of the formulas, but e.g. for $X = X_1 \setminus X_2$: If e.g. $X_1 = X_2$, then any atom X would satisfy the formula given in the above proof.

Note that the formula in the above proof given for $X = (x_1, x_2)$ is *not* bounded, even if x_1, x_2, and X are constants (you have to write down the formula more explicitly to see this). However, this formula is transitively bounded.

Of course, the list from Proposition 3.6 is by far not complete. Anyway, we will tacitly make use of simple extensions of this proposition; for example we consider a formula like $\{x_1, \dots, x_n\} \in X$ as an abbreviation of $(\exists \underline{x} \in X : \underline{x} = \{x_1, \dots, x_n\})$, and so on.

The $*$-*transform* of a transitively bounded formula α with constants in $\widehat{S}$ is the formula ${}^*\alpha$ where any occurring constant c is replaced by ${}^*c$ (in contrast, variables remain unchanged!). In other words: If α is a set-theoretical formula about the superstructure $\widehat{S}$, then ${}^*\alpha$ is the formula which arises from α if each element of $\widehat{S}$ in the formula is replaced by its image under the mapping $*$.

Now we are in a position to formulate the main principle of nonstandard analysis:

Theorem 3.7 (Transfer principle)**.** *Let $* : \widehat{S} \to \widehat{{}^*S}$ be elementary. The transitively bounded formula α with constants in $\widehat{S}$ is true if and only if its $*$-transform ${}^*\alpha$ is true.*

Proof. Since I is the identity, the formula α is true if and only if it is true under the interpretation map I. By the first transfer principle this is the case if and only if α is true under the interpretation map I', i.e. if and only if ${}^{I'}\alpha$ is a true formula in set theory. Now observe that ${}^{I'}\alpha = {}^*\alpha$. □

As remarked earlier, the definition of elementary maps (and thus the transfer principle) slightly varies in literature: While e.g. in [SL76, Lux73] only bounded sentences are considered, in the book [LR94] even sentences which are slightly more general than transitively bounded sentences are considered.

Since already the simple formula $x = (x_1, x_2)$ has no simple bounded formulation, the first approach appears rather restrictive.

On the other hand, all results in this section are of course based on the assumption that there exists a nontrivial elementary map $*$. In the mentioned literature (and also in §4), this map is constructed in the same way. Only the *proof* that the map has the required properties varies. Of course, the more properties one wants to have the more technical is this proof. Our restriction to transitively bounded formulas is reasonably mild in practice and still allows us to give a proof where the main ideas are not hidden beyond technical details. The reader who needs a (slightly) more general form of the transfer principle is referred to [LR94].

3.2 The Standard Definition Principle

The transfer principle implies that the mapping $*$ preserves much of the structure of $\widehat{S}$. In fact, we have:

Theorem 3.8 (Standard Definition Principle)**.** *An entity $A \in \widehat{{}^*S}$ is standard if and only if it can be written in the form*

$$A = \{\underline{x} \in B : \alpha(\underline{x})\}$$

where α is a transitively bounded predicate with $\underline{x}$ as its only free variable, and B and all elements (=constants) occurring in α are standard elements.

More precisely, if $B = {}^*B_0$*, and the elements occurring in* α *are* ${}^*B_1, \ldots, {}^*B_n$ *(we write* $\alpha = \alpha(\underline{x}, {}^*B_1, \ldots, {}^*B_n)$ *to point this out), then*

$$A = \{\underline{x} \in {}^*B_0 : \alpha(\underline{x}, {}^*B_1, \ldots, {}^*B_n)\} = {}^*\{\underline{x} \in B_0 : \alpha(\underline{x}, B_1, \ldots, B_n)\}$$

where $\alpha(\underline{x}, B_1, \ldots, B_n)$ *denotes the formula* α *where all occurrences of* B_i *are replaced by* ${}^*B_i$ $(i = 1, \ldots, n)$*.*

Proof. Necessity is clear, since we have $A = \{\underline{x} \in B : \underline{x} = \underline{x}\}$ for the choice $B = A$. Sufficiency follows from the second statement. To prove the latter, let $A_0 := \{\underline{x} \in B_0 : \alpha(\underline{x}, B_1, \ldots, B_n)\}$ and note that

$$\forall \underline{x} \in B_0 : (\underline{x} \in A_0 \iff \alpha(\underline{x}, B_1, \ldots, B_n))$$

is a true sentence which is transitively bounded. Hence, the transfer principle implies that its $*$-transform is true. But the latter reads

$$\forall \underline{x} \in {}^*B_0 : (\underline{x} \in {}^*A_0 \iff \alpha(\underline{x}, {}^*B_1, \ldots, {}^*B_n)).$$

Since ${}^*A_0 \subseteq {}^*B_0$ by Lemma 3.5, this means that ${}^*A_0 = \{\underline{x} \in {}^*B_0 : \alpha(\underline{x}, {}^*B_1, \ldots, {}^*B_n)\}$, i.e. ${}^*A_0 = A$, as claimed. □

Together with Proposition 3.6 we find that each elementary embedding is a superstructure monomorphism:

Definition 3.9. A map $* : \widehat{S} \to \widehat{T}$ is a *superstructure monomorphism* if it is one-to-one, and if for all entities $A, B \in \widehat{S}$ and all $x_1, \ldots, x_n \in \widehat{S}$ the following holds:

1. $*$ preserves $\in$: $a \in A$ implies ${}^*a \in {}^*A$.
2. $*$ preserves finite sets:
$$ {}^*\{x_1, \ldots, x_n\} = \{{}^*x_1, \ldots, {}^*x_n\}.$$
3. $*$ preserves grouping:
$$ {}^*(x_1, \ldots, x_n) = ({}^*x_1, \ldots, {}^*x_n).$$
4. $*$ preserves basic set operations: ${}^*\varnothing = \varnothing$, ${}^*(A \cup B) = {}^*A \cup {}^*B$, ${}^*(A \cap B) = {}^*A \cap {}^*B$, ${}^*(A \setminus B) = {}^*A \setminus {}^*B$, ${}^*(A \times B) = {}^*A \times {}^*B$.
5. $*$ preserves domains and ranges of n-ary relations Φ in $\widehat{X}$ and commutes with permutations of the variables:

   For example, if Φ is a binary relation, then ${}^*\Phi$ is a relation, and $\mathrm{dom}({}^*\Phi) = {}^*\mathrm{dom}(\Phi)$, $\mathrm{rng}({}^*\Phi) = {}^*\mathrm{rng}(\Phi)$. If another relation Ψ on B, A has the property that $(y, x) \in \Phi$ if and only if $(x, y) \in \Psi$, then $(z, w) \in {}^*\Phi$ if and only if $(w, z) \in {}^*\Psi$ (i.e. ${}^*(\Phi^{-1}) = ({}^*\Phi)^{-1}$). An analogous statement holds for relations of more than two variables.

6. $*$ preserves the equality relation:

$$^*\{(\underline{x},\underline{x}) : \underline{x} \in A\} = \{(\underline{x},\underline{x}) : \underline{x} \in {}^*A\}.$$

7. $*$ preserves atomic standard definitions of sets:

$$^*\{(\underline{x},\underline{y}) : \underline{x} \in \underline{y} \in A\} = \{(\underline{x},\underline{y}) : \underline{x} \in \underline{y} \in {}^*A\}.$$

We note that many of the properties in Definition 3.9 are actually redundant. For example, the relation ${}^*(A \setminus B) = {}^*A \setminus {}^*B$ implies for the choice $A = B$ that ${}^*\varnothing = \varnothing$.

Theorem 3.10. *Each elementary map $*$ is a superstructure monomorphism.*

Proof. The injectivity of $*$ has already been proved in Lemma 3.5. Concerning the other properties:

1. This was proved in Lemma 3.5.
2. Let C denote the entity $\{x_1, \ldots, x_n\}$. Then $C = \{x_1, \ldots, x_n\}$ is a true and bounded sentence by Proposition 3.6 with C and $x_1, \ldots, x_n$ as the only constants, and so its $*$-transform ${}^*C = \{{}^*x_1, \ldots, {}^*x_n\}$ is true.
3. Analogously with $C = (x_1, \ldots, x_n)$. The only difference is that this statement is only *transitively* bounded.
4. Let C denote the entity $A \setminus B$. Then $C = A \setminus B$ is a true and transitively bounded sentence by Proposition 3.6, and so its $*$-transform ${}^*C = {}^*A \setminus {}^*B$ is true. Note that ${}^*C$ is actually an entity by Lemma 3.5. The proof of the other statements is similar. To see that ${}^*\varnothing = \varnothing$, we may argue as remarked before, or apply the transfer principle to the bounded true sentence $\forall \underline{x} \in \varnothing : \underline{x} \neq \underline{x}$.
5. Let Φ be an entity in $\widehat{S}$ which is a binary relation. We have $\Phi \in S_n$ for some n. Since S_n is transitive (Theorem 2.1), the relation $(\underline{x},\underline{y}) \in \Phi$ implies $\underline{x},\underline{y} \in S_n$. Consequently, $\Phi \subseteq S_n \times S_n$. By Lemma 3.5 and 4., we have ${}^*\Phi \subseteq {}^*(S_n \times S_n) = {}^*S_n \times {}^*S_n$.

If $C := \operatorname{dom}(\Phi)$, then $C = \{\underline{x} \in S_n : (\exists \underline{y} \in S_n : (\underline{x},\underline{y}) \in \Phi)\}$ where $(\underline{x},\underline{y}) \in \Phi$ is the shortcut of Proposition 3.6. The standard definition principle implies ${}^*C = \{\underline{x} \in {}^*S_n \mid \exists \underline{y} \in {}^*S_n : (\underline{x},\underline{y}) \in {}^*\Phi\}$, i.e. ${}^*C = {}^*\operatorname{dom}(\Phi)$. The formula ${}^*\operatorname{rng}(\Phi) = \operatorname{rng}({}^*\Phi)$ is proved analogously.

If Ψ is a relation which satisfies $(y,x) \in \Psi$ if and only if $(x,y) \in \Phi$, then

$$\Psi = \{\underline{z} \in S_n \times S_n \mid \exists \underline{x},\underline{y} \in S_n : (\underline{z} = (\underline{x},\underline{y}) \wedge (\underline{y},\underline{x}) \in \Phi)\},$$

and the standard definition principle implies

$$^*\Psi = \{\underline{z} \in {}^*(S_n \times S_n) \mid \exists \underline{x},\underline{y} \in {}^*S_n : (\underline{z} = (\underline{y},\underline{x}) \wedge (\underline{x},\underline{y}) \in {}^*\Phi)\},$$

which in view of ${}^*(S_n \times S_n) = {}^*S_n \times {}^*S_n$ implies that ${}^*\Psi = \{(y,x) : (x,y) \in \Phi\}$, as claimed.

The case of relations which are not binary is similar and left to the reader.

6. Let $C := \{(\underline{x},\underline{x}) : \underline{x} \in A\} = \{\underline{y} \in A \times A : (\exists \underline{x} \in A : \underline{y} = (\underline{x},\underline{x}))\}$ where $\underline{y} = (\underline{x},\underline{x})$ is the transitively bounded predicate from Proposition 3.6. The standard definition principle thus implies ${}^*C = \{\underline{y} \in {}^*(A \times A) : (\exists \underline{x} \in {}^*A : \underline{y} = (\underline{x},\underline{x})\}$. Since ${}^*(A \times A) = {}^*A \times {}^*A$ by 4., we find ${}^*C = \{(\underline{x},\underline{x}) : \underline{x} \in {}^*A\}$, as claimed.

7. At first glance, one might suspect that for the proof of 7., one might argue similarly to the proof of 6. However, from the definition of the set $C := \{(\underline{x},\underline{y}) : \underline{x} \in \underline{y} \in A\}$ we may not apply the standard definition principle, since it is not clear from which entity the elements $(\underline{x},\underline{y})$ are taken: We need a "universal" entity U with $(\underline{x},\underline{y}) \in U$ whenever $\underline{x} \in \underline{y} \in A$.

Such an entity indeed exists: Since A is an entity, we find some n with $A \in S_n$ (with S_n as in Section 2.1). Since S_n is transitive (Theorem 2.1, 2.), the relation $\underline{y} \in A$ implies $\underline{y} \in S_n$; using the transitivity of S_n once more, we find that $\underline{x} \in \underline{y}$ also implies $\underline{x} \in S_n$. Consequently, $(\underline{x},\underline{y}) \in S_n \times S_n$ whenever $\underline{x} \in \underline{y} \in A$. Hence, $U := S_n \times S_n$ is the required universal entity (U actually *is* an entity by Theorem 2.1).

Now the proof is straightforward: By our choice of U, we have $\forall \underline{y} \in A : \forall \underline{x} \in \underline{y} : (\underline{x},\underline{y}) \in U$, and so $C = \{(\underline{x},\underline{y}) \in U : \underline{x} \in \underline{y} \in A\}$ which implies by the standard definition principle that ${}^*C = \{(\underline{x},\underline{y}) \in {}^*U : \underline{x} \in \underline{y} \in {}^*A\}$. Now observe that an application of the transfer principle to the above mentioned transitively bounded true sentence implies that $\forall \underline{y} \in {}^*A : \forall \underline{x} \in \underline{y} : (\underline{x},\underline{y}) \in {}^*U$ is true. Using this fact in the above formula for ${}^*C$, we have ${}^*C = \{(\underline{x},\underline{y}) : \underline{x} \in \underline{y} \in {}^*A\}$. □

We note that, conversely, each superstructure monomorphism satisfies the transfer principle for bounded sentences and also the standard definition principle for bounded predicates; a proof of this result can be found in [RZ69]. It appears rather reasonable that even superstructure monomorphisms must satisfy the transfer principle for *transitively* bounded sentences, but we did not try to prove this.

Corollary 3.11. *Let $*$ be a nonstandard embedding. Then for any entity A in the standard universe we have ${}^\sigma A \subseteq {}^*A$ with equality if and only if A is finite.*

Proof. The inclusion ${}^\sigma A \subseteq {}^*A$ has been proved in Lemma 3.5, and the equality for finite sets follows from the fact that $*$ is a superstructure monomorphism. The fact that we have inequality for infinite sets is the definition of a nonstandard embedding. □

In particular, if S is finite and so all entities of the standard universe are finite, we always have ${}^\sigma A = {}^*A$. In this case, everything is trivial: After a renaming of the atoms in ${}^*S = {}^\sigma S$, we may assume that ${}^*S = S$ and that $* : S \to {}^*S$ is the identity. Since ${}^\sigma A = {}^*A$ for all sets A, we may conclude that $* : \widehat{S} \to \widehat{{}^*S}$ is the identity.

Corollary 3.12 (Standard Definition Principle for Relations)**.** *An n-ary relation $A \in \widehat{{}^*S}$ is standard if and only if it can be written in the form*

$$A = \{(\underline{x}_1, \dots, \underline{x}_n) \in B_1 \times \cdots \times B_n : {}^*\alpha(\underline{x}_1, \dots, \underline{x}_n)\}$$

where ${}^*\alpha$ is a transitively bounded predicate with $\underline{x}_1, \dots, \underline{x}_n$ as its only free variables, and B and all elements (=constants) occurring in ${}^*\alpha$ are standard elements.

More precisely, if $B_k = {}^*A_k$, then

$$A = {}^*\{(\underline{x}_1, \dots, \underline{x}_n) \in A_1 \times \cdots \times A_n : \alpha(\underline{x}_1, \dots, \underline{x}_n)\}$$

where α denotes the formula ${}^*\alpha$ where all occurrences of constants ${}^*B$ are replaced by B.

Proof. Put $B_0 = A_1 \times \cdots \times A_n$. Then ${}^*B_0 = B_1 \times \cdots \times {}^*B_n$, because $*$ is a superstructure monomorphism. Putting $B = {}^*B_0$, we thus have

$$A = \{\underline{y} \in B \mid \exists \underline{x}_1 \in B_1, \dots, \underline{x}_n \in B_n : (\underline{y} = (\underline{x}_1, \dots, \underline{x}_n) \wedge {}^*\alpha(\underline{x}_1, \dots, \underline{x}_n))\}.$$

The standard definition principle implies

$$A = {}^*\{\underline{y} \in B_0 \mid \exists \underline{x}_1 \in A_1, \dots, \underline{x}_n \in A_n : (\underline{y} = (\underline{x}_1, \dots, \underline{x}_n) \wedge \alpha(\underline{x}_1, \dots, \underline{x}_n))\},$$

and so the statement follows. □

Exercise 5. Prove that it is not possible to describe nonstandard elements by a standard predicate. More precisely, if A is a standard entity and $c \in {}^*A \setminus {}^\sigma A$ is a nonstandard element, prove that there is no standard predicate $\alpha(\underline{x})$ (i.e. all constants in α are standard) such that $\alpha(\underline{x})$ is true for $\underline{x} \in {}^*A$ if and only if $\underline{x} = c$.

In nonstandard analysis it is an important fact that each function in the standard world may be "extended" to a function acting on the corresponding nonstandard sets:

Theorem 3.13. *Let $*$ be elementary. If A, B are standard entities and $f : A \to B$, then ${}^*f : {}^*A \to {}^*B$. Moreover:*

1. *f is one-to-one if and only if ${}^*f$ is one-to-one, i.e. if and only if ${}^*f$ is invertible on its range* $\operatorname{rng}({}^*f) = {}^*\operatorname{rng}(f)$, *and we have* $({}^*f|_{\operatorname{rng}({}^*f)})^{-1} = {}^*(f^{-1}|_{\operatorname{rng}(f)})$.
2. *f is onto if and only if ${}^*f$ is onto.*
3. ${}^*(f(a)) = ({}^*f)({}^*a)$ *for each $a \in A$.*
4. *If $C \subseteq A$, then ${}^*(f|_C) = ({}^*f)|_{{}^*C}$.*

5. *${}^*(f(C)) = {}^*f({}^*C)$ $(C \subseteq A)$ and ${}^*(f^{-1}(D)) = ({}^*f)^{-1}({}^*D)$ $(D \subseteq B)$.*
6. *If $g : B \to C$, then ${}^*(g \circ f) = ({}^*g) \circ ({}^*f)$.*

Proof. The transfer principle applied to the formula $f : A \to B$ from Proposition 3.6 shows that ${}^*f : {}^*A \to {}^*B$.

1. f is one-to-one if and only if the transitively bounded sentence

$$\forall \underline{x}_1, \underline{x}_2 \in A : (\underline{x}_1 \neq \underline{x}_2 \implies f(\underline{x}_1) \neq f(\underline{x}_2))$$

is true, where $f(\underline{x}_1) \neq f(\underline{x}_2)$ is a shortcut for $\exists \underline{y}_1, \underline{y}_2 \in \operatorname{rng}(f) : (\underline{y}_1 = f(\underline{x}_1) \wedge \underline{y}_2 = f(\underline{x}_2) \wedge \underline{y}_1 \neq \underline{y}_2)$ (where we used the shortcut from Proposition 3.6). The $*$-transform of this sentence reads

$$\forall \underline{x}_1, \underline{x}_2 \in {}^*A : (\underline{x}_1 \neq \underline{x}_2 \implies {}^*f(\underline{x}_1) \neq {}^*f(\underline{x}_2))$$

which is true if and only if ${}^*f$ is one-to-one. By the transfer principle the sentences are either both true or both false. The formulas for the range and the inverse have been proved in Theorem 3.10.

2. f is onto if and only if $\operatorname{rng}(f) = B$ which by the injectivity of $*$ is the case if and only if ${}^*\operatorname{rng}(f) = {}^*B$. By Theorem 3.10, this means ${}^*B = \operatorname{rng}({}^*f)$, i.e. ${}^*f$ is onto.

3. Let c denote the constant $f(a)$. Then the sentence $(a, c) \in f$ is true, and so by the transfer principle ${}^*(a, c) \in {}^*f$. Since $*$ is a superstructure monomorphism, we have $({}^*a, {}^*c) = {}^*(a, c)$, and so $({}^*a, {}^*c) \in {}^*f$, i.e. ${}^*c = {}^*f({}^*a)$.

4. We have

$$f|_C = \{(\underline{x}, \underline{y}) \in f : \underline{x} \in C\}.$$

The standard definition principle implies

$${}^*(f|_C) = \{(\underline{x}, \underline{y}) \in f : \underline{x} \in {}^*C\}$$

which means that ${}^*(f|_C) = ({}^*f)|_{{}^*C}$, as claimed.

5. ${}^*(f(C)) = {}^*\operatorname{rng}(f|_C) = \operatorname{rng}({}^*(f|_C)) = \operatorname{rng}({}^*f|_{{}^*C}) = {}^*f({}^*C)$. Since

$$f^{-1}(D) = \{\underline{x} \in A : f(\underline{x}) \in D\},$$

the standard definition principle implies

$${}^*(f^{-1}(D)) = \{\underline{x} \in {}^*A : {}^*f(\underline{x}) \in {}^*D\} = ({}^*f)^{-1}({}^*D).$$

6. Let $h = g \circ f$. Then the sentence $\forall \underline{x} \in A : h(\underline{x}) = g(f(\underline{x}))$ with the evident abbreviations is transitively bounded and true. The transfer principle implies $\forall \underline{x} \in {}^*A : {}^*h(\underline{x}) = {}^*g({}^*f(\underline{x}))$ which means ${}^*h = ({}^*g) \circ ({}^*f)$. □

From the previous proofs, the reader might get the wrong impression that all useful sentences are transitively bounded so that the whole structure of $\widehat{S}$ is preserved by an elementary map $*$. This impression is false:

Although it actually is the case that any useful sentence can be formulated as a transitively bounded formula, their "natural" formulation often is *not* a transitively bounded formula. But the transfer principle only applies to the transitively bounded formulation.

Let us give a key example: Let α denote the sentence

$$\forall A \subseteq B : \beta(A)$$

with some infinite set $B \in \widehat{S}$, and assume that α is true. The sentence α is in general *not* bounded (unless B or β are rather trivial). However, α is equivalent to the sentence

$$\forall A \in \mathcal{P}(B) : \beta(A).$$

This sentence is bounded (here, $\mathcal{P}(B)$ is considered as a constant). Hence, the transfer principle implies that

$$\forall A \in {}^*\mathcal{P}(B) : \beta(A)$$

is true. In contrast, the $*$-transform of the original sentence α would read $\forall A \subseteq {}^*B : \beta(A)$ which can be rephrased as $\forall A \in \mathcal{P}({}^*B) : \beta(A)$. The transfer principle does *not* imply that the latter sentence is true. In fact, this sentence may fail, because we have

$${}^*\mathcal{P}(B) \neq \mathcal{P}({}^*B)$$

for any infinite set $B \in \widehat{S}$ if $*$ is a nonstandard map, as we shall see later.

3.3 The Internal Definition Principle

The reader might have observed that almost all of the arguments in the previous proofs followed the same scheme. Only in the proof of Theorem 3.10 in step 7., an essentially different argument was needed. Actually, this argument is the key to the definition of the nonstandard universe and the internal definition principle which we shall introduce now.

Let us first point out that transitive subsets of $\widehat{{}^*S}$ play a special role in connection with transitively bounded sentences (this is the reason for our terminology *transitively bounded*):

Consider a simple transitively bounded sentence like $\forall \underline{x} \in C : \forall \underline{y} \in \underline{x} : \exists \underline{z} \in \underline{y} : \alpha$ under an interpretation map $I' : \mathrm{cns}(\mathscr{L}) \to \widehat{{}^*S}$ where C is interpreted as an element from a transitive subset $\mathscr{T} \subset \widehat{{}^*S}$. Since $\mathscr{T}$ is transitive, this element is a subset of $\mathscr{T}$, and so $\underline{x}$ runs under the interpretation actually only through elements of $\mathscr{T}$. For the same reason, also $\underline{y}$ runs only through elements of $\mathscr{T}$, and hence also $\underline{z}$ runs only through elements of $\mathscr{T}$. The same argument holds for any transitively bounded sentence. Thus, roughly speaking, for transitively bounded sentences with constants from $\mathscr{T}$, it suffices to know the "universe" $\mathscr{T}$ to find the correct interpretation: It is not necessary to know the whole superstructure $\widehat{{}^*S}$.

Since we are mainly interested in the $*$-transform of sentences, i.e. sentences with only standard constants, it suffices for us to consider the smallest transitive set which contains all the standard constants. This set will be defined as the *nonstandard universe*. It turns out that this set has a simple characterization: The elements of this set are precisely all elements of all standard entities. To see this, we need some preparation:

Lemma 3.14. *Let* $*:\widehat{S}\to\widehat{{}^*S}$ *be elementary. For the sets* S_n *from* Section 2.1*, we have*

1. ${}^*S_n$ *is transitive.*
2. ${}^*S_0\subseteq{}^*S_1\subseteq\cdots\subseteq\widehat{{}^*S}$.
3. ${}^*S_1\in{}^*S_1\in\cdots\in\widehat{{}^*S}$.

Proof. By Theorem 2.1, we know that S_n is transitive, $S_0\subseteq S_1\subseteq\cdots\subseteq\widehat{S}$, and $S_0\in S_1\in\cdots\in\widehat{S}$. Since "$S_n$ is transitive" is a transitively bounded sentence (Proposition 3.6), the transfer principle implies that ${}^*S_n$ is transitive. The inclusions ${}^*S_n\subseteq{}^*S_{n+1}$ and ${}^*S_n\in{}^*S_{n+1}$ follow from Lemma 3.5, and ${}^*S_n\in\widehat{{}^*S}$ follows from $*:\widehat{S}\to\widehat{{}^*S}$. This also implies ${}^*S_n\subseteq\widehat{{}^*S}$, since the superstructure $\widehat{{}^*S}$ is transitive by Theorem 2.1. □

It is important to know that the sets ${}^*S_n$ are *not* the level sets of the superstructure $\widehat{{}^*S}$ as in Section 2.1: The superstructure $\widehat{{}^*S}$ is much larger than the union of the sets ${}^*S_n$, as we shall see. Instead, it turns out that the elements of ${}^*S_n$ form what we call the *nonstandard universe*:

Definition 3.15. Let $*:\widehat{S}\to\widehat{{}^*S}$ be elementary. The elements of standard sets are called *internal*. The *nonstandard universe* $\mathscr{I}$ is the system of all internal elements, i.e.

$$\mathscr{I}:=\bigcup\{{}^*A : A \text{ is an entity in } \widehat{S}\}.$$

An element $x\in\widehat{{}^*S}$ which is not internal is called *external*.

In particular, each *atom* of $\widehat{{}^*S}$ is an internal element (but not each entity, as we shall see).

Proposition 3.16. *The nonstandard universe* $\mathscr{I}$ *is a transitive subset of* $\widehat{{}^*S}$*. Each standard element is internal. Conversely, each internal entity is a subset of some* *standard entity. Moreover,*

$$\mathscr{I} = \bigcup_{n=0}^{\infty} {}^*S_n \tag{3.1}$$

where S_n *are defined as in* Section 2.1*. Hence,* $\mathscr{I}$ *is the smallest transitive subset of* $\widehat{{}^*S}$ *which contains all sets* ${}^*S_n$ *as subsets (or, equivalently, which contains all standard elements).*

Proof. If $x \in \widehat{S}$, then $x \in S_n$ for some n, and so ${}^*x \in {}^*S_n$ by Lemma 3.5. Since S_n is an entity of $\widehat{S}$ (Theorem 2.1, 1.), $\mathscr{I}$ contains all sets ${}^*S_n$ as subsets (which proves one inclusion of (3.1)), and hence in particular ${}^*x \in \mathscr{I}$.

If x is an internal element, then $x \in {}^*A$ for some entity $A \in \widehat{S}$. Then $A \in S_n$ for some n. Recall that S_n is transitive, and so $A \subseteq S_n$, hence ${}^*A \subseteq {}^*S_n$ by Lemma 3.5, i.e. $x \in {}^*S_n$. The formula (3.1) and $\mathscr{I} \subseteq \widehat{{}^*S}$ now follows. If additionally x is no atom, the transitivity of ${}^*S_n$ (Lemma 3.14) implies $x \subseteq {}^*S_n \subseteq \mathscr{I}$ which shows that each internal entity is a subset of some standard entity and, moreover, that $\mathscr{I}$ is transitive. □

Roughly speaking, all elements which can be "described" by transitively bounded formulas are internal (we will make this more precise and give a rigorous proof in the internal definition principle below). At this point, we only want to point out that in a sense the internal sets are precisely those sets which can be "explicitly" constructed within the language of the nonstandard universe. For this reason, internal sets are sometimes called "definable". Nevertheless, many important sets will turn out to be external. The distinction of internal and external sets is one of the most useful concepts in nonstandard analysis (this will become clear later in the applications). Thus, one can hardly overestimate the value of theorems stating that certain sets are internal or external.

Now that we know the above described intuitive "transfer principle for internal sets", it is not surprising that we also have an analogue of the standard definition principle for internal sets which can be formulated more rigorously.

Recall that by Proposition 3.3 we have much freedom in the choice of our language $\mathscr{L}$: As long as $\mathrm{dom}(I)$ and the restriction of I to $\mathrm{dom}(I)$ remains unchanged, we may add or delete any constants to $\mathrm{cns}(\mathscr{L})$. In particular, it is no loss of generality to assume that $\mathrm{rng}(I') = \mathscr{I}$: We just have to establish a one-to-one correspondence between the constants from $\mathrm{cns}(\mathscr{L}) \setminus \mathrm{dom}(I)$ (we may assume that this set has the right cardinality) and the nonstandard elements from $\mathscr{I}$. We call a formula α in this language *internal*.

In other words: A (set-theoretical) formula containing as constants only elements from $\widehat{{}^*S}$ is internal if and only if the elements occurring in the formula all belong to $\mathscr{I}$.

Theorem 3.17 (Internal Definition Principle)**.** *An entity $A \in \widehat{{}^*S}$ is internal if and only if it can be written in the form*

$$A = \{\underline{x} \in B : \alpha(\underline{x})\}$$

where B is an internal entity, and α is a transitively bounded internal predicate with $\underline{x}$ as its only free variable.

Proof. Necessity is trivial: Put $B = A$, and let $\alpha(\underline{x})$ be $\underline{x} = \underline{x}$. To prove sufficiency, let B be internal, and $\alpha(\underline{x})$ be a transitively bounded internal predicate with $\underline{x}$ as its only free variable. Let $B_1, \ldots, B_k$ be the constants (internal elements) which occur in α. The essential step in the proof is to observe that there is some n such that $B =: B_0, B_1, \ldots, B_k \in {}^*S_n$: Indeed, by Proposition 3.16, we find for any i some n with $B_i \in {}^*S_n$. Since ${}^*S_0 \subseteq {}^*S_1 \subseteq \cdots$ by Lemma 3.14, we may assume that n is independent of i. We denote by $\alpha(\underline{x}, \underline{y}_1, \ldots, \underline{y}_k)$ the formula which arises from $\alpha(\underline{x})$ if we replace any occurrence of B_i by $\underline{y}_i$ ($i = 1, \ldots, k$) (we assume that the variables $\underline{y}_i$ did not occur before in α). Now observe that the transitively bounded sentence

$$\begin{gathered}\forall \underline{y}_1, \ldots, \underline{y}_k, \underline{y} \in S_n : \exists \underline{z} \in S_{n+1} : \forall \underline{x} \in S_{n+1} : \\ (\underline{x} \in \underline{z} \iff (\underline{x} \in \underline{y} \wedge \alpha(\underline{x}, \underline{y}_1, \ldots, \underline{y}_k)))\end{gathered}$$

is true: In fact, if $\underline{y}_1, \ldots, \underline{y}_k, \underline{y} \in S_n$ are given, the set $\underline{z} := \{\underline{x} \in \underline{y} : \alpha(\underline{x}, \underline{y}_1, \ldots, \underline{y}_k)\}$ exists by the *comprehension axiom* of formal set theory. Since S_n is transitive, we have $z \subseteq S_n$, and so $\underline{z} \in S_{n+1}$ (a more careful analysis shows that we even have $z \in S_n$, but we do not need this fact). The transfer principle thus implies that

$$\begin{gathered}\forall \underline{y}_1, \ldots, \underline{y}_k, \underline{y} \in {}^*S_n : \exists \underline{z} \in {}^*S_{n+1} : \forall \underline{x} \in {}^*S_{n+1} : \\ (\underline{x} \in \underline{z} \iff (\underline{x} \in \underline{y} \wedge \alpha(\underline{x}, \underline{y}_1, \ldots, \underline{y}_k)))\end{gathered}$$

is true. For the choice $\underline{y}_i = B_i \in {}^*S_n$, $\underline{y} = B \in {}^*S_n$, we thus find that there is a set $Z := \underline{z} \in {}^*S_{n+1}$ which satisfies $Z \cap {}^*S_{n+1} = A \cap {}^*S_{n+1}$. Noting that ${}^*S_{n+1}$ is transitive by Lemma 3.14, we have $Z \subseteq {}^*S_{n+1}$ and $A \subseteq B \subseteq {}^*S_n \subseteq {}^*S_{n+1}$. Hence, $A = Z \in {}^*S_{n+1}$ is an internal set. □

Corollary 3.18. *An entity $A \in \widehat{{}^*S}$ is internal if and only if it can be written in the form*

$$A = \{\underline{x} \in {}^*B : \alpha(\underline{x})\}$$

where $B \in \widehat{S}$, and α is a transitively bounded internal predicate with $\underline{x}$ as its only free variable.

Proof. Since ${}^*B$ is internal by Proposition 3.16, the internal definition principle implies that each set A of the described form is internal. Conversely, if A is internal, then Proposition 3.16 implies that $A \subseteq {}^*B$ for some standard entity ${}^*B$. Hence, we have the required representation $A = \{\underline{x} \in {}^*B : \underline{x} \in A\}$. □

Theorem 3.19. *The nonstandard universe $\mathscr{I}$ satisfies:*

1. *If $A, B \in \mathscr{I}$ are entities, then $A \cup B, A \cap B, A \setminus B, A \times B \in \mathscr{I}$.*
2. *If $\mathscr{A}$ is an internal system of entities, then $\bigcup \mathscr{A}$ and $\bigcap \mathscr{A}$ are internal.*
3. *For binary internal relations φ, the sets* $\operatorname{dom}(\varphi)$ *and* $\operatorname{rng}(\varphi)$ *are internal. Also φ^{-1} is internal.*
4. *Images and preimages of internal sets under internal functions are internal.*

Proof. 1. By Proposition 3.16 and Lemma 3.14, we find some index n with $A, B \in {}^*S_n$, and so $A, B \subseteq {}^*S_n$. Hence, $A \cup B = \{\underline{x} \in {}^*S_n : \underline{x} \in A \vee \underline{x} \in B\}$ is internal by the internal definition principle. $A \cap B \in \mathscr{I}$ and $A \setminus B \in \mathscr{I}$ follows analogously. Since ${}^*S_n \times {}^*S_n = {}^*(S_n \times S_n)$ (since $*$ is a superstructure monomorphism) is internal, also

$$A \times B = \{\underline{z} \in {}^*S_n \times {}^*S_n \mid \exists \underline{x} \in A, \underline{y} \in B : \underline{z} = (\underline{x}, \underline{y})\}$$

(Proposition 3.6) is internal by the internal definition principle.
2. If $\mathscr{A}$ is an internal system of entities, then Proposition 3.16 implies that there is some n with $\mathscr{A} \in {}^*S_n$, and then $\mathscr{A} \subseteq {}^*S_n$. Since ${}^*S_n$ is transitive, we have

$$\bigcup \mathscr{A} = \{\underline{x} \in {}^*S_n \mid \exists \underline{y} \in \mathscr{A} : \underline{x} \in \underline{y}\}$$

which is internal by the internal definition principle. Analogously,

$$\bigcap \mathscr{A} = \{\underline{x} \in {}^*S_n \mid \forall \underline{y} \in \mathscr{A} : \underline{x} \in \underline{y}\}$$

is internal.
3. If φ is a binary internal relation, we find some n with $\varphi \in {}^*S_n$. Since ${}^*S_n$ is transitive, it follows that $(\underline{x}, \underline{y}) \in \varphi$ implies $\underline{x}, \underline{y} \in {}^*S_n$. Hence, $\operatorname{dom}(\varphi) = \{\underline{x} \in {}^*S_n \mid \exists \underline{y} \in {}^*S_n : (\underline{x}, \underline{y}) \in \varphi\}$, which implies by the internal definition principle that $\operatorname{dom}(\varphi)$ is internal. $\operatorname{rng}(\varphi) \in \mathscr{I}$ is proved analogously. If $\psi = \{(\underline{y}, \underline{x}) : (\underline{x}, \underline{y}) \in \varphi\}$, then $\psi = \{(\underline{y}, \underline{x}) \in {}^*S_n \times {}^*S_n : (\underline{x}, \underline{y}) \in \varphi\}$ is internal by the internal definition principle.
4. If $\varphi : A \to B$ is internal and $M \subseteq A$ is internal, then the image of M is $\{\underline{y} \in \operatorname{rng}(\varphi) \mid \exists \underline{x} \in M : (\underline{x}, \underline{y}) \in \Phi\}$ which is internal by the internal definition principle. An analogous proof shows that preimages of internal sets are internal. □

We have some analogue to Theorem 2.1: A union of a system of internal sets need not be internal; but it *is* internal if the system is finite or if the system is internal.

However, in contrast to Theorem 2.1, subsets of internal sets are not internal, in general (and so we have a similar restriction as for the union also for the intersection of internal sets).

Corollary 3.20 (Internal Definition Principle for Relations)**.** *An n-ary relation $\varphi \in \widehat{{}^*S}$ is internal if and only if it can be written in the form*

$$\varphi = \{(\underline{x}_1, \ldots, \underline{x}_n) \in B_1 \times \cdots \times B_n : \alpha(\underline{x}_1, \ldots, \underline{x}_n)\}$$

where $B_1, \ldots, B_n$ are internal entities, and α is a transitively bounded internal predicate with $\underline{x}_1, \ldots, \underline{x}_n$ as its only free variables.

Proof. We have

$$A = \{\underline{z} \in B_1 \times \cdots \times B_n \mid \exists \underline{x}_1 \in B_1, \ldots, \underline{x}_n \in B_n : \\ (\underline{z} = (\underline{x}_1, \ldots, \underline{x}_n) \wedge \alpha(\underline{x}_1, \ldots, \underline{x}_n))\}.$$

Since $B_1 \times \cdots \times B_n$ is internal by Theorem 3.19, the internal definition principle implies that A is internal. □

Exercise 6. If $x_1, \ldots, x_n$ are internal, prove that $\{x_1, \ldots, x_n\}$ and $(x_1, \ldots, x_n)$ are internal. Conclude that any external subset of an internal entity is infinite.

Exercise 7. Let $f : A \to B$ and $g : B \to C$ be internal functions. Prove that $g \circ f$ is an internal function.

Exercise 8. Prove the following:

1. Let $A_1, \ldots, A_n$ be pairwise disjoint internal entities, and $A := A_1 \cup \cdots \cup A_n$. Let f_i $(i = 1, \ldots, n)$ be internal functions with $\mathrm{dom}(f_i) \supseteq A_i$ and $\mathrm{rng}(f_i) \subseteq B$. Then there exists an internal function $f : A \to B$ satisfying $f(x) = f_i(x)$ for $x \in A_i$ $(i = 1, \ldots, n)$.
2. Let A, B be internal entities, $B \neq \varnothing$, $A_0 \subseteq A$, and let $f : A_0 \to B$ be an internal function. Then f may be extended to an internal function $F : A \to B$, i.e. $F(x) = f(x)$ for all $x \in A_0$.
3. The restriction of an internal function to an internal set is internal.

From the above observations one might guess that $\mathscr{I}$ is the same as $\widehat{{}^*S}$, because many "natural" operations appear to remain within the nonstandard universe $\mathscr{I}$. In fact, the earlier mentioned approach to nonstandard analysis by Nelson only "knows" internal sets: This approach is more or less an axiomatic description of set theory within $\mathscr{I}$, the so-called *internal set theory* (this is not quite precise, but gives a rather good idea of Nelson's approach).

However, the above impression is misleading: $\widehat{{}^*S}$ is actually much larger than $\mathscr{I}$. Although $\mathscr{I}$ has many properties analogous to Theorem 2.1, one important property is missing: The relation $A \subseteq B$ for an internal set B does not imply that A is internal. In particular, the powerset of an internal set is not internal, in general. For this reason one has to take extreme care if the symbols $\subseteq$ and $\mathcal{P}$ are involved.

For some purposes it is very convenient to calculate with $\widehat{{}^*S}$ in place of $\mathscr{I}$; this possibility gets lost in Nelson's nonstandard analysis. On the other hand, some other properties can be described more consistently by internal set theory. A comparison of the two approaches can be found in [DS88] (see also [CK90, LR94]). We will not discuss Nelson's nonstandard analysis any more.

Now that we know internal sets, we can "explicitly" calculate the $*$-value of $\mathcal{P}(A)$ and of the set B^A of all functions $f : A \to B$:

Theorem 3.21. *Let $* : \widehat{S} \to \widehat{{}^*S}$ be elementary. If $A \in \widehat{S}$ is an entity, then*

$$^*\mathcal{P}(A) = \{M \subseteq {}^*A : M \text{ is internal}\}.$$

If $A, B \in \widehat{S}$ are entities, then

$$^*(B^A) = \{f \mid f : {}^*A \to {}^*B, \text{ and } f \text{ is internal}\}.$$

Proof. Note that $C := \mathcal{P}(A)$ and $D := B^A$ belong to $\widehat{S}$. Hence, we find some index n with $C, D \subseteq S_n$ (recall that S_n is transitive), and so ${}^*C, {}^*D \subseteq {}^*S_n$ (Lemma 3.5). Using the shortcuts from Proposition 3.6, the sentences $\forall \underline{x} \in S_n : (\underline{x} \in C \iff \underline{x} \subseteq A)$ and $\forall \underline{x} \in S_n : (\underline{x} \in D \iff (\underline{x} : A \to B))$ are true. The transfer principle implies that $\forall \underline{x} \in {}^*S_n : (\underline{x} \in {}^*C \iff \underline{x} \subseteq {}^*A)$ and $\forall \underline{x} \in {}^*S_n : (\underline{x} \in {}^*D \iff (\underline{x} : {}^*A \to {}^*B))$ are true. In view of ${}^*C, {}^*D \subseteq {}^*S_n$, this implies the statement. □

Now we know all definitions and results which are needed to understand the appendix, where $*$-values of other important classes of sets are calculated. Occassionally, some of these values are needed in the later sections. However, reading the appendix *now* may appear boring and not too easy in the moment. Thus, although from a mathematical point of view the appendix should be read *now*, the reader is advised to read it at a later time, after having more experience with nonstandard analysis (e.g. when reference to a result in the appendix is made). It is a good idea to read the appendix parallel to §6.

3.4 Existence of External Sets

Let us now prove that there are a lot of external sets. Simultaneously, we show that all "nontrivial" elementary embeddings $*$ are actually nonstandard embeddings:

Theorem 3.22. *Let $* : \widehat{S} \to \widehat{{}^*S}$ be elementary. If $\widehat{S}$ contains an infinite entity, then $*$ is a nonstandard embedding if and only if ${}^\sigma B \neq {}^*B$ for some infinite countable entity $B \in \widehat{S}$. In this case, for any entity $A \in \widehat{S}$ the following holds:*

1. *If A is infinite, then ${}^\sigma A$ is external.*
2. *If A is infinite, then $\mathcal{P}({}^*A)$ is external, and*

$$ {}^\sigma\mathcal{P}(A) \subsetneq {}^*\mathcal{P}(A) \subsetneq \mathcal{P}({}^*A). $$

3. *If S is infinite, then ${}^*S \setminus {}^\sigma S$ is nonempty and contains only elements which are internal but not standard.*

Proof. If $\widehat{S}$ contains an infinite entity B_0 and $*$ is a nonstandard embedding, choose some infinite countable $B \subseteq B_0$. Then $B \in \widehat{S}$ by Theorem 2.1, and by definition ${}^\sigma B \neq {}^*B$.

Conversely, suppose that there is some infinite countable $B \in \widehat{S}$ such that ${}^\sigma B \neq {}^*B$. We shall show 1. from this assumption. This implies that $*$ is a nonstandard embedding, since we must have ${}^\sigma A \neq {}^*A$, because ${}^*A$ is internal by Proposition 3.16, but ${}^\sigma A$ is not.

1. Recall that ${}^\sigma B \subsetneq {}^*B$ by Corollary 3.11. We show first that $C := {}^*B \setminus {}^\sigma B$ is external:

Assume the contrary, that C is internal. Let $b_1, b_2, \ldots$ be an enumeration of the elements of B. Define a relation $\varphi \subseteq B \times B$ by $(b_n, b_k) \in \varphi \iff n \leq k$. Then φ defines a well-order on B, i.e. each nonempty subset of B has a smallest element:

$$ \forall \underline{x} \in \mathcal{P}(B) : (\underline{x} \neq \varnothing \implies \exists \underline{y} \in \underline{x} : \forall \underline{z} \in \underline{x} : (\underline{y}, \underline{z}) \in \varphi). $$

The transfer principle thus implies

$$ \forall \underline{x} \in {}^*\mathcal{P}(B) : (\underline{x} \neq {}^*\varnothing \implies \exists \underline{y} \in \underline{x} : \forall \underline{z} \in \underline{x} : (\underline{y}, \underline{z}) \in {}^*\varphi). $$

Since C is internal by assumption, Theorem 3.21 implies $C \in {}^*\mathcal{P}(B)$. Since $C \neq \varnothing = {}^*\varnothing$, the above statement for $\underline{x} = C$ thus implies that there is an element $y \in C$ such that $\underline{z} \in C$ implies $(\underline{y}, \underline{z}) \in {}^*\varphi$.

We shall follow this to a contradiction. We note first that an induction by n implies $({}^*b_n, y) \in {}^*\varphi$: Indeed, the transfer principle applied to the transitively bounded sentence

$$ \forall \underline{y} \in B : ((\underline{y} \neq b_1 \wedge \cdots \wedge \underline{y} \neq b_{n-1}) \implies (b_n, \underline{y}) \in \varphi) $$

(Proposition 3.6) shows that the sentence

$$ \forall \underline{y} \in {}^*B : ((\underline{y} \neq {}^*b_1 \wedge \cdots \wedge \underline{y} \neq {}^*b_{n-1}) \implies ({}^*b_n, \underline{y}) \in {}^*\varphi) $$

is true. Since $\underline{y} = y \in {}^*B \setminus {}^\sigma B$, we thus have $({}^*b_n, y) \in {}^*\varphi$ by induction assumption (the case $n = 1$ is analogous).

Now we consider the map $p : B \to B$ defined by $p(b_n) := b_{n-1}$ $(n \geq 1)$ and $p(b_1) := b_1$ and note that

$$\forall \underline{y} \in B : (\underline{y} \neq b_1 \implies (\underline{y}, p(\underline{y})) \notin \varphi).$$

Here, $(\underline{y}, p(\underline{y})) \notin \varphi$ is a shortcut for $\exists \underline{z} \in \operatorname{rng} p : (\underline{z} = p(\underline{y}) \wedge (\underline{y}, \underline{z}) \notin \varphi)$ (henceforth, we will no longer write down such shortcuts). The transfer principle now implies that $(y, {}^*p(y)) \notin {}^*\varphi$. We thus have a contradiction if we can prove that $z := {}^*p(y)$ does not belong to $C = {}^*B \setminus {}^\sigma B$. Since ${}^*p : {}^*B \to {}^*B$ (Theorem 3.13), we have $z \in {}^*B$, so it suffices to prove $z \notin {}^\sigma B$. But this would mean $z = {}^*b_n$ for some n. The transfer of the sentence

$$\forall \underline{y} \in B : (p(\underline{y}) = b_n \implies (\underline{y} = b_{n+1} \vee \underline{y} = b_1))$$

thus implies $y = {}^*b_{n+1}$ or $y = {}^*b_1$ which are both not possible, since $y \notin {}^\sigma B$. This contradiction shows that $C = {}^*B \setminus {}^\sigma B$ is indeed external.

Now it follows that also ${}^\sigma B$ is external, because otherwise $C = {}^*B \setminus {}^\sigma B$ is the difference of two internal sets (Proposition 3.16) and thus C were internal by Theorem 3.19.

Since A is infinite, there is a function $f : A \to B$ which is *onto* B. Then ${}^*f : {}^*A \to {}^*B$, and ${}^*f$ maps ${}^\sigma A$ onto ${}^\sigma B$: Indeed, for any $a \in A$ and any $b \in B$ for which the sentence $f(a) = b$ is true, the transfer principle implies ${}^*f({}^*a) \in {}^*b$. In particular ${}^*f({}^*a) \in {}^\sigma B$ for all ${}^*a \in {}^\sigma A$, and for each ${}^*b \in {}^\sigma B$, we find indeed some preimage ${}^*a \in {}^\sigma A$, since f is onto. We may now conclude that ${}^\sigma A$ is external, because otherwise the image of ${}^\sigma A$ under the internal map ${}^*f$ (Proposition 3.16) would be internal by Theorem 3.19; but this image is ${}^\sigma B$ and thus external. This contradiction shows that ${}^\sigma A$ is external.
2. The inclusion ${}^\sigma\mathcal{P}(A) \subsetneq {}^*\mathcal{P}(A)$ follows from Corollary 3.11. Now we note that $\mathcal{P}({}^*A)$ consists of all subsets of ${}^*A$, while ${}^*\mathcal{P}(A)$ consists of all *internal* such subsets. Hence, ${}^*\mathcal{P}(A) \subseteq \mathcal{P}({}^*A)$, and the inclusion is strict since by 1. there actually is a subset of ${}^*A$ which is not internal, namely ${}^\sigma A$.
3. From Corollary 3.11, we conclude that ${}^*S \setminus {}^\sigma S$ is nonempty. Let $a \in {}^*S$. Then a is an internal element and also an atom by Lemma 3.5. Using Lemma 3.5, we find that a is a standard element if and only if $a = {}^*b$ for some atom $b \in S$, i.e. if and only if $a \in {}^\sigma S$. □

The previous proof might appear rather artificial to the reader: What we did in fact was to identify $B = \mathbb{N}$ and then proved that all internal subsets of ${}^*\mathbb{N}$ have a smallest element, but that ${}^*\mathbb{N} \setminus {}^\sigma\mathbb{N}$ contains no smallest element. We will repeat this argument in §5.

To give the reader an impression about internal sets, let us already note the following theorem, although we have to postpone the proof:

Theorem 3.23. *Let* $* : \widehat{S} \to \widehat{{}^*S}$ *be a nonstandard embedding. Then an internal entity is either finite or has at least the cardinality of the continuum. In particular, there are no countable internal sets.*

Theorem 3.23 can be proved by purely model theoretic considerations [CK90]. However, we will give a real "nonstandard" proof for this result in Section 7.1.

Corollary 3.24. *If* $* : \widehat{S} \to \widehat{{}^*S}$ *is a nonstandard embedding, then each infinite internal entity has an external subset.*

Proof. Just choose some countable subset. □

The previous results already give a rather good impression of the nature of internal sets: They may not be "too discrete".

For example if $S = \mathbb{R}$, the subset ${}^\sigma\mathbb{R}$ of the internal set ${}^*\mathbb{R}$ is "too discrete" to be internal; also the complement ${}^*\mathbb{R} \setminus {}^\sigma\mathbb{R}$ is not internal (why?). In this connection it is useful to think of ${}^*\mathbb{R}$ as a "smooth continuum"; the "natural" subsets are always "thick" in the sense that with each point they have to contain a "continuous neighborhood of infinitesimals". These "natural" sets are the internal ones. Of course, it is also "natural" to speak of a single real number (and in fact, sets containing only a single number are also internal (why?)). However, it is not admissible to collect infinitely many single numbers into some set: Already for a countable collection, we get an external set by Theorem 3.23. However, if the collection is "good enough" (in some sense a "smooth continuum"), we have to deal with an internal set.

Exercise 9. All previous examples of external sets have been subsets of standard sets. Is it always true that an external set is a subset of some standard set or at least a subset of some internal set?

§4 Nonstandard Ultrapower Models

The aim of this section is to describe a nonstandard map "explicitly". There is only one inconstructive step in the proof, namely the choice of a so-called ultrafilter. Since there are many possible ultrafilters (recall that we assume the axiom of choice), there are many different ways to define nonstandard maps which usually have different additional properties. Nevertheless, the process of defining the model will be independent of the particular choice of the ultrafilter.

We should point out that the method we present here is not the only possible approach, but the model we obtain has particularly nice properties. Moreover, the fact that the approach is "almost constructive" has the advantage that in special cases one can better see what happens: Up to the ultrafilter one can "calculate" the nonstandard embedding $*$. In particular, the role of the internal sets will become evident in the course of the construction.

The plan of the proof is as follows: Let a superstructure $\widehat{S}$ and a language $\mathscr{L}$ be given whose constants are in a one-to-one relation with the elements of $\widehat{S}$ (as before, we denote this relation by I). We first determine an abstract model $\mathscr{S}$ of $\mathscr{L}$ which is in a certain sense nonstandard and which satisfies the transfer principle even for sentences which are not necessarily bounded; let I_0 denote the corresponding interpretation map. In a second step, we embed $\mathscr{S}$ into a superstructure $\widehat{{}^*S}$ by a map φ such that $I' = \varphi \circ I_0$ is the desired interpretation map. This map φ does not satisfy the transfer principle for all sentences but only for a certain subclass (containing the transitively bounded sentences). In the construction of φ the role of the internal sets will become evident.

To construct the abstract model $\mathscr{S}$, we need some facts about ultrafilters.

4.1 Ultrafilters

Let J be some set. Probably, the reader has already heard the notion that a property holds "almost everywhere" on J: By this, one means that the set of all point $j \in J$ with this property is "large" in a certain specified sense. For example, one may mean that the complement of this set is finite (if J is infinite); if J is a measure space, one can also mean that the complement of this set is a null set. (Recall Exercises 3 and 4).

If we want to introduce a general definition of the term "almost everywhere" which contains the two cases above, we should fix a family $\mathscr{F}$ of subsets of J and say that a property holds *almost everywhere* if the set of all points $j \in J$ with this property is an element of $\mathscr{F}$. It is natural to require that for any set $A \in \mathscr{F}$, $\mathscr{F}$ should also contain all sets which are larger than A. Moreover, we should require that if a property P_1 holds almost everywhere and a property P_2

also holds almost everywhere, then $P_1 \wedge P_2$ also holds almost everywhere. To avoid trivial cases, we also require that if a property holds nowhere, then it does not hold almost everywhere. These requirements may be formulated in terms of $\mathscr{F}$ as follows:

Definition 4.1. A set $\mathscr{F}$ of subsets of J is called a *filter* on J, if it has the following properties:

1. If $A \in \mathscr{F}$ and $A \subseteq B \subseteq J$, then $B \in \mathscr{F}$.
2. If $A, B \in \mathscr{F}$, then $A \cap B \in \mathscr{F}$.
3. $\varnothing \notin \mathscr{F}$.

We say that a property holds *almost everywhere* on J (with respect to $\mathscr{F}$), if the set of all $j \in J$ with this property belongs to $\mathscr{F}$; we briefly say that this property holds for *almost every* $j \in J$.

The examples mentioned above come from the following filters:

Example 4.2. Let J be infinite, and $\mathscr{F}$ be the system of all sets of the form $J \setminus J_0$ where $J_0 \subseteq J$ is finite. Then $\mathscr{F}$ is a filter.

Example 4.3. Let J be a measure space ($\operatorname{mes} J \neq 0$), and $\mathscr{F}$ be the system of all sets of the form $J \setminus N$ where N is a subset of some set of measure 0. Then $\mathscr{F}$ is a filter.

Recall that not any subset of a set of measure 0 must be measurable; but if we define a null set as an arbitrary subset of a set of measure 0, we get a filter in the sense of Definition 4.1. In this sense, we can say that the complements of sets of measure 0 generate the filter from Example 4.3:

Definition 4.4. A system $\mathscr{B}$ of sets has the *finite intersection property*, if for each finitely many sets of $\mathscr{B}$ the intersection is nonempty. If $\mathscr{B}$ is a system of subsets of J with the finite intersection property, then the *filter generated* by $\mathscr{B}$ is the system $\mathscr{F}$ of all subsets $B \subseteq J$ for which there exist finitely many $A_1, \ldots, A_n \in \mathscr{B}$ with $B \supseteq A_1 \cap \cdots \cap A_n$.

The proof of the following observation is straightforward and left to the reader.

Proposition 4.5. *Let $\mathscr{F}$ be the filter generated by $\mathscr{B}$. Then $\mathscr{F}$ is a filter in the sense of Definition* 4.1*. Moreover, $\mathscr{F}$ is the smallest filter which contains all sets from $\mathscr{B}$.* □

If we can say that a property holds *almost everywhere*, it also makes sense to say that a property holds *almost nowhere* if this property fails almost everywhere. Since a property is either true or false, we have that a property holds almost nowhere on J (with respect to $\mathscr{F}$), if the *complement* of the set of all points $j \in J$ with this property belongs to $\mathscr{F}$. It is a natural question whether there exists a filter $\mathscr{F}$ such that each property holds either almost everywhere or almost nowhere. The filters with this property are called *ultrafilters*:

Definition 4.6. A filter $\mathscr{U}$ on J is an *ultrafilter* if for any $A \subseteq J$ with $A \notin \mathscr{U}$, we have $J \setminus A \in \mathscr{U}$.

Note that $A \in \mathscr{U}$ implies $J \setminus A \notin \mathscr{U}$, since the intersection of these sets is empty and thus cannot be contained in the filter $\mathscr{U}$.

Proposition 4.7. *A filter $\mathscr{U}$ on J is an ultrafilter, if and only if it is not contained in a strictly larger filter on J.*

Proof. Assume that $\mathscr{U}$ is an ultrafilter. If $\mathscr{U}$ is contained in a strictly larger filter $\mathscr{F}$, we find some $A \in \mathscr{F}$ which does not belong to $\mathscr{U}$. By assumption, the set $B = J \setminus A$ belongs to $\mathscr{U}$ and thus to $\mathscr{F}$. Since $\mathscr{F}$ is a filter, we must have $A \cap B \in \mathscr{F}$. But this means $\varnothing \in \mathscr{F}$, a contradiction.

Conversely, suppose that $\mathscr{U}$ is not contained in a strictly larger filter. If $\mathscr{U}$ is not an ultrafilter, we find some $A \subseteq J$ such that $A \notin \mathscr{U}$ or $J \setminus A \notin \mathscr{U}$. Then the set $\mathscr{B} = \mathscr{U} \cup \{A\}$ has the finite intersection property: To see this, it suffices to prove that $A \cap B \neq \varnothing$ for any $B \in \mathscr{U}$, because $\mathscr{U}$ is a filter. But if $A \cap B = \varnothing$ for some $B \in \mathscr{U}$, we have $B \subseteq J \setminus A$, and so $J \setminus A \in \mathscr{U}$, contradicting our assumption on A. Hence, $\mathscr{B}$ has the finite intersection property and thus generates a filter which is strictly larger than $\mathscr{U}$. □

A trivial example of an ultrafilter is the following: Fix some element $j_0 \in J$, and let $\mathscr{U}$ be the system of all subsets of J which contain the element j_0. Then $\mathscr{U}$ is a filter which contains $A \subseteq J$ if and only if $j_0 \in A$; if $A \notin \mathscr{U}$, we have $j_0 \in J \setminus A$, and thus $J \setminus A \in \mathscr{U}$. We want to exclude this example:

Definition 4.8. A filter $\mathscr{F}$ is called *free*, if $\bigcap \mathscr{F} = \varnothing$.

Exercise 10. Prove that an ultrafilter $\mathscr{U}$ is free if and only if it does not have the form described above, i.e. if and only if we do not have

$$\mathscr{U} = \{U \subseteq J : j_0 \in U\}$$

for some $j_0 \in J$.

It is not obvious whether there exist free ultrafilters. Consider, for example, the filter $\mathscr{F}$ of all subsets of an infinite set J with finite complements. Then $\mathscr{F}$ is free. Thus, any ultrafilter containing $\mathscr{F}$ is free. However, even for $J = \mathbb{N}$ it is not possible to describe an ultrafilter containing $\mathscr{F}$. Nevertheless, the axiom of choice implies that such an ultrafilter exists:

Theorem 4.9. *Each filter $\mathscr{F}$ is contained in some ultrafilter.*

The proof is a straightforward application of e.g. Zorn's Lemma (or, alternatively, of Hausdorff's maximality principle or of the well-ordering theorem) and is left to the reader. We remark that, in contrast to Zorn's lemma, the statement of Theorem 4.9 is not equivalent to the axiom of choice [Pin73]. In other words, if we take Theorem 4.9 as an axiom, this axiom is strictly less restrictive than the axiom of choice (usually, this axiom is called the *maximal ideal theorem* because it is equivalent to the fact that on any Boolean algebra there exists a maximal ideal [Sik64, Lux69c]).

Now we can prove that over each infinite set J there exists a free ultrafilter (using the axiom of choice):

Exercise 11. Prove that for an ultrafilter $\mathscr{U}$ over an infinite set J the following statements are equivalent:

1. $\mathscr{U}$ is free.
2. $\mathscr{U}$ contains the filter of Example 4.2 (observe that by Theorem 4.9 such ultrafilters exist for each infinite set J).

Actually, we do not really need free ultrafilters but δ-incomplete ultrafilters:

Definition 4.10. A filter $\mathscr{F}$ is called *δ-incomplete* if there is a countable subset $\mathscr{F}_0 \subseteq \mathscr{F}$ with $\bigcap \mathscr{F}_0 \notin \mathscr{F}$.

Note that for a filter $\mathscr{F}$, the intersection of *finite* subsets of $\mathscr{F}$ belongs to $\mathscr{F}$, so that $\mathscr{F}_0$ must actually be infinite.

Exercise 11 (and Theorem 4.9) imply that at least over countable sets J there exist δ-incomplete ultrafilters:

Proposition 4.11. *If $\mathscr{F}$ is a free filter on a countable set J, then $\mathscr{F}$ is δ-incomplete.*

Proof. Let $J = \{j_1, j_2, \ldots\}$. Since $\bigcap \mathscr{F} = \varnothing$, we find for each n some set $F_n \in \mathscr{F}$ with $j_n \notin F_n$. Hence, $\bigcap_n F_n = \varnothing \notin \mathscr{F}$. □

The importance of δ-incomplete filters lies in the following fact:

Proposition 4.12. *If a filter $\mathscr{F}$ on J is δ-incomplete then there is a partition of J into countably many pairwise disjoint sets $J_0, J_1, \ldots$ such that none of these sets belongs to $\mathscr{F}$.*

Conversely, if $\mathscr{U}$ is an ultrafilter for which such a partition exists, then $\mathscr{U}$ is δ-incomplete and free.

Proof. If $\mathscr{F}$ is δ-incomplete, there exist countably many $A_1, A_2, \ldots \in \mathscr{F}$ such that $J_0 := \bigcap_n A_n \notin \mathscr{F}$. Then we define by induction $J_n := J \setminus (J_0 \cup \cdots \cup J_{n-1} \cup A_n)$ $(n = 1, 2, \ldots)$. By construction, we have $J_n \cap J_k = \varnothing$ for $k < n$, and $J \setminus J_n \subseteq A_n$. The latter implies $J \setminus \bigcup_{n \geq 1} J_n = \bigcap_{n \geq 1} (J \setminus J_n) \subseteq J_0$, and so $J_0, J_1, \ldots$ is indeed a partition of J. Since the set $J_0 \cup \cdots \cup J_{n-1} \cup A_n$ belongs to $\mathscr{F}$ (because $\mathscr{F}$ is a filter which contains A_n), the complement J_n does not belong to $\mathscr{F}$.

Conversely, if $J_0, J_1, \ldots$ is a partition of J into (at most) countably many pairwise disjoint sets with $J_n \notin \mathscr{U}$, then $J \setminus J_n \in \mathscr{U}$, because $\mathscr{U}$ is an ultrafilter.

Hence, $\mathscr{U}_0 = \{J \backslash J_n : n = 0, 1, \ldots\}$ is a countable subset of $\mathscr{U}$ with $\bigcap \mathscr{U}_0 = \varnothing \notin \mathscr{U}$. Hence, $\mathscr{U}$ is δ-incomplete. Since $\bigcap \mathscr{U} \subseteq \bigcap \mathscr{U}_0 = \varnothing$, $\mathscr{U}$ is free. □

Corollary 4.13. *Any δ-incomplete ultrafilter is free.* □

On any uncountable set J, there is a free filter which fails to be δ-incomplete, namely the filter of all sets with countable complements. However, it is not clear whether any free *ultrafilter* must be δ-incomplete (i.e. whether the converse of Corollary 4.13 holds). This question is known as "Ulam's measure problem". It is consistent with the axioms of ZF set theory to assume that any free filter is δ-incomplete, but it is not provable that it is consistent to assume the converse: If there actually should exist a free ultrafilter on J which fails to be δ-incomplete, then J must have an extremely large cardinality, namely the cardinality of at least a so-called measurable cardinal which in turn has the cardinality of at least a so-called inaccessible cardinal. It is not provable that inaccessible cardinals exist. Moreover, it is not even provable that it is consistent (with the axioms of ZF set theory) to assume that such cardinals exist (see [Jec97]).

4.2 Ultrapowers

Let $\widehat{S}$ be a superstructure, and $\mathscr{L}$ a language with a surjective interpretation map I onto $\widehat{S}$.

Let J be an infinite set, and $\mathscr{U}$ an ultrafilter on J. We define now an abstract model $\mathscr{S}$ of the language $\mathscr{L}$, the so-called *ultrapower of* $\widehat{S}$ modulo $\mathscr{U}$.

We start with the set $\widehat{S}^J$ of all functions $f : J \to \widehat{S}$. On this set, we introduce a natural equivalence relation: We call two such functions f, g equivalent, if $f(j) = g(j)$ for almost all $j \in J$, i.e. if $J_{f,g} = \{j \in J : f(j) = g(j)\} \in \mathscr{U}$. This is an equivalence relation: This is trivial except for the transitivity. But if f is equivalent to g and g is equivalent to h, then $J_{f,g}, J_{g,h} \in \mathscr{U}$ which implies $J_{f,g} \cap J_{g,h} \in \mathscr{U}$. Since $J_{f,h} \supseteq J_{f,g} \cap J_{g,h}$, this implies $J_{f,h} \in \mathscr{U}$.

Let $\mathscr{S}$ be the set of equivalence classes of $\widehat{S}^J$ with respect to this equivalence relation. The interpretation map $I_0 : \widehat{S} \to \mathscr{S}$ is simply the map which associates to each constant whose interpretation under I is $c \in \widehat{S}$ the class which contains the constant function f_c, defined by $f_c(j) = c$ $(j \in J)$. To speak of an abstract model we have to equip $\mathscr{S}$ with two relations $\in_{\mathscr{U}}$ and $=_{\mathscr{U}}$. These are defined as follows:

Denote the class containing a function $f : J \to \widehat{S}$ by $[f]$. Then we define

$$[f] \in_{\mathscr{U}} [g] \iff f(j) \in g(j) \text{ for almost all } j \in J,$$

and

$$[f] =_{\mathscr{U}} [g] \iff f(j) = g(j) \text{ for almost all } j \in J.$$

We have to prove that this definition is independent of the particular choice of the representing elements f, g: Concerning $=_{\mathscr{U}}$, this is trivial, since $=_{\mathscr{U}}$ is just the usual equality relation of equivalence classes. Concerning $\in_{\mathscr{U}}$, assume that $[f_1] = [f_2]$ and $[g_1] = [g_2]$ and $f_1(j) \in g_1(j)$ for almost all j. We have to prove that $f_2(j) \in g_2(j)$ for almost all j. By assumption, the sets J_{f_1,f_2}, J_{g_1,g_2}, and $\{j \in J : f_1(j) \in g_1(j)\}$ belong to $\mathscr{U}$, and so also their intersection. This intersection is contained in the set $\{j \in J : f_2(j) \in g_2(j)\}$ which thus also belongs to $\mathscr{U}$. Hence, $f_2(j) \in g_2(j)$ for almost all j, as claimed.

So far, we have only used the fact that $\mathscr{U}$ is a filter (if $\mathscr{U}$ is not necessarily an ultrafilter, one calls $\mathscr{S}$ also the *reduced power* of $\widehat{S}$). But the fact that $\mathscr{U}$ is an ultrafilter is needed for the following important theorem:

Theorem 4.14 (Łoś and Luxemburg). *A sentence in $\mathscr{L}$ is true under the interpretation map I if and only if it is true under the interpretation map I_0.*

Proof. Let α be a formula in $\mathscr{L}$ (not necessarily a sentence!). Let $\underline{x}_1, \ldots, \underline{x}_n$ be all free variables of α ($n = 0$ is not excluded). For $f_i \in \widehat{S}^J$, let ${}^{I_0}\alpha([f_1], \ldots, [f_n])$ denote the formula where all free occurrences of $\underline{x}_i$ are replaced by $[f_i]$ ($i = 1, \ldots, n$), all constants are replaced by their image under the interpretation map I_0, and the symbols $\in$ and $=$ are replaced by $\in_{\mathscr{U}}$ and $=_{\mathscr{U}}$, respectively. Similarly, let ${}^{I}\alpha(f_1(j), \ldots, f_n(j))$ denote the formula where all free occurrences of $\underline{x}_i$ are replaced by $f_i(j)$, and all constants are replaced by their image under the canonical interpretation map I. We will show that

$$ {}^{I}\alpha(f_1(j), \ldots, f_n(j)) \text{ is true for almost all } j \iff {}^{I_0}\alpha([f_1], \ldots, [f_n]) \text{ is true.} \tag{4.1}$$

Then the statement follows from the special case $n = 0$: If α is a sentence, then α is true under the canonical interpretation map I if and only if ${}^{I}\alpha$ holds (for all j, since it is independent of j). By (4.1), this is equivalent to the fact that ${}^{I_0}\alpha$ is true which means that α is true under the interpretation map I_0.

Let us now prove (4.1). By the usual logical transformations, we may assume that the only logical connectives used in α are $\neg$ and $\wedge$. Moreover, replacing the formula $\forall \underline{x} : \beta$ by the equivalent formula $\neg \exists \underline{x} : \neg\beta$, we may assume that $\exists$ is the only quantifier used in α. We now prove the statement by induction over the structure of the sentence α (i.e. on the number of the symbols $\neg$, $\wedge$, and $\exists$). For the induction assumption, we only have to consider the elementary formulas $x = y$ and $x \in y$ where x and y are either free variables or constants. In this case, (4.1) follows immediately from the definition of the relations $=_{\mathscr{U}}$ and $\in_{\mathscr{U}}$. For the induction step, we have to consider three cases:

1. α has the form $\exists \underline{x} : \beta(\underline{x}, \underline{x}_1, \ldots, \underline{x}_n)$: (Note that even if α is a sentence, $\beta(\underline{x})$ might be a formula with a free variable: For this reason, even though we are

actually only interested in (4.1) for sentences ($n = 0$), we have to consider more general formulas with free variables for our induction proof).

Assume that ${}^{I}\alpha(f_1(j), \dots, f_n(j))$ is true for almost all j. Then we find for almost all j some $f_0(j)$ such that ${}^{I}\beta(f_0(j), f_1(j), \dots, f_n(j))$ is true. We consider f_0 is a function (axiom of choice!). By induction hypothesis, this implies that ${}^{I'}\beta([f_0], [f_1], \dots, [f_n])$ is true, and so ${}^{I'}\alpha([f_1], \dots, [f_n])$ is true.

Conversely, if ${}^{I}\alpha([f_1], \dots, [f_n])$ is true, then there is some $[f_0] \in \mathscr{S}$ such that ${}^{I'}\beta([f_0], [f_1], \dots, [f_n])$ is true. By induction assumption, this implies that ${}^{I}\beta(f_0(j), f_1(j), \dots, f_n(j))$ is true for almost all j. For all these j, we thus have that ${}^{I}\alpha(f_1(j), \dots, f_n(j))$ is true.

2. α has the form $\neg\beta(\underline{x}_1, \dots, \underline{x}_n)$:

If ${}^{I}\alpha(f_1(j), \dots, f_n(j))$ is true for almost all j, then ${}^{I}\beta(f_1(j), \dots, f_n(j))$ is false for almost all j. In particular, we do not have for almost all j that ${}^{I}\beta(f_1(j), \dots, f_n(j))$ is true. By induction assumption, this means that ${}^{I_0}\beta([f_1], \dots, [f_n])$ is not true, i.e. ${}^{I_0}\alpha([f_1], \dots, [f_n])$ is true.

Conversely, assume that ${}^{I_0}\alpha([f_1], \dots, [f_n])$ is true. Then ${}^{I_0}\beta([f_1], \dots, [f_n])$ is false, and by induction assumption it is not the case that, for almost all j, ${}^{I}\beta(f_1(j), \dots, f_n(j))$ is true. Since $\mathscr{U}$ is an ultrafilter, we may conclude that ${}^{I}\beta(f_1(j), \dots, f_n(j))$ is true for almost no j, i.e. ${}^{I}\beta(f_1(j), \dots, f_n(j))$ is false for almost all j. Hence, ${}^{I}\alpha(f_1(j), \dots, f_n(j))$ is true for almost all j.

3. α has the form $\beta_1 \wedge \beta_2$:

${}^{I}\alpha(f_1(j), \dots, f_n(j))$ is true for almost all j if and only if ${}^{I}\beta_i(f_1(j), \dots, f_n(j))$ ($i = 1, 2$) are both true for almost all j (because $A, B \in \mathscr{U}$ implies $A \cap B \in \mathscr{U}$). This is the case if and only if ${}^{I_0}\beta_i([f_1], \dots, [f_n])$ ($i = 1, 2$) are both true, i.e. if and only if ${}^{I_0}\alpha([f_1], \dots, [f_n])$ is true. $\square$

We note that we used the axiom of choice in the previous proof to find the function f_0; however, if e.g. J is countable, only a countable form of the axiom of choice is needed.

We call the interpretation map I_0 *nonstandard*, if for any constant A whose interpretation under I is an infinite entity ${}^{I}A$, the sets $A_* := \{c \in \mathscr{S} : c \in_{\mathscr{U}} {}^{I_0}A\}$ and $A_\sigma := \{{}^{I_0}a : a \in A \text{ is true}\}$ differ.

Theorem 4.15 (Luxemburg). *If the ultrafilter $\mathscr{U}$ is δ-incomplete, then the above defined interpretation map I_0 is nonstandard. More precisely, we have for any constant A with infinite ${}^{I}A$ that $A_* \supsetneq A_\sigma$.*

Conversely, if there is a constant A with countable infinite ${}^{I}A$ such that $A_* \neq A_\sigma$, then $\mathscr{U}$ is δ-incomplete.

Proof. Note that A_* consists of the equivalence classes of the functions $f : J \to \widehat{S}$ such that $f(j) \in {}^{I}A$ for almost all j, and A_σ consists of the equivalence classes of the constant functions $f : J \to \widehat{S}$ with values in ${}^{I}A$. Hence, $A_\sigma \subseteq A_*$, and we have $A_* \neq A_\sigma$ if and only if there is a function $f : J \to {}^I A$ which does not belong to the equivalence class of a constant function.

If $\mathscr{U}$ is δ-incomplete, we find a partition $J_1, J_2, \ldots$ of J such that no J_n belongs to $\mathscr{U}$. If A is a constant with infinite ${}^I A$, we find a sequence of pairwise distinct elements $a_1, a_2, \ldots \in {}^I A$. Putting $f(j) := a_n$ for $j \in J_n$, the function f is not almost everywhere equal to a constant function, since $J_n \notin \mathscr{U}$. Hence, $[f] \in A_* \setminus A_\sigma$, and the interpretation map I_0 is nonstandard.

Conversely, assume there is a countable infinite ${}^I A$ and a function $f : J \to {}^I A$ which does not belong to the equivalence class of a constant function. Let a_n be an enumeration of the elements of the image, and J_n be the preimage of a_n. Then J_n is an (at most) countable partition of J into pairwise disjoint sets which do not belong to $\mathscr{U}$: If $J_n \in \mathscr{U}$, then $f(j) = a_n$ for almost all j, a contradiction. Proposition 4.12 thus implies that $\mathscr{U}$ is δ-incomplete. □

4.3 Embedding in a Superstructure

So far, we were only able to build an *abstract* nonstandard model, i.e. the relations $\in$ and $=$ are not interpreted in the usual set-theoretic sense but instead by the relations $\in_{\mathscr{U}}$ and $=_{\mathscr{U}}$. But to have a nonstandard embedding in the sense of Definition 3.2, we want the interpretation of $\in$ and $=$ in the set-theoretic sense, and in fact we want an interpretation map I' with values in a superstructure $\widehat{{}^*S}$ (and not just in an abstract set $\mathscr{S}$ of equivalence classes).

As long as we restrict our attention to transitively bounded sentences, we may embed the abstract model $\mathscr{S}$ defined in Section §4.2 into a superstructure. More precisely, we are going to define a set ${}^*S$ and a map φ of a subset $\mathscr{S}_{\mathscr{I}} \subseteq \mathscr{S}$ into $\widehat{{}^*S}$ such that any transitively bounded sentence which is true under the interpretation map I is also true under the interpretation map $I' = \varphi \circ I_0$. Moreover, $\mathscr{S}_{\mathscr{I}}$ is sufficiently large so that $\widehat{{}^*S}$ will be a nonstandard model if $\mathscr{S}$ was a nonstandard model.

Let S_n be the levels of the superstructure $\widehat{S}$ as defined in Section 2.1. We consider the sets

$$\mathscr{I}_n := \{[f] \mid f : J \to S_n\}$$

and note that $S_0 \subseteq S_1 \subseteq \cdots$ implies $\mathscr{I}_0 \subseteq \mathscr{I}_1 \subseteq \cdots$; put $\mathscr{I}' := \bigcup_n \mathscr{I}_n$.

Lemma 4.16. *We have*

$$\mathscr{I}_n = \{[f] \in \mathscr{S} : \textit{There is some } [g] \in \mathscr{I}_{n+1} \textit{ with } [f] \in_{\mathscr{U}} [g]\}.$$

Proof. If $[f] \in \mathscr{I}_n$ is given, consider the constant function $g : J \to S_{n+1}$, defined by $g(j) := S_n$. Then $[g] \in \mathscr{I}_{n+1}$, and $[f] \in_{\mathscr{U}} [g]$.

Conversely, if $[g] \in \mathscr{I}_{n+1}$, and $[f] \in_{\mathscr{U}} [g]$, we have $f(j) \in g(j) \in S_{n+1}$ for almost all j. Since $S_{n+1} = S_0 \cup \mathcal{P}(S_n)$, this implies $f(j) \in S_n$ for almost all j, and so we may assume $f : J \to S_n$. Hence, $[f] \in \mathscr{I}_n$. □

We now let ${}^*S := \mathscr{I}_0$ be the set of atoms ${}^*S$ of our new superstructure $\widehat{{}^*S}$. Let T_n denote the level sets of that superstructure, i.e. in particular $T_0 = {}^*S$.

The function φ that we are looking for is an injection $\varphi : \mathscr{I}' \to \widehat{{}^*S}$ such that

$$\varphi([g]) = \begin{cases} [g] & \text{if } [g] \in \mathscr{I}_0 = {}^*S, \\ \{\varphi([f]) : [f] \in_{\mathscr{U}} [g]\} & \text{if } [g] \notin \mathscr{I}_0. \end{cases} \tag{4.2}$$

It turns out that the range of φ consists precisely of the internal sets. To prove that the function we are going to construct is actually injective, we need the following lemma:

Lemma 4.17. *If elements $[f], [g] \in \mathscr{S} \setminus \mathscr{I}_0$ satisfy*

$$\{[h] \in \mathscr{S} : [h] \in_{\mathscr{U}} [f]\} = \{[h] \in \mathscr{S} : [h] \in_{\mathscr{U}} [g]\}, \tag{4.3}$$

then $[f] = [g]$.

Proof. Let $k : J \to \widehat{S}$ be the constant function $k(j) = S$. The relation $[f] \notin \mathscr{I}_0$ means that $f(j) \in S_0 = k(j)$ does not hold almost everywhere, i.e. $[f] \in_{\mathscr{U}} [k]$ is not true. Hence, we have $[f] \notin_{\mathscr{U}} [k]$. Analogously, $[g] \notin_{\mathscr{U}} [k]$.

Note now that the sentence (in the language $\mathscr{L}$)

$$\forall \underline{x}, \underline{y} \notin S : ((\forall \underline{z} : (\underline{z} \in \underline{x} \iff \underline{z} \in \underline{y})) \implies \underline{x} = \underline{y})$$

is true under the canonical interpretation map I. By the theorem of Łoś/Luxemburg (Theorem 4.14), this sentence must also be true in the abstract model $\mathscr{S}$. (Recall that it is not required in Theorem 4.14 that the sentence be transitively bounded). In the model $\mathscr{S}$ the above sentence reads

$$\forall \underline{x}, \underline{y} \notin_{\mathscr{U}} [k] : ((\forall \underline{z} : (\underline{z} \in_{\mathscr{U}} \underline{x} \iff \underline{z} \in_{\mathscr{U}} \underline{y})) \implies \underline{x} = \underline{y}).$$

Putting $\underline{x} = [f]$ and $\underline{y} = [g]$ in this sentence, the statement follows. □

Exercise 12. Give a straightforward proof of Lemma 4.17, i.e. without applying Theorem 4.14.

Now we are going to define by induction on n sets ${}^*S_n$ and functions φ_n such that the following holds. (The idea is that φ_n is the restriction $\varphi|_{\mathscr{I}_n}$ where φ satisfies (4.2); however, we still have to prove that such a function φ actually exists).

1. $\varphi_n : \mathscr{I}_n \to {}^*S_n$ is bijective.

2. ${}^*S_n \subseteq T_n$.
3. We have

$$\varphi_n([g]) = \begin{cases} [g] & \text{if } [g] \in \mathscr{I}_0 = {}^*S, \\ \{\varphi_n([f]) : [f] \in_{\mathscr{U}} [g]\} & \text{if } [g] \in \mathscr{I}_n \setminus \mathscr{I}_0. \end{cases} \tag{4.4}$$

4. If $n \geq 1$, we have ${}^*S_{n-1} \subseteq {}^*S_n$, and moreover, φ_{n-1} is the restriction of φ_n to the set $\mathscr{I}_{n-1}$.

Then we may define the desired function φ by putting $\varphi([g]) := \varphi_n([g])$ for $[g] \in \mathscr{I}_n$. Note that φ must be injective, since each φ_n is injective and $\mathscr{I}_0 \subseteq \mathscr{I}_1 \subseteq \cdots$.

We stress that the inclusion ${}^*S_n \subseteq T_n$ is in general strict (which is a deeper reason for the existence of external sets).

For $n = 0$, we observe that $\mathscr{I}_0 = {}^*S = T_0$, and so we may put ${}^*S_0 := {}^*S$, and $\varphi_0([g]) := [g]$ for $[g] \in \mathscr{I}_0$.

Assume now by induction hypothesis that the set ${}^*S_n$ and the map $\varphi_n : \mathscr{I}_n \to {}^*S_n$ with the additional properties have already been defined. Then we let ${}^*S_{n+1}$ consist of all elements of ${}^*S$ and furthermore of all *elements* $A \in \mathcal{P}({}^*S_n)$ with the following property: There is some function $[g] \in \mathscr{I}_{n+1} \setminus \mathscr{I}_0$ such that $A = \{x \in {}^*S_n : \varphi_n^{-1}(x) \in_{\mathscr{U}} [g]\}$. In this case, we put $\varphi_{n+1}([g]) := A$. For $[g] \in \mathscr{I}_0 = {}^*S$, we put $\varphi_{n+1}([g]) := [g]$.

By construction, ${}^*S_{n+1} \subseteq {}^*S \cup \mathcal{P}({}^*S_n)$, which in view of the induction assumption ${}^*S_n \subseteq T_n$ implies ${}^*S_{n+1} \subseteq T_0 \cup \mathcal{P}(T_n) = T_{n+1}$. Since by induction assumption (4.4), ${}^*S_{n-1} \subseteq {}^*S_n$, and since any $[g] \in \mathscr{I}_n$ belongs to $\mathscr{I}_{n+1}$, our construction implies that φ_n is indeed a restriction of φ_{n+1} and thus that (4.4) holds for $n+1$.

It is also clear from the construction that φ_{n+1} is onto. It remains to prove that φ_{n+1} is one-to-one. Assume that $\varphi_{n+1}([f]) = \varphi_{n+1}([g])$. We prove that this implies $[f] = [g]$ by distinguishing three cases: If $[f] \in \mathscr{I}_0$, then $\varphi_{n+1}([f]) = [f] \in \mathscr{I}_0$; hence $\varphi_{n+1}([g]) = [f]$ is an equivalence class (and not a subset of T_n), which by construction of φ_{n+1} is only possible if $[g] \in \mathscr{I}_0$ in which case we must have $[f] = \varphi_{n+1}([g]) = [g]$. In the case $[g] \in \mathscr{I}_0$, it follows analogously that $[f] = [g]$. In the remaining case, we have $[f], [g] \notin \mathscr{I}_0$. Then the construction of φ_{n+1} implies

$$\{x \in {}^*S_n : \varphi_n^{-1}(x) \in_{\mathscr{U}} [f]\} = \{x \in {}^*S_n : \varphi_n^{-1}(x) \in_{\mathscr{U}} [g]\}.$$

Since φ_n is a bijection onto ${}^*S_n$, we thus find

$$\{[h] \in \mathscr{I}_n : [h] \in_{\mathscr{U}} [f]\} = \{[h] \in \mathscr{I}_n : [h] \in_{\mathscr{U}} [g]\}.$$

Lemma 4.16 now implies that (4.2) holds which by Lemma 4.17 implies that $[f] = [g]$, as claimed.

The function φ can now be defined, and the range of φ is the set $\mathscr{I} := \bigcup_n {}^*S_n$.

Roughly speaking, it is now clear that φ preserves the truth of transitively bounded sentences which deal with internal sets: It "preserves" the relations $\in$ and $=$ (for $\in$ observe (4.2), and for $=$ use Lemma 4.17). Moreover, this mapping is *onto*, and thus only provides a "renaming" of the constants. A more rigorous proof reads as follows:

Theorem 4.18. *Let α be a transitively bounded formula in the language whose constants are taken from $\mathscr{I}'$, and let $\underline{x}_1, \dots, \underline{x}_n$ denote the free variables of α ($n = 0$ is not excluded). For $x_i \in \mathscr{I}'$, let α_0 denote the formula where all free occurrences of $\underline{x}_i$ are replaced by x_i ($i = 1, \dots, n$). Then α_0 is true under the interpretation map φ if and only if it is true interpreted by the inclusion i into $\mathscr{S}$.*

In particular, a transitively bounded sentence with constants taken from $\mathscr{I}'$ is true under the interpretation map φ if and only if it is true under the inclusion i into $\mathscr{S}$.

Proof. Similarly as in the proof of Theorem 4.14, we may assume that the only logical connectives used in α are $\neg$ and $\wedge$ and that the only quantifier used is $\exists$ (in a transitively bounded form). The proof is by induction on the structure of the formula α. For induction assumption, we have to consider the elementary formulas $x = y$ and $x \in y$ where x and y are either free variables or constants. In general, we have to distinguish the following cases:

1. α has the form $x = y$: If α_0 is true under the interpretation map φ, then it is also true under the interpretation map i, since φ is one-to-one. The converse is trivial.
2. α has the form $x \in y$: Then the statement follows immediately from (4.2).
3. α has the form $\neg\beta$ or $\beta_1 \wedge \beta_2$: These cases are trivial, since only the constants are exchanged.
4. α has the form $\exists \underline{x} \in y : \beta(\underline{x})$ where y is either a free variable or a constant. Then α_0 has the form $\exists \underline{x} \in y : \beta_0(\underline{x})$ where β_0 is derived from β by replacing the free occurrences of $\underline{x}_1, \dots, \underline{x}_n$ by $x_1, \dots, x_n$, respectively.

If ${}^{\varphi}\alpha_0$ is true, then there is some x in the set ${}^{\varphi}y$ such that ${}^{\varphi}\beta_0(x)$ holds. By 2., the element $c := \varphi^{-1}(x)$ then satisfies $c \in_{\mathscr{U}} y$, and by induction assumption ${}^{i}\beta_0(c)$ holds. Hence, ${}^{i}\alpha_0$ is true.

Conversely, if ${}^{i}\alpha_0$ is true, we find some $x \in \mathscr{I}'$ such that $x \in_{\mathscr{U}} y$ and ${}^{i}\beta_0(x)$ holds. By 2., we then have $c := {}^{\varphi}x \in {}^{\varphi}y$, and by induction assumption ${}^{\varphi}\beta_0(c)$ is true. Hence, ${}^{\varphi}\alpha_0$ is true. □

Theorem 4.18 is the reason why we can prove the transfer principle only for *transitively bounded* sentences. If one needs the transfer principle for a particular class of sentences which are not transitively bounded, one "just" has to check whether Theorem 4.18 can be generalized to this class.

Proposition 4.19. *Let $I' := \varphi \circ I_0$. Then $* := I' \circ I^{-1} : \widehat{S} \to \widehat{{}^*S}$ has the following property: We have $x \in {}^*A$ for some entity $A \in \widehat{S}$ if and only if $x = \varphi([f])$ for some function $f : J \to A$. Moreover, S_n is actually mapped into the set ${}^*S_n$ as defined above.*

Proof. Let some entity $A \in \widehat{S}$ be given, and consider the constant function $f_A : J \to \widehat{S}$, defined by $f_A(j) := A$. Note that $I_0 \circ I^{-1}$ maps A into $[f_A]$. If $f : J \to A$, then $[f] \in_{\mathscr{U}} [f_A]$, and so Theorem 4.18 (or also (4.2)) implies $\varphi([f]) \in \varphi([f_A]) = {}^*A$. Conversely, if $x \in {}^*A$, then x belongs to the range of φ, i.e. $x = \varphi([f])$ for some $f : J \to \widehat{S}$. Since $\varphi([f]) \in {}^*A = \varphi([f_A])$, Theorem 4.18 implies $[f] \in_{\mathscr{U}} [f_A]$, i.e. $f(j) \in A$ for almost all j. By choosing a different representative if necessary, we may assume that $f(j) \in A$ even for all j, i.e. $f : J \to A$.

For the second statement, apply what we just proved for $A := S_n$: We have

$$ {}^*A = \{\varphi([f]) \mid f : J \to S_n\} = \{\varphi(x) : x \in \mathscr{I}_n\}. $$

But by construction, we have that $\varphi_n : \mathscr{I}_n \to {}^*S_n$ is onto where φ_n is the restriction of φ to $\mathscr{I}_n$. Hence, ${}^*A = \{y : y \in {}^*S_n\} = {}^*S_n$, as claimed. □

Let us collect the main result of §4:

Theorem 4.20. *Let $I' := \varphi \circ I_0$. Then $* := I' \circ I^{-1} : \widehat{S} \to \widehat{{}^*S}$ is an elementary embedding which maps S_n into ${}^*S_n$.*

If the ultrafilter $\mathscr{U}$ is δ-incomplete, then $*$ is a nonstandard embedding. Conversely, if $*$ is a nonstandard embedding and S is infinite, then $\mathscr{U}$ is δ-incomplete.

Proof. By Proposition 4.19, S_0 is mapped into ${}^*S_0 = {}^*S = T_0$, as required in Definition 3.1.

If α is a transitively bounded sentence which is true under the interpretation map I, then ${}^{I_0}\alpha$ is true by Theorem 4.14. Interpreting this sentence by φ, we thus get a true sentence by Theorem 4.18. But this interpreted sentence is just ${}^{I'}\alpha$.

Let the ultrafilter $\mathscr{U}$ be δ-incomplete. By Theorem 4.15, we have for each constant A with infinite ${}^{I}A$ that the inclusion $A_\sigma \subsetneq A_*$ is strict. For each such A, we find some $c \in_{\mathscr{U}} {}^{I_0}A$ which cannot be written in the form ${}^{I_0}a$ where the sentence $a \in A$ is true. Note that c belongs to $\mathscr{I}'$, since ${}^{I_0}A$ is the equivalence class of a constant mapping $f : J \to \widehat{S}$. Similarly, all constant ${}^{I_0}a$ with $a \in A$ belong to $\mathscr{I}'$. Thus each of the sentences $c \neq {}^{I_0}a$ can be formulated in the language of Theorem 4.18 and is true. Theorem 4.18 thus implies ${}^{\sigma}A \neq {}^*A$.

Conversely, if S is infinite and $*$ is a nonstandard embedding, let A be an infinite countable subset of S. Then ${}^*A \subsetneq {}^{\sigma}A$, i.e. we find some $c \in {}^{\sigma}A$ which cannot be written in the form ${}^*a$ with $a \in A$. Since c is an internal element, it occurs in the image of φ. With a similar argument as above, Theorem 4.18 now implies $A_\sigma \neq A_*$, and Theorem 4.15 shows that the ultrafilter $\mathscr{U}$ is δ-incomplete. $\square$

Together with Proposition 4.19, we find a natural characterization of internal sets in our model:

Corollary 4.21. *With the above notation, we have: An element x is internal if and only if it arises from a map $f : J \to A$ with an entity $A \in \widehat{S}$ in the sense that $x = \varphi([f])$. Moreover, $x = {}^*a$ is standard if and only if f can be chosen constant $f(j) = a$.* $\square$

Example 4.22. For the map $*$ from our ultrapower model (Theorem 4.20), one has a simple interpretation for standard relations ${}^*\Phi$ where $\Phi \subseteq X_1 \times \cdots \times X_n$:

Note that ${}^*\Phi = \varphi(f_\Phi)$ where $f_\Phi : J \to \widehat{S}$ denotes the constant function $f_\Phi(j) := \Phi$. In view of Proposition 4.19, each element $x_k \in {}^*X_k$ has a representation $x_k = \varphi([f_k])$ where $f_k : J \to X_k$. We claim that

$$(x_1, \ldots, x_n) \in {}^*\Phi \iff (f_1(j), \ldots, f_n(j)) \in \Phi \text{ for almost all } j.$$

Indeed, Theorem 4.18 implies that $(x_1, \ldots, x_n) \in \Phi$ is true if and only if the corresponding formalized sentence is true if interpreted in the abstract model $\mathscr{S}$. But this just means that $(f_{x_1}(j), \ldots, f_{x_n}(j)) \in f_\Phi(j)$ for almost all j.

Example 4.23. Let us illustrate the importance of Example 4.22 for a standard function ${}^*f$ when $f : X \to Y$:

If $x = \varphi([f_x])$ with $f_x : J \to X$, then ${}^*f(x) = \varphi([f_y])$ where $f_y : J \to Y$ is given by $f_y(j) = f(f_x(j))$. Thus, the extension of a function f to a function ${}^*f$ is indeed defined in the canonical way announced in Section 1.1.

Although the model of Theorem 4.20 is rather "constructive", the reader should be aware that actually the ultrafilter $\mathscr{U}$ is a rather "unknown" element: Except for very special cases it is impossible to decide from the representation of two nonstandard elements x, y whether they satisfy e.g. a simple sentence like $x \in y$ or $x = y$:

Exercise 13. Consider the model of Theorem 4.20 with a countable infinite set J. Let $A \in \widehat{S}$, and $x, y \in {}^*A$, i.e. $x = \varphi([f_x])$ and $y = \varphi([f_y])$ where $f_x, f_y : J \to A$. Give a necessary and sufficient condition on f_x and f_y such that the identity $x = y$ holds for any choice of a δ-incomplete ultrafilter $\mathscr{U}$.

For our particular map $*$, we can prove a special case of Theorem 3.23 already. (Actually, the model theoretic proof of Theorem 3.23 which can be found in [CK90] reduces the general result to a variation of the following special case; our proof in Section 7.1 will be rather different).

Exercise 14. Prove from the definition of the map $*$ in Theorem 4.20 that

1. ${}^*A = {}^\sigma A$ if A is a finite entity.
2. ${}^*A$ is uncountable if A is infinite and $\mathscr{U}$ is δ-incomplete (in particular, ${}^*A \neq {}^\sigma A$ for countable A).

   Hint: If $f_1, f_2, \ldots : J \to A$, construct a function $f : J \to A$ such that $f(j) \neq f_1(j)$ everywhere, $f(j) \neq f_2(j)$ almost everywhere, $f(j) \neq f_3(j)$ on a smaller set but still almost everywhere, etc.

In particular, standard entities are either finite or uncountable if the mapping $*$ in Theorem 4.20 is a nonstandard embedding.

Chapter 3

Nonstandard Real Analysis

Throughout this chapter, let $* : \widehat{S} \to \widehat{{}^*S}$ be a nonstandard embedding, and $\mathbb{R} \in \widehat{S}$ be an entity.

§5 Hyperreal Numbers

5.1 Hyperreal and Hypernatural Numbers

Let ${}^*\mathbb{R}$ be the value of the $*$-transform of $\mathbb{R}$. The elements of ${}^*\mathbb{R}$ are called the *hyperreal numbers*.

We first introduce the notation for the most important functions on ${}^*\mathbb{R}$: These are $+ : \mathbb{R}^2 \to \mathbb{R}$ (defined in the usual way), and similarly subtraction, multiplication, division, and exponentiation. These functions are mapped by $*$ into functions ${}^*(\mathbb{R}^2) \to {}^*\mathbb{R}$. Note that ${}^*(\mathbb{R}^2) = ({}^*R)^2$, and so in particular, e.g. ${}^*+ : ({}^*\mathbb{R})^2 \to {}^*\mathbb{R}$. Instead of writing ${}^*+(a, b)$ for $a, b \in {}^*\mathbb{R}$, we use the traditional notation $a{}^*{+}b$. For the sake of convenience, we will later also drop the symbol * in this connection and just write $a + b$. However, for the beginning and to avoid confusion, we will keep this symbol.

Proposition 5.1. *If $a, b \in \mathbb{R}$, then ${}^*(a + b) = {}^*a \, {}^*{+} \, {}^*b$; similarly for multiplication, division, and exponentiation.*

Proof. If c denotes the constant $a + b$, then $c = a + b$ is a true bounded sentence, and so its $*$-transform ${}^*c = {}^*a \, {}^*{+} \, {}^*b$ is true by the transfer principle. □

Thus, $*$ is an isomorphism of $\mathbb{R}$ into ${}^\sigma\mathbb{R}$. In particular:

Corollary 5.2. ${}^\sigma\mathbb{R}$ *is a field.*

However, the hyperreal numbers would not be useful if we had the field property only for the copy ${}^{\sigma}\mathbb{R}$ of $\mathbb{R}$: We want to have the same rules even for the larger set ${}^*\mathbb{R}$. The essential point in the following result is that the statement holds not only for elements of the form ${}^*a$ where a is from the standard universe, but even for nonstandard elements:

Proposition 5.3. *The set* ${}^*\mathbb{R}$ *equipped with the relations* ${}^*\!+$ *and* ${}^*\cdot$ *is a field. The neutral element of addition and multiplication is* ^{*0} *and* ^{*1}*, respectively. The inverse element of* $a \in {}^*\mathbb{R}$ *for addition is* ${}^*0 \,{}^*\!- a$*, and the inverse element of* $a \in {}^*\mathbb{R} \setminus \{{}^*0\}$ *for multiplication is* ${}^*1 \,{}^*\!/\, a$.

Proof. The commutative law for addition follows by applying the transfer principle for the transitively bounded true sentence

$$\forall \underline{x}, \underline{y} \in \mathbb{R} : \underline{x} + \underline{y} = \underline{y} + \underline{x}.$$

The commutative law for multiplication, and the associative and the distributive laws are proved analogously. The transfer of the formula $\forall \underline{x} \in \mathbb{R} : \underline{x} + 0 = \underline{x}$ shows that ^{*0} is the neutral element of the addition ${}^*\!+$, and the transfer of the formula

$$\forall \underline{x} \in \mathbb{R} : \underline{x} + (0 - \underline{x}) = 0$$

with the evident shortcuts implies

$$\forall \underline{x} \in {}^*\mathbb{R} : \underline{x} \,{}^*\!+ ({}^*0 \,{}^*\!- \underline{x}) = {}^*0$$

which means that ${}^*0 \,{}^*\!- a$ is the inverse element of addition. The proof concerning multiplication is similar, if one observes that ${}^*(\mathbb{R} \setminus \{0\}) = {}^*\mathbb{R} \setminus {}^*\{0\} = {}^*\mathbb{R} \setminus \{{}^*0\}$. □

We note that the inverse element of a with respect to addition and multiplication is usually denoted by $-a$ resp. a^{-1}. One might thus define functions $f_1(a) = -a$ and $f_2(a) = a^{-1}$ on $\mathbb{R}$. The $*$-transform gives us hyperreal functions ${}^*f_i : {}^*\mathbb{R} \to {}^*\mathbb{R}$. One may ask whether ${}^*f_1(a)$ is the inverse element of a with respect to addition even for all hyperreal numbers $a \in {}^*\mathbb{R}$. The transfer of the formula $\forall \underline{x} \in \mathbb{R} : \underline{x} + f_1(\underline{x}) = 0$ shows that this is indeed the case. Similarly, ${}^*f_2(a)$ is the inverse of a with respect to multiplication for any $a \in {}^*\mathbb{R} \setminus \{{}^*0\}$.

One might also interpret $f_2(a)$ as the result of the exponential function $e(a, b) := a^b$ with $b := -1$ and may ask whether the $*$-transform of e yields the same function, i.e. whether ${}^*f_2(a) = {}^*e(a, {}^*(-1))$ for any hyperreal number $a \in {}^*\mathbb{R}$. This is indeed the case, as follows by the transfer principle from the sentence

$$\forall \underline{x} \in \mathbb{R} \setminus \{0\} : f_2(\underline{x}) = e(\underline{x}, -1)$$

in view of the fact that ${}^*(\mathbb{R}\setminus\{0\}) = {}^*\mathbb{R}\setminus\{{}^*0\}$ (recall the proof of Proposition 5.3). Thus the $*$-transform of $a \mapsto a^{-1}$ yields always the same nonstandard function, no matter how the symbol a^{-1} is interpreted.

We hope that the reader already has the impression that all elementary properties of $\mathbb{R}$ carry over to ${}^*\mathbb{R}$ in the canonical way. The limitations of this transfer will be made clear later.

Let us now also consider the order properties of ${}^*\mathbb{R}$: The relation $\leq$ on $\mathbb{R}$ is described by a subset of $\mathbb{R}^2$, namely $\leq := \{(a,b) : a \leq b\}$. The $*$-transform of this set is a relation on ${}^*\mathbb{R}$. We write $a \,{}^*{\leq}\, b$ (or later more briefly $a \leq b$), if the pair $(a,b) \in ({}^*\mathbb{R})^2$ belongs to ${}^*{\leq}$. Similarly, we define the meaning of symbols like ${}^*{<}$ or ${}^*{>}$ for hyperreal numbers.

The transfer principle of course implies that the relation $a \leq b$ for elements a, b of the standard universe gives ${}^*a {}^*{\leq} {}^*b$. In particular, ${}^*{\leq}$ is a total order on the standard copy ${}^\sigma\mathbb{R}$ of $\mathbb{R}$. However, ${}^*{\leq}$ is even a total order in the nonstandard universe, as we shall show now. Observe that this does not follow from the above argument, since ${}^\sigma\mathbb{R}$ is a *strict* subset of ${}^*\mathbb{R}$.

We note that we could alternatively have defined ${}^*{<}$ on ${}^*\mathbb{R}$ by

$$a \,{}^*{<}\, b \iff (a \,{}^*{\leq}\, b \wedge a \neq b).$$

The following result implies that these two possible definitions coincide. An analogous remark holds for ${}^*{>}$ and ${}^*{\geq}$.

Proposition 5.4. *${}^*{\leq}$ defines a total order on ${}^*\mathbb{R}$. Moreover, for hyperreal numbers $a, b \in {}^*\mathbb{R}$ the following holds true:*

1. *We have $a \,{}^*{\leq}\, b$ if and only if $a {}^*{<} b$ or $a = b$.*
2. *We have $a \,{}^*{\geq}\, b$ if and only if $a {}^*{>} b$ or $a = b$.*
3. *Precisely one of the three relations $a \,{}^*{<}\, b$, $a \,{}^*{>}\, b$, $a = b$ holds.*

Proof. Let α be the bounded sentence $\forall \underline{x} \in \mathbb{R} : \underline{x} \leq \underline{x}$. This sentence is true, and so the transfer principle implies that $\forall \underline{x} \in {}^*\mathbb{R} : \underline{x} \,{}^*{\leq}\, \underline{x}$, i.e. $a \,{}^*{\leq}\, a$ for any hyperreal number $a \in {}^*\mathbb{R}$. Similarly, the transfer of the sentence $\forall \underline{x}, \underline{y} \in \mathbb{R} : ((\underline{x} \leq \underline{y} \wedge \underline{y} \leq \underline{x}) \implies \underline{x} = \underline{y})$ shows that the relations $a \leq b$ and $b \leq a$ for hyperreal numbers a, b imply $a = b$. The proof of the other properties is similar. □

Let us now summarize:

Theorem 5.5. *With the above notation, ${}^\sigma\mathbb{R}$ and ${}^*\mathbb{R}$ are ordered fields, and ${}^\sigma\mathbb{R}$ is isomorphic to $\mathbb{R}$.*

Proof. The statement for ${}^\sigma\mathbb{R}$ is trivial, since by the transfer principle, $* : \mathbb{R} \to {}^\sigma\mathbb{R}$ is an isomorphism. Concerning ${}^*\mathbb{R}$, we have proved already everything up to the relations between arithmetic and order operations. But these follow by the transfer principle from the sentences

$$\forall \underline{x}, \underline{y}, \underline{z} \in \mathbb{R} : \underline{x} < \underline{y} \implies (\underline{x} + \underline{z} < \underline{y} + \underline{z})$$

and

$$\forall \underline{x}, \underline{y}, \underline{z} \in \mathbb{R} : (\underline{x} < \underline{y} \wedge \underline{z} > 0) \implies (\underline{x} \cdot \underline{z} < \underline{y} \cdot \underline{z})$$

where we used evident abbreviations (henceforth, we will use such shortcuts without further mention). □

We also use notation like ${}^*|\cdot|$, ${}^*\max(\cdot, \cdot)$ for the transfer of the functions with their evident meanings.

All elementary formulas like ${}^*|a| = {}^*\max(a, {}^*{-a})$ (for *hyperreal* numbers $a \in {}^*\mathbb{R}$) follow immediately from the transfer principle and will be used henceforth without further mention. Moreover, we will henceforth drop the symbol * on such simple functions and on simple constants, if no confusion arises. Thus, 0 may either mean the element 0 in the set $\mathbb{R}$, or the element ^{*0} in the set ${}^*\mathbb{R}$ (or in ${}^\sigma\mathbb{R}$).

Example 5.6. For the map $*$ from our ultrapower model (Theorem 4.20), one has a simple interpretation for the elementary operations: Recall that any element $x \in {}^*\mathbb{R}$ has the form $x = \varphi([f])$ with a function $f : J \to \mathbb{R}$ (recall Proposition 4.19). To simplify notation, we write f_x for such an f. Note that if $x \in {}^\sigma\mathbb{R}$, then x is standard, and so f_x may be chosen constant. We claim that

$$\begin{aligned} z = x + y &\iff f_z(j) = f_x(j) + f_y(j) \text{ for almost all } j, \\ z = x \cdot y &\iff f_z(j) = f_x(j) \cdot f_y(j) \text{ for almost all } j, \\ x \leq y &\iff f_x(j) \leq f_y(j) \text{ for almost all } j. \end{aligned}$$

To see this, recall that $+$, $\cdot$ and $\leq$ are just standard relations, and apply Example 4.22.

We define the *hypernatural numbers* as the elements of ${}^*\mathbb{N}$, and similarly the *hyperinteger numbers* and the *hyperrational numbers* as the elements of ${}^*\mathbb{Z}$ and ${}^*\mathbb{Q}$, respectively. Note that the relations $\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R}$ imply ${}^*\mathbb{N} \subseteq {}^*\mathbb{Z} \subseteq {}^*\mathbb{Q} \subseteq {}^*\mathbb{R}$.

There are two natural definitions for the order on ${}^*\mathbb{N}$: Either, we can define the order as the restriction of the order on ${}^*\mathbb{R}$ to ${}^*\mathbb{N}$, or we can use the $*$-transform of the order of $\mathbb{N}$. As the reader might have expected, the two definitions actually coincide: One may apply the transfer principle to see this, or just has to recall Theorem 3.13.

We will show soon that ${}^*\mathbb{R}$ contains, besides the copy ${}^\sigma\mathbb{R}$ of $\mathbb{R}$, also infinite and infinitesimal numbers:

Definition 5.7. A hyperreal number $x \in {}^*\mathbb{R}$ is called

1. *finite*, if there is some $n \in {}^\sigma\mathbb{N}$ such that $|x| < n$,
2. *infinite*, if for any $n \in {}^\sigma\mathbb{N}$ we have $|x| > n$,
3. *infinitesimal*, if $|x| < n^{-1}$ for any $n \in {}^\sigma\mathbb{N}$.

We use the notation

$$\mathrm{fin}({}^*\mathbb{R}) := \{x \in {}^*\mathbb{R} : x \text{ is finite}\},$$
$$\inf({}^*\mathbb{R}) := \{x \in {}^*\mathbb{R} : x \text{ is infinitesimal}\}.$$

The notation $\inf({}^*\mathbb{R})$ is of course ambiguous, since the symbol inf is usually reserved for the infimum; however, we hope that no confusion will arise.

It will be convenient to use the notation $\mathbb{R}_+ := \{\underline{x} \in \mathbb{R} : \underline{x} > 0\}$. Then we have ${}^\sigma\mathbb{R}_+ = \{\underline{x} \in {}^\sigma\mathbb{R} : \underline{x} > 0\}$ and ${}^*\mathbb{R}_+ = \{\underline{x} \in {}^*\mathbb{R} : \underline{x} > 0\}$.

Any $x \in {}^*\mathbb{R}$ is either finite or infinite.

Proposition 5.8. *A number $x \in {}^*\mathbb{R}$ is*

1. *finite, if and only if $|x| < y$ for some $y \in {}^\sigma\mathbb{R}$,*
2. *infinite, if and only if $|x| > y$ for any $y \in {}^\sigma\mathbb{R}$,*
3. *infinitesimal, if and only if $|x| < \varepsilon$ for any $\varepsilon \in {}^\sigma\mathbb{R}_+$.*

Proof. One implication follows immediately from ${}^\sigma\mathbb{N} \subseteq {}^\sigma\mathbb{R}$. For the converse implication note that ${}^\sigma\mathbb{R}$ is Archimedean, i.e. for any $y \in {}^\sigma\mathbb{R}$ we find some $n \in {}^\sigma\mathbb{N}$ such that $n > y$. Thus, if x is infinite, we have $|x| > n > y$. Similarly, if x is infinitesimal and $\varepsilon \in {}^\sigma\mathbb{R}_+$, we put $y := \varepsilon^{-1}$ and then find some $n \in {}^\sigma\mathbb{N}$ with $|x| < n^{-1} < y^{-1} = \varepsilon$. □

Note that there is of course no $x \in {}^*\mathbb{R}_+$ which satisfies $x < \varepsilon$ for any *hyperreal* number $\varepsilon \in {}^*\mathbb{R}_+$ (It is not necessary to invoke the transfer principle to see this: The choice $\varepsilon := x$ is enough).

Although ${}^*\mathbb{R}$ contains both, infinite numbers and infinitesimals, one might expect that ${}^*\mathbb{N}$ contains, besides ${}^\sigma\mathbb{N}$, only infinite numbers. This is indeed the case. Moreover, there *are* infinite numbers in ${}^*\mathbb{N}$:

Recall that, since $*$ is a nonstandard embedding, we have ${}^\sigma\mathbb{N} \subsetneq {}^*\mathbb{N}$. For evident reasons, we put

$$\mathbb{N}_\infty := {}^*\mathbb{N} \setminus {}^\sigma\mathbb{N}.$$

Proposition 5.9. *If $h \in \mathbb{N}_\infty$, then h is infinite, i.e.*

$${}^*N \cap \mathrm{fin}({}^*\mathbb{R}) = {}^\sigma\mathbb{N}.$$

Proof. Given some $N \in \mathbb{N}$, consider the sentence

$$\forall \underline{x} \in \mathbb{N} : (\underline{x} \neq 1 \wedge \underline{x} \neq 2 \wedge \cdots \wedge \underline{x} \neq N \implies \underline{x} > N).$$

The transfer principle implies that any $h \in {}^*\mathbb{N}$ which does not have the form $h = {}^*k$ with $k = 1, 2, \ldots, N$ (in particular, any $h \in \mathbb{N}_\infty$) satisfies $h > {}^*N$. Since $N \in \mathbb{N}$ was arbitrary, the statement follows. □

Corollary 5.10. *In* ${}^*\mathbb{R}$ *there are infinite numbers and nonzero infinitesimal numbers.*

Proof. By Proposition 5.9, there is an infinite number h, and so $1/h$ is infinitesimal (and different from 0). □

Exercise 15. Show that for each $x \in {}^*\mathbb{R}$ there is precisely one $h \in {}^*\mathbb{N}$ with $h \le |x| < h + 1$. Moreover, prove that x is infinite if and only if $h \notin {}^\sigma\mathbb{N}$.

Example 5.11. For the map $*$ from our ultrapower model (Theorem 4.20), it is easy to characterize the finite numbers: Recall that $x \in {}^*\mathbb{R}$ if and only if $x = \varphi([f])$ where $f : J \to \mathbb{R}$. By definition, x is finite if and only if $|x| \le {}^*n$ for some $n \in \mathbb{N}$. Example 5.6 implies that the relation $|x| \le {}^*n$ for $x = \varphi([f])$ is equivalent to $f(j) \le n$ for almost all j. Hence, $x = \varphi([f])$ is finite if and only if $[f]$ has a representing function which is bounded on J.

Now we come to the limitations of the transfer principle:

Theorem 5.12. *The set* ${}^*\mathbb{N}$ *is* not *well-ordered. More precisely,* $\mathbb{N}_\infty$ *has no smallest element. However, any nonempty* internal *subset of* ${}^*\mathbb{N}$ *has a smallest element.*

Proof. Assume by contradiction that $\mathbb{N}_\infty$ has a smallest element h. Then $h > {}^*n$ for each $n \in \mathbb{N}$, and so $h - 1 > {}^*n$ for each $n \in \mathbb{N}$. Since no element $n_0 \in \mathbb{N}$ satisfies ${}^*n_0 > {}^*n$ for all $n \in \mathbb{N}$, we may conclude that $h - 1 \notin {}^\sigma\mathbb{N}$. Hence, the element $h_0 := h - 1$ belongs to $\mathbb{N}_\infty$ and is strictly smaller than h, a contradiction.

Since $\mathbb{N}$ is well-ordered, the sentence

$$\forall \underline{x} \in \mathcal{P}(\mathbb{N}) : (\underline{x} \ne \varnothing \implies \exists \underline{y} \in \underline{x} : \forall \underline{z} \in \underline{x} : \underline{y} \le \underline{z})$$

is true. The transfer of this sentence implies that any nonempty $\underline{x} \in {}^*\mathcal{P}(\mathbb{N})$ has a smallest element. By Theorem 3.21, the set ${}^*\mathcal{P}(\mathbb{N})$ consists precisely of all internal subsets of ${}^*\mathbb{N}$. □

Theorem 5.12 implies that ${}^*N \setminus {}^\sigma\mathbb{N}$ is external, and so ${}^\sigma\mathbb{N}$ is external. In the light of this proof, the reader might want to reconsider the proof of Theorem 3.22: This proof is actually just a repetition of the arguments that we used above.

The reason why the transfer principle does not apply for the sentence "$\mathbb{N}$ is well-ordered" is that the natural formalization of this sentence is unbounded (namely it has the form $\forall \underline{x} \subseteq \mathbb{N} : \ldots$); recall in this connection the remark following Theorem 3.13.

Proposition 5.9 shows another limitation of the transfer principle:

Theorem 5.13. *The ordered field* ${}^*\mathbb{R}$ *is not Archimedean and not isomorphic to a subfield of* $\mathbb{R}$.

Proof. Using the notation of Section 1.2 with $X := {}^*\mathbb{R}$, we evidently have $X_{\mathbb{Q}} = {}^\sigma\mathbb{Q}$ and $X_{\mathbb{N}} = {}^\sigma\mathbb{N}$. Choose some $h \in {}^*N \setminus {}^\sigma\mathbb{N}$. Since ${}^*\mathbb{N} \subseteq {}^*\mathbb{R}$, we then have $h \in {}^*\mathbb{R}$, but by Proposition 5.9 there is no $n \in {}^\sigma\mathbb{N}$ with $n > h$. This means that $X = {}^*\mathbb{R}$ is not Archimedean. The second statement follows from the first by Theorem 1.3. □

Theorem 5.14. *The set* ${}^*\mathbb{R}$ *is not Dedekind complete. However, any nonempty* internal *subset of* ${}^*\mathbb{R}$ *which is bounded from above has a least upper bound.*

Proof. The subset ${}^\sigma\mathbb{N} \subseteq {}^*\mathbb{R}$ is bounded from above by Proposition 5.9. However, ${}^\sigma\mathbb{N}$ has no least upper bound: If $x \in {}^*\mathbb{R}$ is an upper bound for ${}^\sigma\mathbb{N}$, i.e. $x > n$ for each $n \in {}^\sigma\mathbb{N}$, then also $x - 1 > n$ for each $n \in {}^\sigma\mathbb{N}$, i.e. $x - 1 \in {}^*\mathbb{R}$ is an upper bound for ${}^\sigma\mathbb{N}$ which is strictly smaller than x.

For the second statement, observe that $\mathbb{R}$ is Dedekind complete which means that

$$\forall \underline{x} \in \mathcal{P}(\mathbb{R}) : (\underline{x} \neq \varnothing \wedge \alpha(\underline{x}, \mathbb{R})) \implies \beta(\underline{x}, \mathbb{R})$$

is true where $\alpha(\underline{x}, \mathbb{R})$ and $\beta(\underline{x}, \mathbb{R})$ are transitively bounded formulas with the meaning "$\underline{x}$ has an upper bound in $\mathbb{R}$" and "$\underline{x}$ has a smallest upper bound in $\mathbb{R}$", respectively (we leave the precise formulation of the formulas to the reader). The transfer of the above statement means that any internal subset of ${}^*\mathbb{R}$ which is nonempty and bounded from above in ${}^*\mathbb{R}$ possesses a smallest upper bound in ${}^*\mathbb{R}$ (recall that ${}^*\varnothing = \varnothing$ and that ${}^*\mathcal{P}(\mathbb{R})$ consists by Theorem 3.21 precisely of all internal subsets of ${}^*\mathbb{R}$). □

The fact that ${}^*\mathbb{R}$ is not Dedekind complete should not be too surprising to the reader, since the transfer principle simply does not provide much information on external sets: The "natural" formalization of the sentence that a field is Dedekind complete is not transitively bounded.

However, the reader might be surprised that ${}^*\mathbb{R}$ is not Archimedean, because the sentence that $\mathbb{R}$ is Archimedean can in a natural way be formalized by the transitively bounded sentence

$$\forall \underline{x} \in \mathbb{R} : \exists \underline{y} \in \mathbb{N} : \underline{y} > \underline{x}.$$

Of course, the $*$-transform of this sentence must be true:

$$\forall \underline{x} \in {}^*\mathbb{R} : \exists \underline{y} \in {}^*\mathbb{N} : \underline{y} > \underline{x}.$$

However, this sentence does *not* mean that ${}^*\mathbb{R}$ is Archimedean, because for $X = {}^*\mathbb{R}$ the set $\mathbb{N}_X$ from Section 1.2 is ${}^\sigma\mathbb{N}$ and not ${}^*\mathbb{N}$.

Proposition 5.15. $\mathrm{fin}({}^*\mathbb{R})$ *is an Archimedean subring of* ${}^*\mathbb{R}$ *without zero-divisors which contains* ${}^\sigma\mathbb{R}$ *and* $\inf({}^*\mathbb{R})$.

$\mathrm{fin}({}^*\mathbb{R})$ *is not a field. More precisely, we have* $1/x \in \mathrm{fin}({}^*\mathbb{R})$ *for some* $x \in {}^*\mathbb{R}$ $(x \neq 0)$ *if and only if* $x \notin \inf({}^*\mathbb{R})$.

Proof. ${}^\sigma\mathbb{R} \subseteq \mathrm{fin}({}^*\mathbb{R})$ follows from Proposition 5.8, and the fact that $\mathrm{fin}({}^*\mathbb{R})$ is Archimedean follows from the definition (note that for $X := \mathrm{fin}({}^*\mathbb{R})$, we have $\mathbb{N}_X = {}^\sigma\mathbb{N}$). If $\mathrm{fin}({}^*\mathbb{R})$ would have zero-divisors, we would have $x \cdot y = 0$ for $x, y \in \mathrm{fin}({}^*\mathbb{R}) \subseteq {}^*\mathbb{R}$, contradicting the fact that ${}^*\mathbb{R}$ is a field. $\mathrm{fin}({}^*\mathbb{R})$ is a subring:

If $x, y \in \mathrm{fin}({}^*\mathbb{R})$, then $x+y$, $x-y$, and xy also belong to $\mathrm{fin}({}^*\mathbb{R})$: Indeed, there are $n, m \in {}^\sigma\mathbb{N}$ with $|x| \leq n$, $|y| \leq m$ which implies $|x \pm y| \leq |x|+|y| \leq n+m \in {}^\sigma\mathbb{N}$, and similarly $|xy| \leq nm \in {}^\sigma\mathbb{N}$.

If $x \in \inf({}^*\mathbb{R})$, then $|x| < n^{-1} < 1$ for all $n \in {}^\sigma\mathbb{N}$. Hence, $x \in \mathrm{fin}({}^*\mathbb{R})$ and $y := 1/x \notin \mathrm{fin}({}^*\mathbb{R})$ (if $x \neq 0$) because $|y| > n$ for all $n \in {}^\sigma\mathbb{N}$.

Conversely, if $x \notin \inf({}^*\mathbb{R})$, then we find some n with $|x| \geq n^{-1}$. Hence, $y := 1/x$ satisfies $|y| \leq n$ and does not belong to $\mathrm{fin}({}^*\mathbb{R})$. □

Definition 5.16. We say that two hyperreal numbers $x, y \in {}^*\mathbb{R}$ are *infinitely close* to each other, if $x - y \in \inf({}^*\mathbb{R})$. We then write $x \approx y$.

Proposition 5.17. $\approx$ *is an equivalence relation on* ${}^*\mathbb{R}$. *Moreover, if* $x_1 \approx y_1$ *and* $x_2 \approx y_2$ *we have:*

1. $x_1 \pm x_2 \approx y_1 \pm y_2$.
2. $x_1 \cdot x_2 \approx y_1 \cdot y_2$ *if* x_1 *and* x_2 *are finite.*
3. $x_1/x_2 \approx y_1/y_2$ *if* x_1 *is finite and* $x_2 \not\approx 0$.

Proof. Since $0 \in \inf({}^*\mathbb{R})$, we have $x \approx x$. Since $|x-y| = |y-x|$, the relation $x \approx y$ implies $y \approx x$. Finally, if $x \approx y$ and $y \approx z$, then $|x-z| \leq |x-y| + |y-z| < 2\varepsilon$ for any $\varepsilon \in {}^\sigma\mathbb{R}_+$ which implies $x \approx z$.

For each $n \in {}^\sigma\mathbb{N}$ we have $|x_i - y_i| < n^{-1}$ $(i = 1, 2)$. Hence, $|(x_1 \pm x_2) - (y_1 \pm y_2)| \leq 2n^{-1}$ for each $n \in {}^\sigma\mathbb{N}$. If x_1 and x_2 are finite, we find some $N \in {}^\sigma\mathbb{N}$ with $|x_i| \leq N$ $(i = 1, 2)$. Since $|x_2 - y_2| \leq 1$, we find $|y_2| \leq N + 1$, and so $|x_1 \cdot x_2 - y_1 \cdot y_2| = |x_1(x_2 - y_2) + (x_1 - y_1)y_2| \leq Nn^{-1} + n^{-1}N = 2Nn^{-1}$ for each $n \in {}^\sigma\mathbb{N}$. Finally, if $x_2 \not\approx 0$, then we find some $N \in {}^\sigma\mathbb{N}$ with $|x_2| \geq N^{-1}$. Since $|x_2 - y_2| \leq N^{-1}/2$, we also have $|y_2| \geq N^{-1}/2$. Consequently,

$$\left|\frac{1}{x_2} - \frac{1}{y_2}\right| = \left|\frac{y_2 - x_2}{x_2 y_2}\right| \leq \frac{n^{-1}}{N^{-1}N^{-1}/2}$$

for each $n \in {}^\sigma\mathbb{N}$ which proves $x_2^{-1} \approx y_2^{-1}$. Since x_2^{-1} is finite, it follows by what we just proved that $x_1/x_2 = x_1 x_2^{-1} \approx y_1 y_2^{-1} = y_1/y_2$. □

The reader might have observed that the above proof essentially repeats the argument of the classical limit rules like $\lim(x_n \pm y_n) = \lim x_n \pm \lim y_n$, etc. In fact, we will see later that Proposition 5.17 implies these limit rules.

Corollary 5.18. *For any* $x \in \inf({}^*\mathbb{R})$, $y \in \mathrm{fin}({}^*\mathbb{R})$, *we have* $x \cdot y \in \inf({}^*\mathbb{R})$.

Proof. $x \approx 0$ and $y \approx y$ implies $x \cdot y \approx 0 \cdot y = 0$. □

Exercise 16. What can be said about the "equations" $x^2 = 2$ and $x^2 \approx 2$ in ${}^*\mathbb{Q}$?

Exercise 17. Show that for any $x \in {}^*\mathbb{R}$ there is some $q \in {}^*\mathbb{Q}$ with $x \approx q$.

Theorem 5.19. *For each* $x \in \mathrm{fin}({}^*\mathbb{R})$ *there is precisely one* $\hat{x} \in \mathbb{R}$ *with* $x \approx {}^*\hat{x}$.

Proof. Uniqueness: If $\hat{x}, \hat{y} \in {}^\sigma\mathbb{R}$ satisfy ${}^*\hat{x} \approx x \approx {}^*\hat{y}$, then $|{}^*\hat{x} - {}^*\hat{y}| < {}^*\varepsilon$ for any $\varepsilon \in \mathbb{R}_+$. The inverse transfer principle implies $|\hat{x} - \hat{y}| < \varepsilon$ for any $\varepsilon \in \mathbb{R}_+$, and so $\hat{x} = \hat{y}$.

Existence: Consider the set $A := \{y \in {}^\sigma\mathbb{R} : y < x\}$. Since $x \in \mathrm{fin}({}^*\mathbb{R})$, the set A is nonempty and bounded from above, and since ${}^\sigma\mathbb{R}$ is Dedekind complete (because it is isomorphic to $\mathbb{R}$), it has a least upper bound $s \in {}^\sigma\mathbb{R}$, i.e. $s = {}^*\hat{x}$ for some $\hat{x} \in \mathbb{R}$. Given some $n \in {}^\sigma\mathbb{N}$, we have $s \geq x - n^{-1}$ since otherwise $s + n^{-1} \in A$ would contradict the fact that s is an upper bound for A. But we also have $s \leq x + n^{-1}$, since otherwise $s - n^{-1}$ would be an upper bound for A which is strictly smaller than s. Hence $|s - x| \leq n^{-1}$ for each $n \in {}^\sigma\mathbb{N}$, i.e. $x \approx s = {}^*\hat{x}$. □

We emphasize that the proof of Theorem 5.19 made essential use of the Dedekind completeness of $\mathbb{R}$.

Definition 5.20. Let $\mathrm{st} : \mathrm{fin}({}^*\mathbb{R}) \to \mathbb{R}$ be the map $x \mapsto \hat{x}$ from Theorem 5.19. We call $\mathrm{st}(x) = \hat{x}$ the *standard part* of $x \in \mathrm{fin}({}^*\mathbb{R})$, and st the *standard part homomorphism*.

Theorem 5.21. $\mathrm{st} : \mathrm{fin}({}^*\mathbb{R}) \to \mathbb{R}$ *is a surjective order-preserving ring-homomorphism with kernel* $\inf({}^*\mathbb{R})$, *i.e. for all* $x, x_1, x_2 \in \mathrm{fin}({}^*\mathbb{R})$ *we have*

1. $\mathrm{st}(x) = 0$ *if and only if* $x \in \inf({}^*\mathbb{R})$,
2. $\mathrm{st}(x_1 \pm x_2) = \mathrm{st}(x_1) \pm \mathrm{st}(x_2)$,
3. $\mathrm{st}(x_1 \cdot x_2) = \mathrm{st}(x_1) \cdot \mathrm{st}(x_2)$,
4. $\mathrm{st}(x_1/x_2) = \mathrm{st}(x_1)/\mathrm{st}(x_2)$ *if* $\mathrm{st}(x_2) \neq 0$, *and*
5. $x_1 \leq x_2$ *implies* $\mathrm{st}(x_1) \leq \mathrm{st}(x_2)$.

Hence, st *induces an order-preserving ring-isomorphism*

$$\mathrm{fin}({}^*\mathbb{R}) / \inf({}^*\mathbb{R}) \cong \mathbb{R}.$$

Since $\mathbb{R}$ *is a field, also* $\mathrm{fin}({}^*\mathbb{R})/\inf({}^*\mathbb{R})$ *is a field (and* $\inf({}^*\mathbb{R})$ *is a maximal ideal in* $\mathrm{fin}({}^*\mathbb{R})$*).*

Proof. Let $y_1 := \mathrm{st}(x_1)$ and $y_2 := \mathrm{st}(x_2)$. Then $x_i \approx y_i$, and Proposition 5.17 implies $x_1 + x_2 \approx y_1 + y_2$. Since $y_1 + y_2 \in {}^\sigma\mathbb{R}$, we find $\mathrm{st}(x_1 + x_2) = y_1 + y_2$, and so st commutes with the addition. Analogously, st commutes with multiplication and thus is a ring homomorphism. If $x_1 \le x_2$, then we have for any $n \in {}^\sigma\mathbb{N}$ in view of $|x_i - y_i| \le n^{-1}$ that $y_1 < x_1 + n^{-1} \le x_2 + n^{-1} \le y_2 + n^{-1}$. Since $y_1, y_2 \in {}^\sigma\mathbb{R} \cong \mathbb{R}$, this implies $y_1 \le y_2$.

st is onto ${}^\sigma\mathbb{R}$, since $\mathrm{st}(x) = x$ for $x \in {}^\sigma\mathbb{R}$. Moreover, $\mathrm{st}(x_1) = 0$ if and only if $x_1 \approx 0$, i.e. if and only if $x_1 \in \inf({}^*\mathbb{R})$. □

Definition 5.22. For $x \in \mathbb{R}$, we define the *monad of x* as the set

$$\mathrm{mon}(x) := \{y \in {}^*\mathbb{R} : y \approx {}^*x\}.$$

We have

$$\mathrm{mon}(x) = {}^*x + \inf({}^*\mathbb{R}) = \{y : \mathrm{st}(y) = x\}.$$

Indeed, $y \approx {}^*x$ if and only if $y - x \in \inf({}^*\mathbb{R})$. Now observe that $\inf({}^*\mathbb{R})$ is the kernel of the standard part homomorphism st which satisfies $\mathrm{st}({}^*x) = x$ for $x \in \mathbb{R}$.

Up to now we know that ${}^*\mathbb{R}$ contains more elements than ${}^\sigma\mathbb{R} \cong \mathbb{R}$; we also know that under these new elements are the infinite and the nonzero infinitesimal elements. However, ${}^*\mathbb{R}$ contains even more elements, namely all which are infinitely close to some $x \in {}^\sigma\mathbb{R}$, i.e. all which belong to some monad. The question arises whether there are other "exotic" elements contained in ${}^*\mathbb{R}$. The answer is "no":

Proposition 5.23. *The set* $\mathrm{fin}({}^*\mathbb{R})$ *is the disjoint union of all monads. The elements of* ${}^*\mathbb{R} \setminus \mathrm{fin}({}^*\mathbb{R})$ *are the inverses of the nonzero elements of* $\inf({}^*\mathbb{R}) = \mathrm{mon}(0)$.

Proof. For any $y \in \mathrm{fin}({}^*\mathbb{R})$, we have $y \in \mathrm{mon}(\mathrm{st}(y))$, and so $\mathrm{fin}({}^*\mathbb{R})$ is contained in the union of all monads. Conversely, each monad $\mathrm{mon}(x)$ is contained in $\mathrm{fin}({}^*\mathbb{R})$, since ${}^*x \in \mathrm{fin}({}^*\mathbb{R})$ and $y \approx {}^*x$ implies $y \in \mathrm{fin}({}^*\mathbb{R})$ (we have $|x| \le n$ for some $n \in \mathbb{N}$, and so $|y| \le |{}^*x| + 1 \le {}^*n + 1$). To see that the monads are disjoint, assume that for $x_1, x_2 \in \mathbb{R}$ we find some $y \in \mathrm{mon}(x_1) \cap \mathrm{mon}(x_2)$. Then $\mathrm{st}(y) = x_1$ and $\mathrm{st}(y) = x_2$, i.e. $x_1 = x_2$.

The second statement is a reformulation of the second part of Proposition 5.15. □

Most of the objects we considered so far are actually external. In this connection recall that ${}^*A$ is internal and ${}^\sigma A$ is external for infinite sets A (in particular, for $A = \mathbb{R}, \mathbb{N}, \mathbb{Z}, \mathbb{Q}$).

Theorem 5.24. *The sets* $\inf({}^*\mathbb{R})$, $\mathrm{fin}({}^*\mathbb{R})$, *and* $\mathrm{mon}(x)$ *are external. Moreover, also the mapping* $s : x \mapsto {}^*(\mathrm{st}(x))$ *is external.*

Proof. The set $\inf({}^*\mathbb{R}) \subseteq {}^*\mathbb{R}$ has no least upper bound (and so Theorem 5.14 implies that it is external): Assume that $x \in {}^*\mathbb{R}_+$ is such a bound. If $x \in \inf({}^*\mathbb{R})$, then $2x \in \inf({}^*\mathbb{R})$ (Corollary 5.18) contradicts the fact that x is an upper bound. Hence, $x \notin \inf({}^*\mathbb{R})$ which implies $x/2 \notin \inf({}^*\mathbb{R})$ (since otherwise $x = 2(x/2) \in \inf({}^*\mathbb{R})$). This contradicts the fact that x is the *least* upper bound.

If $\text{fin}({}^*\mathbb{R})$ were internal, then the internal definition principle would imply that

$$\inf({}^*\mathbb{R}) = \{\underline{x} \in {}^*\mathbb{R} : \underline{x} \neq 0 \wedge 1/\underline{x} \in \text{fin}({}^*\mathbb{R})\}$$

is internal, a contradiction. Similarly, if $\text{mon}(x)$ were internal, the internal definition principle would imply that $\inf({}^*\mathbb{R}) = \text{mon}(x) - x$ were internal.

If s were internal, then $\text{dom}(s) = \text{fin}({}^*\mathbb{R})$ were internal by Theorem 3.19 (and even $\text{rng}(s) = {}^\sigma\mathbb{R}$ were internal). □

Exercise 18. Prove that $\{y \in {}^*\mathbb{R} : y \approx x\}$ is external for any $x \in {}^*\mathbb{R}$.

Exercise 19. Prove that any $h \in {}^*\mathbb{N}$ is either even or odd (and also not both), i.e. precisely one of the following alternatives holds:

1. There is some $n \in {}^*\mathbb{N}$ such that $h = 2n$, or
2. There is some $n \in {}^*\mathbb{N}$ such that $h = 2n - 1$.

Prove the statement also for the map $*$ of Theorem 4.20 directly from the definition.

Let us now point out that some "curiosities" of the set ${}^*\mathbb{R}$ correspond to the classical paradoxa which have been associated with the "continuum" (e.g. by Leibniz):

Historically, the continuum has been considered as a line. It was possible to divide this line at any point, but it did not make much sense to speak of a "point" of the line: The "point" could only be considered as an "endpoint" of e.g. a line segment; if there is no line segment, then there is also no "point" to speak of. This intuitive idea is reflected by the fact that monads are external (if we consider only internal subsets of ${}^*\mathbb{R}$ as "reasonable") and that infinite internal sets are always "large" in a certain sense (recall the remarks following Theorem 3.23).

Another correspondence to Leibniz's intuitive ideas is the treatment of infinitesimals: An infinitesimal $dx > 0$ was by Leibniz only considered as a "placeholder for many possibilities" (with the only requirement that it be less than any positive number). This is reflected by the fact, that the standard formal language has no symbols for infinitesimals: We can describe infinitesimals only by means of variables within quantified expressions like $\forall \underline{x} \in {}^*\mathbb{R} : \ldots$. In particular, although we can say in a proof "Fix some infinitesimal c", we actually do not know precisely which infinitesimal will be fixed: Although we can specify even more properties of the infinitesimal, e.g. that it be of the form $1/h$ with $h \in \mathbb{N}_\infty$, we cannot give a property which describes it completely: Indeed, assume that such a property would exist. Formally, this means that there is a standard predicate $\alpha(\underline{x})$ such that $\{\underline{x} \in {}^*\mathbb{R} : \alpha(\underline{x})\} = \{c\}$. But by the standard definition principle, this would imply that $\{c\}$ is a standard set which is not the case (Exercise 5).

5.2 Interpretation of the Standard Part Homomorphism

The standard part homomorphism st is one of the most important functions in nonstandard analysis. Let us interpret this function in the situation of the ultrapower model (Theorem 4.20) (with $\mathscr{U}$ being δ-incomplete).

Recall (Example 5.6) that $x \in \operatorname{fin}({}^*\mathbb{R})$ if and only if $x = \varphi([f])$ where $f : J \to \mathbb{R}$ is bounded.

Proposition 5.25. *If in the above situation $J = \mathbb{N}$ and $\lim_{j\to\infty} f(j) = c$, then* $\operatorname{st}(x) = c$.

Proof. Given $\varepsilon \in \mathbb{R}_+$, the set $N = \{j : |f(j) - c| > \varepsilon\}$ is finite, and so its complement belongs to $\mathscr{U}$ by Exercise 11. Hence $|f(j) - c| \leq \varepsilon$ for almost all j which in view of Example 5.6 means $|x - {}^*c| \leq {}^*\varepsilon$. Hence, $x \approx {}^*c$ which in view of $c \in \mathbb{R}$ actually implies $c = \operatorname{st}(x)$. □

Thus, st is equal to "lim" (if it makes sense to speak of "lim"). Moreover, Theorem 5.21 shows that st has also many properties analogous to "lim". However, st is even defined whenever $f : J \to \mathbb{R}$ is just bounded. Thus, in a certain sense we may consider st as a generalization of a limit to all bounded functions. We will discuss such limits later on.

At the moment we will just recall a definition of "limit" from general topology: The reader who is not familiar with topology may in the following just consider $X = \mathbb{R}$; this is the only case that is actually needed at the moment. However, since it makes no essential difference, we formulate the following results already in a more general context (the reader may want to reread the following after having read §12).

For the rest of this section, let X be some Hausdorff space i.e. a topological space with the property that each two points $x \neq y$ have disjoint open neighborhoods.

Let $\mathscr{F}$ be a filter over some set J, and $f : J \to X$ be a function. One calls the filter generated by $\{f(F) : F \in \mathscr{F}\}$ the *image filter* of $\mathscr{F}$ and denotes this filter by $f(\mathscr{F})$. One says that f converges with respect to $\mathscr{F}$ to some point $x \in X$ if any open neighborhood $U \subseteq X$ of x is an element of the image filter $f(\mathscr{F})$. If f converges with respect to $\mathscr{F}$, we write

$$\lim_{j\to\mathscr{F}} f(j) = x.$$

This notation is justified:

Proposition 5.26. *If X is a Hausdorff space and $\mathscr{F}$ some filter, then f converges to at most one point with respect to $\mathscr{F}$.*

Proof. Assume that f converges to x and to y with respect to $\mathscr{F}$, and that $x \neq y$. Since X is a Hausdorff space, the points x and y have disjoint open neighborhoods, say U_x and U_y. Then $U_x, U_y \in f(\mathscr{F})$, and so $\varnothing = U_x \cap U_y \in f(\mathscr{F})$, contradicting the fact that $f(\mathscr{F})$ is a filter. □

Lemma 5.27. *$f(\mathscr{F})$ consists precisely of those sets $U \subseteq X$ for which $f^{-1}(U) = \{j : f(j) \in U\}$ belongs to $\mathscr{F}$.*

Proof. If $F = f^{-1}(U)$ belongs to $\mathscr{F}$, then $U \supseteq f(F) \in f(\mathscr{F})$. Since $f(\mathscr{F})$ is a filter, this implies $U \in f(\mathscr{F})$. Conversely, if $U \in f(\mathscr{F})$, then there are finitely many $F_1, \dots, F_n \in \mathscr{F}$ such that $U \supseteq f(F_1) \cap \cdots \cap f(F_n)$. The set $F := F_1 \cap \cdots \cap F_n$ belongs to $\mathscr{F}$ (because $\mathscr{F}$ is a filter), and the set $f(F)$ is contained in $f(F_1) \cap \cdots \cap f(F_n)$, and so $U \supseteq f(F)$, i.e. $f^{-1}(U) \supseteq F$. This implies $f^{-1}(U) \in \mathscr{F}$. □

The definition of convergence with respect to a filter contains the usual notions of convergence as special cases:

Proposition 5.28. *Let $J = \mathbb{N}$, X be a topological space, and $\mathscr{F}$ be the filter of Example 4.2. Then*

$$\lim_{j \to \mathscr{F}} f(j) = x \iff \lim_{j \to \infty} f(j) = x.$$

Proof. We have $\lim_{j\to\infty} f(j) = x$ if and only if for any open neighborhood U of x we have $f(j) \in U$ for all except finitely many j. The latter means $f^{-1}(U) = \{j : f(j) \in U\} \in \mathscr{F}$ which by Lemma 5.27 is equivalent to $U \in f(\mathscr{F})$. Hence, $\lim_{j\to\infty} f(j) = x$ if and only if any open neighborhood U of x is contained in $f(\mathscr{F})$, i.e. if and only if $\lim_{j\to\mathscr{F}} f(j) = x$. □

Exercise 20. Let J, X be topological spaces, $j_0 \in J$, and let $\mathscr{F}$ be the filter generated by all sets of the form $J_0 \setminus \{j_0\}$ where J_0 is an open neighborhood of j_0 (we assume that none of these sets is empty which implies that the family of these sets indeed has the finite intersection property). Prove that

$$\lim_{j \to \mathscr{F}} f(j) = x \iff \lim_{j \to j_0} f(j) = x.$$

The finer the filter, the more convergent functions exist:

Proposition 5.29. *Let $\mathscr{F}_1, \mathscr{F}_2$ be filters over J with $\mathscr{F}_1 \subseteq \mathscr{F}_2$. If f converges to x with respect to $\mathscr{F}_1$, then f converges to x with respect to $\mathscr{F}_2$.*

Proof. We have $f(\mathscr{F}_1) \subseteq f(\mathscr{F}_2)$. Hence, if any open neighborhood of x is contained in $f(\mathscr{F}_1)$, it is also contained in $f(\mathscr{F}_2)$. □

The crucial point for ultrafilters is that all bounded functions have a limit:

Theorem 5.30. *If $\mathscr{U}$ is an ultrafilter over J and $f : J \to \mathbb{R}$ is bounded, then f converges with respect to $\mathscr{U}$ to some point $x \in \mathbb{R}$.*

Proof. Since f is bounded, the image $f(J)$ is contained in a compact interval $[a, b]$. Let $\mathscr{B}$ be the system of all open sets which do not belong to $f(\mathscr{U})$. We claim that $\bigcup \mathscr{B}$ does not contain $[a, b]$: Indeed, since $[a, b]$ is compact, we find otherwise finitely many $B_1, \dots, B_n \in \mathscr{B}$ with $[a, b] \subseteq B_1 \cup \dots \cup B_n$. Put $A_i := \{j : f(j) \in B_i\}$ $(i = 1, \dots, n)$. Then $A_i \notin \mathscr{U}$, since otherwise $f(A_i) \in f(\mathscr{U})$ which would imply the contradiction $B_i \in f(\mathscr{U})$. Since $\mathscr{U}$ is an ultrafilter, the sets $J_i := J \setminus A_i$ belong to $\mathscr{U}$. Hence, $C := J_1 \cap \dots \cap J_n = J \setminus (A_1 \cup \dots \cup A_n)$ belongs to $\mathscr{U}$. But $A_1 \cup \dots \cup A_n \subseteq \{j : f(j) \in [a, b]\} = J$ which implies $C = \varnothing$, a contradiction.

This contradiction shows that there is indeed some point $x \in [a, b]$ which is not contained in $\bigcup \mathscr{B}$, i.e. x is not contained in an open set which does not belong to $f(\mathscr{U})$. But this means that all open neighborhoods of x belong to $f(\mathscr{U})$, i.e. f converges to x with respect to $\mathscr{U}$. □

An inspection of the proof of Theorem 5.30 shows that it is actually not required that the image space of f be $\mathbb{R}$; in fact, it suffices that $f(J)$ is contained in a compact space.

The limit value depends only on the equivalence class of f:

Lemma 5.31. *If $\mathscr{F}$ is a filter and $f_1, f_2 : J \to X$ such that $f_1(j) = f_2(j)$ for almost all j, then f_1 converges to x if and only if f_2 converges to x.*

Proof. By assumption, $F := \{j : f_1(j) = f_2(j)\}$ belongs to $\mathscr{F}$. We claim that this implies $f_1(\mathscr{F}) = f_2(\mathscr{F})$ which of course implies the statement. Thus, given $A \in f_1(\mathscr{F})$, we have to prove that $A \in f_2(\mathscr{F})$ (the converse follows analogously). We have $A \supseteq f_1(F_0)$ for some $F_0 \in \mathscr{F}$ (Lemma 5.27). But then $A \supseteq f_1(F_0 \cap F) = f_2(F_0 \cap F)$ which in view of $F_0 \cap F \in \mathscr{F}$ implies that $A \in f_2(\mathscr{F})$, as claimed. □

Theorem 5.32. *Consider the map $*$ of the ultrapower model of Theorem 4.20. If $x = \varphi([f])$ is finite, then*

$$\operatorname{st}(x) = \lim_{j \to \mathscr{U}} f(j),$$

where the limit on the right-hand side exists and is independent of the particular choice of f.

Proof. First note that, since x is finite, we have $x = \varphi([f])$ for some bounded $f : J \to \mathbb{R}$ (Example 5.11) so that the limit $x_0 = \lim_{j \to \mathscr{U}} f(j)$ exists by Theorem 5.30. Moreover, by Lemma 5.31, the limit exists also (and has the same value) if we choose another representative f. Given $\varepsilon \in \mathbb{R}_+$, the open neighborhood $U := (x_0 - \varepsilon, x_0 + \varepsilon)$ belongs to $f(\mathscr{U})$. By Lemma 5.27, this means that $\{j : f(j) \in U\}$ belongs to $\mathscr{U}$. But this means $|f(j) - x_0| < \varepsilon$ almost everywhere, and so $x \approx {}^*x_0$ in view of Example 5.6. We thus must have $x_0 = \operatorname{st}(x)$. $\square$

Proposition 5.25 is a special case of Theorem 5.32: Indeed, if $J = \mathbb{N}$ and $\lim_{j\to\infty} f(j) = x_0$ exists, then $\lim_{j\to\mathscr{F}} f(j) = x_0$ for the filter of Example 4.2 by Proposition 5.28, and so $\lim_{j\to\mathscr{U}} f(j) = x_0$ by Proposition 5.29 (observe that $\mathscr{F} \subseteq \mathscr{U}$ by Exercise 11, since $\mathscr{U}$ is δ-incomplete and thus free).

§6 The Permanence Principle and *-finite Sets

One of the most important principles in nonstandard analysis is the so-called *permanence principle* which is a simple consequence of the fact that ${}^\sigma\mathbb{N}$ is external and that all elements in $\mathbb{N}_\infty$ are infinite:

Theorem 6.1 (Permanence Principle)**.** *Let $\alpha(\underline{n})$ be an internal predicate with $\underline{n}$ as its only free variable.*

1. *If $\alpha(n)$ holds for all sufficiently large finite $n \in {}^\sigma\mathbb{N}$, $n \geq n_0$, then there is some infinite $h \in \mathbb{N}_\infty$ such that $\alpha(n)$ holds for all $n \in {}^*\mathbb{N}$ with $n_0 \leq n \leq h$. In particular, $\alpha(h)$ holds for some infinite $h \in \mathbb{N}_\infty$.*
2. *If $\alpha(h)$ holds for all infinite $h \in {}^*\mathbb{N}_\infty$, then there is some $n_0 \in {}^\sigma\mathbb{N}$ such that $\alpha(n)$ holds for all $n \in {}^*\mathbb{N}$ with $n \geq n_0$. In particular, $\alpha(n_0)$ holds for some finite $n_0 \in {}^\sigma\mathbb{N}$.*

Proof. 1. Let

$$M := \{\underline{n} \in {}^*\mathbb{N} \mid \forall \underline{y} \in {}^*\mathbb{N} : (n_0 \leq \underline{y} \leq \underline{n} \implies \alpha(\underline{y}))\}.$$

Then M is internal by the internal definition principle, and the assumption implies that any $n \in {}^\sigma\mathbb{N}$ belongs to M: Indeed, any $n_1 \in {}^*\mathbb{N}$ with $n_1 \leq n$ is finite and thus belongs to ${}^\sigma\mathbb{N}$ by Proposition 5.9. Hence, ${}^\sigma\mathbb{N} \subseteq M$. Since M is internal by the internal definition principle and ${}^\sigma\mathbb{N}$ is external, we have $M \neq {}^\sigma\mathbb{N}$. Hence, there is some $h \in M \setminus {}^\sigma\mathbb{N}$, which thus has the required properties.
2. Let

$$M := \{\underline{n} \in {}^*\mathbb{N} \mid \forall \underline{y} \in {}^*\mathbb{N} : (\underline{y} \geq \underline{n} \implies \alpha(\underline{y}))\}.$$

If $h \in \mathbb{N}_\infty$ is infinite (Proposition 5.9), then no $h_1 \in {}^*\mathbb{N}$ with $h_1 \geq h$ belongs to ${}^\sigma\mathbb{N}$, and so the assumption implies $h \in M$. Hence, $\mathbb{N}_\infty \subseteq M$. Since ${}^\sigma\mathbb{N}$ is external, also $\mathbb{N}_\infty$ is external (Theorem 3.19), and so $\mathbb{N}_\infty \neq M$. Hence, there is some $n \in M \setminus \mathbb{N}_\infty = M \cap {}^\sigma\mathbb{N}$. □

The second part of the following consequence is also called the *Cauchy principle*. The name "Cauchy principle" is due to the fact that it allows us to formulate properties which hold for infinitesimals (which have been used by Leibniz) in an ε-δ-type manner as was first propagated by Cauchy (and which is the only reasonable definition in standard analysis).

Corollary 6.2 (Permanence Principle for ${}^*\mathbb{R}$)**.** *Let $\alpha(\underline{\varepsilon})$ be an internal predicate with $\underline{\varepsilon}$ as its only free variable.*

1. *If $\alpha(\varepsilon)$ holds for all sufficiently small standard $\varepsilon \in {}^\sigma\mathbb{R}_+$, $\varepsilon < \varepsilon_0$, then $\alpha(c)$ holds also for some infinitesimal $c \in \inf({}^*\mathbb{R})$, $c > 0$.*

2. *If $\alpha(c)$ holds for all infinitesimals $c \in \inf({}^*\mathbb{R})$, $c > 0$, then there is some standard $\varepsilon_0 \in {}^\sigma\mathbb{R}$, such that $\alpha(c)$ holds for all standard or nonstandard $c \in {}^*\mathbb{R}$ with $0 < c \le \varepsilon_0$.*

   Moreover, if $\alpha(c)$ holds for all infinitesimals $c \in \inf({}^*\mathbb{R})$, then there is some standard $\varepsilon_0 \in {}^\sigma\mathbb{R}$, such that $\alpha(c)$ holds for all standard or nonstandard $c \in {}^*\mathbb{R}$ with $|c| \le \varepsilon_0$.
3. *If $\alpha(x)$ holds for all sufficiently large standard $x \in {}^\sigma\mathbb{R}$, $x > x_0$, then $\alpha(c)$ holds also for some infinite $c > 0$.*
4. *If $\alpha(c)$ holds for all infinite $c > 0$, then there is some finite standard $x_0 \in {}^\sigma\mathbb{R}$ such that $\alpha(x)$ holds for all standard or nonstandard $x \in {}^\sigma\mathbb{R}$, $x \ge x_0$.*

Proof. 1. Since $\alpha(1/n)$ holds for all sufficiently large n, we have $\alpha(1/h)$ for some $h \in {}^\sigma\mathbb{N}$ by Theorem 6.1. By Proposition 5.15, we have $c := 1/h \in \inf({}^*\mathbb{R})$.
2. Let $\beta(\underline{\varepsilon})$ denote the internal predicate

$$\forall \underline{y} \in {}^*\mathbb{R} : (0 < \underline{y} \le 1/\underline{\varepsilon} \implies \alpha(\underline{y})).$$

Then $\beta(h)$ holds for all infinite $h \in \mathbb{N}_\infty$, and Theorem 6.1 thus implies that $\beta(n)$ holds for some finite $n \in {}^\sigma\mathbb{N}$. Now the claim follows with $\varepsilon_0 := 1/n$. The second part follows analogously by the predicate

$$\forall \underline{y} \in {}^*\mathbb{R} : (|\underline{y}| \le 1/\underline{\varepsilon} \implies \alpha(\underline{y})).$$

The remaining claims follow by applying the above proved statements for the predicate $\alpha(1/\underline{\varepsilon})$ in place of α; recall that $c \in \inf({}^*\mathbb{R})$ if and only if $1/c$ is infinite (Proposition 5.23). □

Exercise 21. Prove the following generalizations of the permanence principle: Let $\alpha(\underline{x})$ be an internal predicate with $\underline{x}$ as its only free variable.

1. If there is some $h_0 \in {}^*\mathbb{N}_\infty$ such that $\alpha(h)$ holds for all $h \in {}^*\mathbb{N}_\infty$ with $h < h_0$, then there is some $n_0 \in {}^\sigma\mathbb{N}$ such that $\alpha(n)$ holds for all $n \in {}^\sigma\mathbb{N}$ with $n \ge n_0$.
2. If there is some infinitesimal $c \in \inf({}^*\mathbb{R})$, $c > 0$, such that $\alpha(d)$ holds for all infinitesimals $d \in \inf({}^*\mathbb{R})$ with $d > c$, then there is some standard $\varepsilon_0 \in {}^\sigma\mathbb{R}_+$ such that $\alpha(\varepsilon)$ holds for all standard or nonstandard $\varepsilon \in {}^*\mathbb{R}$ with $c < \varepsilon \le \varepsilon_0$.

Exercise 22. Prove *Robinson's sequential lemma*: If $x : {}^*\mathbb{N} \to {}^*\mathbb{R}$ is an internal sequence such that $x_n \approx 0$ for all $n \in {}^\sigma\mathbb{N}$, then there is some $h \in \mathbb{N}_\infty$ such that $x_n \approx 0$ for all $n \in {}^*\mathbb{N}$ with $n \le h$.

As a sample application of the permanence principle, let us give a simpler proof of Theorem 5.24:

Example 6.3. $\inf({}^*\mathbb{R})$ is external. Indeed, if $\inf({}^*\mathbb{R})$ were internal, then the predicate $\alpha(\underline{x}) \equiv \forall \underline{y} \in \inf({}^*\mathbb{R}) : \underline{y} < \underline{x}$ were internal. Since $\alpha(\varepsilon)$ holds for all standard numbers $\varepsilon > 0$, the permanence principle implies that we have $\alpha(c)$ for some infinitesimal $c \in \inf({}^*\mathbb{R})$, a contradiction.

Note, however, that we *used* the fact that $\mathbb{N}_\infty$ is external for the proof of the permanence principle. Thus, in a sense, the permanence principle is equivalent to the fact that certain entities are external. This is not accidental: Many deep results of nonstandard analysis depend on the fact that certain entities are external.

All nonstandard phenomena we observed so far are based on the fact that ${}^*\mathbb{N} \neq {}^\sigma\mathbb{N}$. The transfer principle implies, roughly speaking, that ${}^*\mathbb{N}$ plays in the nonstandard universe the same role as the set $\mathbb{N}$ plays in the standard universe. But then the word "finite" should be interpreted differently in the nonstandard universe:

Definition 6.4. A set A is called *finite*, if it is in a one-to-one correspondence with a set $\{1, \ldots, n\}$ of natural numbers.

A set A is called *Dedekind finite*, if it is not in a one-to-one correspondence with a proper subset $A_0 \subsetneq A$.

We recall that a countable form of the axiom of choice implies the following well-known result (of the standard world). For the reader unfamiliar with such results, we provide a (standard) proof:

Proposition 6.5. *A set A is finite if and only if it is Dedekind finite.*

Proof. If A is finite and $A_0 \subsetneq A$, then the cardinality of A_0 is strictly smaller than that of A. Since bijections preserve the cardinality, there is no bijection $f : A \to A_0$. Conversely, if A is infinite, we may define inductively an injection $f : \mathbb{N} \to A$ in the following way: Choose $f(1) \in A$ arbitrary, and if $f(1), \ldots, f(n)$ are already defined, choose $f(n+1) \in A$ such that $f(n+1) \notin \{f(1), \ldots, f(n)\}$. Such a value exists, since otherwise we have a bijection witnessing that A is finite. Then we may define a bijection $F : A \to A_0$ where $A_0 = A \setminus \{f(1)\}$ by

$$F(x) = \begin{cases} x & \text{if } x \notin f(\mathbb{N}), \\ f(n+1) & \text{if } x = f(n). \end{cases}$$

Hence, A is not Dedekind finite. □

In the nonstandard world, we have now:

Definition 6.6. An entity $A \in \widehat{{}^*S}$ is called **-finite* or *hyperfinite* if there is some $h \in {}^*\mathbb{N}$ and an internal bijection $f : \{1, \ldots, h\} \to A$. In this case, we define ${}^\# A$ as the hyperfinite number h; otherwise, we put ${}^\# A := \infty$.

Here and in the following, $\{1, \ldots, h\}$ is used as a shortcut for $\{\underline{x} \in {}^*\mathbb{N} : \underline{x} \leq h\}$ if $h \in {}^*\mathbb{N}$. Although this notation is intuitive, the reader should take care:

For $h \in \mathbb{N}_\infty$ the set $\{1, \dots, h\}$ is even uncountable (Theorem 3.23). Moreover, this set is not well-ordered: The set $\{1, \dots, h\} \cap \mathbb{N}_\infty$ has no smallest element by Theorem 5.12 (but *internal* nonempty subsets of $\{1, \dots, h\}$ have a smallest element by Theorem 5.12).

Note that if A is hyperfinite, then $A = \operatorname{rng} f$ is internal by Theorem 3.19.

The transfer principle implies that a standard entity ${}^*A$ is *-finite if and only if A is finite (see Exercise 27 below). However, the crucial point in the above definition is that this definition applies also for *nonstandard* (internal) sets.

For the rest of this section, we discuss the above notion. The following results all sound rather natural. However, the proofs are surprisingly technical. The reason is that sentences about internal sets cannot easily be formulated such that the transfer principle can be applied, i.e. such that all constants are *standard* objects (and not only *internal* objects): To do this, we always have to formulate the sentences as sentences about objects which contain the given internal sets as elements. In other words: We must consider objects of a higher type. It may be a good idea if the reader works through the appendix parallel to this section. Alternatively, the reader may also want to skip the proofs of this section at the first reading and to consider the proofs after more experience.

We have to prove that ${}^\# A$ is well-defined. To do so, we first show the following lemma which we will need later:

Lemma 6.7. *Let $h_1, h_2 \in {}^*\mathbb{N}$, and $f : \{1, \dots, h_1\} \to \{1, \dots, h_2\}$ be internal. If f is onto, then $h_1 \geq h_2$. If f is one-to-one, then $h_1 \leq h_2$.*

Proof. By Exercise 8, we find an internal function $F : {}^*\mathbb{N} \to {}^*\mathbb{N}$ such that $F(h) = f(h)$ for any $h \in \{1, \dots, h_1\}$.

For the first statement, consider the sentence

$$\forall \underline{x} \in \mathbb{N}^{\mathbb{N}}, \underline{n}_1, \underline{n}_2 \in \mathbb{N} : \alpha(\underline{x}, \underline{n}_1, \underline{n}_2) \implies \underline{n}_1 \geq \underline{n}_2$$

where $\alpha(\underline{x}, \underline{n}_1, \underline{n}_2)$ is a shortcut for

$$\forall \underline{z}_2 \in \mathbb{N} : \underline{z}_2 \leq \underline{n}_2 \implies (\exists \underline{z}_1 \in \mathbb{N} : \underline{z}_1 \leq \underline{n}_1 \land (\underline{z}_1, \underline{z}_2) \in \underline{x})$$

which may more intuitively be written as $\{1, \dots, \underline{n}_2\} \subseteq \underline{x}(\{1, \dots, \underline{n}_1\})$. Thus, the sentence means that whenever $\{1, \dots, \underline{n}_2\} \subseteq \underline{x}(\{1, \dots, \underline{n}_1\})$ for some function $\underline{x} : \mathbb{N} \to \mathbb{N}$, we must have $\underline{n}_1 \geq \underline{n}_2$. This is evidently true. Hence, the $*$-transform of this sentence is true. Since ${}^*(\mathbb{N}^{\mathbb{N}})$ contains F by Theorem 3.21, we may conclude that the relation $\{1, \dots, h_2\} \subseteq F(\{1, \dots, h_1\})$ implies $h_1 \geq h_2$.

For the second statement, consider analogously the sentence

$$\forall \underline{x} \in \mathbb{N}^{\mathbb{N}}, \underline{n} \in \mathbb{N} : \text{“}\underline{x} \text{ one-to-one on } \{1, \dots, \underline{n}\}\text{”} \implies \exists \underline{m} \in \mathbb{N} : (\underline{m} \leq \underline{n} \land \underline{x}(\underline{m}) = \underline{n})$$

which is true since any injection $f : \{1, \ldots, n\} \to \mathbb{N}$ attains at least some value $m \geq n$. Hence, the $*$-transform of this sentence is true. Since ${}^*(\mathbb{N}^{\mathbb{N}})$ contains F, we may conclude that $F(\{1, \ldots, h_1\})$ must at least contain some value $h \geq h_1$. In view of $F(\{1, \ldots, h_1\}) = f(\{1, \ldots, h_1\}) = \{1, \ldots, h_2\}$, this implies $h_2 \geq h_1$. □

Proposition 6.8. *${}^\# A$ is well-defined.*

Proof. Let $f_1 : \{1, \ldots, h_1\} \to A$ and $f_2 : \{1, \ldots, h_2\} \to A$ be two internal bijections. By Theorem 3.19, also f_2^{-1} is internal, and by Exercise 7 also the composition $g = f_2^{-1} \circ f_1 = \{1, \ldots, h_1\} \to \{1, \ldots, h_2\}$. Since g is a bijection, Lemma 6.7 now implies $h_1 = h_2$ which means that ${}^\# A$ is well-defined. □

Of course, one expects from the term "*-finite" that all finite sets are *-finite:

Proposition 6.9. *If $A \in \widehat{{}^*S}$ is finite, then A is *-finite, and ${}^\# A$ is the number of elements.*

Proof. Let $a_1, \ldots, a_n$ be all elements of A. Put $f := \{(1, a_1), \ldots, (n, a_n)\}$. Then $f : \{1, \ldots, n\} \to A$ is the desired internal bijection. □

It is a natural question whether we are led to a different definition of *-finite sets, if we start with Dedekind finite sets. The answer is negative:

Call an internal entity $A \in \widehat{{}^*S}$ *Dedekind *-finite* if there is no internal bijection $f : A \to A_0$ where $A_0 \subsetneq A$. Observe that in this definition we require both A and f to be internal.

Theorem 6.10. *An internal entity A is *-finite if and only if it is Dedekind *-finite.*

Proof. Let α be the sentence

$$\exists \underline{x} \in A^A : \text{“}\underline{x} \text{ is one-to-one”} \wedge \exists \underline{y} \in A : \forall \underline{z} \in A : (\underline{z}, \underline{y}) \notin \underline{x},$$

i.e. α is true if and only if there is a bijection f of A onto a proper subset of A_0. Let β be the sentence

$$\exists \underline{x} \in A^{\mathbb{N}}, \underline{n} \in \mathbb{N} : \text{“}\underline{x} \text{ maps } \{1, \ldots, \underline{n}\} \text{ bijectively onto } A\text{”},$$

i.e. β is true if and only if A is finite. Then the sentence

$$(\neg\alpha) \iff \beta$$

is true by Proposition 6.5. Since ${}^*(A^A)$ and ${}^*(A^{\mathbb{N}})$ consist of all internal functions $f : {}^*A \to {}^*A$ resp. $f : {}^*\mathbb{N} \to {}^*A$, the $*$-transform of the above sentence means that ${}^*A$ is Dedekind *-finite if and only if there is an internal function $f : {}^*\mathbb{N} \to {}^*A$ which maps $\{1, \ldots, h\}$ onto ${}^*A$ for some $h \in {}^*\mathbb{N}$. Since any internal function $f : \{1, \ldots, h\} \to {}^*A$ may be extended to a function $f : {}^*\mathbb{N} \to {}^*A$ by Example 8, the claim follows. □

One might also choose the following characterization as the definition of *-finite sets:

Theorem 6.11. *An entity $A \in \widehat{^*S}$ is *-finite if and only if there is some entity $\mathscr{A} \in \widehat{S}$ whose elements are finite entities such that $A \in {^*\mathscr{A}}$.*

Moreover, if $\mathscr{A} \in \widehat{S}$ consists of finite entities and $f : \mathscr{A} \to \mathbb{N}$ is the mapping which associates to each $B \in \mathscr{A}$ its number of elements, then $^*f(A) = {^\#A}$ for each $A \in {^*\mathscr{A}}$.

It may be arranged that all elements of $\mathscr{A}$ have the same type as A.

Proof. Assume that $\mathscr{A} \in \widehat{S}$ is an entity whose elements are finite entities. Note that $U := \bigcup \mathscr{A} \in \widehat{S}$ by Theorem 2.1. Then the following sentence is true:

$$\forall \underline{x} \in \mathscr{A} : \exists \underline{y} \in U^{\mathbb{N}} : \exists \underline{n} \in \mathbb{N} : (\alpha(\underline{y}, \underline{n}, \underline{x}) \wedge \underline{n} = f(\underline{x}))$$

where $\alpha(\underline{y}, \underline{n}, \underline{x})$ is a shortcut for a transitively bounded sentence with the meaning "$\underline{y}$ maps $\{1, \ldots, \underline{n}\}$ bijectively onto $\underline{x}$": This sentence is a reformulation of the fact that each $B \in \mathscr{A}$ is finite and has $f(B)$ elements. The transfer of this sentence implies for any $A \in {^*\mathscr{A}}$ that there is some $y \in {^*(U^{\mathbb{N}})}$ (i.e. some internal function $y : {^*U} \to {^*\mathbb{N}}$ by Theorem 3.21) and some $z \in {^*\mathbb{N}}$ such that y maps $\{1, \ldots, z\}$ bijectively onto A and $z = {^*f(A)}$. Hence, any $A \in {^*\mathscr{A}}$ is *-finite, and $^\#A = z = {^*f(A)}$.

Conversely, let A be *-finite, i.e. there is some $h \in {^*\mathbb{N}}$ and some internal bijection $f : \{1, \ldots, h\} \to A$. Since A is an internal entity, we find by Corollary A.3 an entity $\mathscr{B} \in \widehat{S}$ which consists only of entities such that $A \in {^*\mathscr{B}}$ (we may even assume that all elements of $\mathscr{B}$ have the same type as A). Put $U := \bigcup \mathscr{B}$, and observe that $^*A \subseteq {^*U}$ by Theorem A.4. Let $\mathscr{A} \subseteq \mathscr{B}$ be the collection of all finite entities $B \in \mathscr{B}$. Then the sentence

$$\forall \underline{x} \in \mathscr{B} : ((\exists \underline{y} \in U^{\mathbb{N}}, \underline{n} \in \mathbb{N} : \alpha(\underline{y}, \underline{n}, \underline{x})) \implies \underline{x} \in \mathscr{A})$$

is true, where α is defined as before. The transfer of this sentence means that $^*\mathscr{A}$ contains all elements $A_0 \in {^*\mathscr{B}}$ for which we find an internal function $y : {^*\mathbb{N}} \to {^*U}$ (Theorem 3.21) and some $h \in {^*\mathbb{N}}$ such that y maps $\{1, \ldots, h\}$ bijectively onto A_0. But A is such an element: Indeed, since $^*A \subseteq {^*U}$, we may extend the given function f to an internal function $y : {^*\mathbb{N}} \to {^*U}$. Hence, we have $A \in {^*\mathscr{A}}$. □

If a set A is infinite, then there exists an injection $f : \mathbb{N} \to A$ (by a countable form of the axiom of choice). We have an analogue in the nonstandard world:

Theorem 6.12. *For each internal entity A precisely one of the following alternatives holds:*

1. *A is *-finite, or*

2. *There is an internal injection* $f : {}^*\mathbb{N} \to A$.

Proof. By Corollary A.3, we find some entity $\mathscr{A} \in \widehat{S}$ whose elements are all entities such that $A \in {}^*\mathscr{A}$. Let $U := \bigcup \mathscr{A}$. Then the transitively bounded sentence

$$\forall \underline{x} \in \mathscr{A} : ((\exists \underline{y} \in U^{\mathbb{N}} : \alpha(\underline{y}, \underline{x}, U, \mathbb{N})) \iff \neg(\exists \underline{y} \in U^{\mathbb{N}}, \underline{n} \in \mathbb{N} : \beta(\underline{y}, \underline{n}, \underline{x}, U, \mathbb{N})))$$

is true, where $\alpha(\underline{y}, \underline{x}, U, \mathbb{N})$ is a shortcut of a transitively bounded formula with the meaning "$\underline{y}$ maps $\mathbb{N}$ injectively into $\underline{x}$", and $\beta(\underline{y}, \underline{n}, \underline{x}, U, \mathbb{N})$ means similarly "$\underline{y}$ maps $\{1, \dots, \underline{n}\}$ bijectively onto $\underline{x}$". The transfer of the above sentence reads:

$$\forall \underline{x} \in {}^*\mathscr{A} : ((\exists \underline{y} \in {}^*(U^{\mathbb{N}}) : \alpha(\underline{y}, \underline{x}, {}^*U, {}^*\mathbb{N}))\dot{\vee}$$
$$(\exists \underline{y} \in {}^*(U^{\mathbb{N}}), \underline{n} \in {}^*\mathbb{N} : \beta(\underline{y}, \underline{n}, \underline{x}, {}^*U, {}^*\mathbb{N}))).$$

By Theorem 3.21, ${}^*(U^{\mathbb{N}})$ consists of all internal functions $f : {}^*\mathbb{N} \to {}^*U$. For the choice $\underline{x} = A \in {}^*\mathscr{A}$, the above sentence thus means in view of ${}^*A \subseteq {}^*U$ (because $A \subseteq U$): Precisely one of the following alternatives holds: Either there is an internal injection $f : {}^*\mathbb{N} \to A$, or there is some $h \in {}^*\mathbb{N}$ and an internal function $g : {}^*\mathbb{N} \to {}^*U$ such that $g : \{1, \dots, h\} \to A$ is bijective. Since the restriction of such an internal function g to $\{1, \dots, h\}$ is internal and since conversely any internal bijection $g : \{1, \dots, h\} \to A$ can be extended to an internal function $g : {}^*\mathbb{N} \to {}^*U$ (both follows from Exercise 8), the statement follows. □

Recall that we agreed to write ${}^\# A = \infty$ if A is not *-finite. In this connection, we define $\infty > h$ for any $h \in {}^*\mathbb{N}$.

Theorem 6.13. *Let* $A \in \widehat{{}^*S}$ *be an internal entity.*

1. *If* $B \subseteq A$ *is internal, then* ${}^\# B \leq {}^\# A$.
2. *If* $B \subsetneq A$ *is internal and* A *is* **-finite, then* ${}^\# B \leq {}^\# A - 1$.
3. *If* $f : A \to B$ *is an internal injection, then* ${}^\# A \leq {}^\# B$.
4. *If* $f : A \to B$ *is an internal surjection, then* ${}^\# A \geq {}^\# B$.

Proof. 1. If ${}^\# A = \infty$, there is nothing to prove. Thus, assume that A is *-finite. This implies that B is *-finite: Otherwise, Theorem 6.12 would imply that there is an internal injection $f : {}^*\mathbb{N} \to B$. But then f is also an internal injection from ${}^*\mathbb{N}$ into A which by Theorem 6.12 contradicts our assumption that A is *-finite.

Hence, we may assume that $h := {}^\# A$ and $k := {}^\# B$ both belong to ${}^*\mathbb{N}$, and that there are internal bijections $f : \{1, \dots, h\} \to A$ and $g : \{1, \dots, k\} \to B$. Then f^{-1} is internal by Theorem 3.19, and thus $G := f^{-1} \circ g : \{1, \dots, k\} \to \{1, \dots, h\}$ is an internal injection (Exercise 7). Lemma 6.7 implies $h \geq k$, i.e. ${}^\# A \geq {}^\# B$.
2. With $G : \{1, \dots, k\} \to \{1, \dots, h\}$ as in 1., the function G is not onto, i.e. there is some $j \in C := \{1, \dots, h\}$ with $j \notin \operatorname{rng} G$. Observe that the sets $C_1 := \{x \in C : G(x) < j\}$ and $C_2 := \{x \in C : G(x) > j\}$ are internal by the internal definition

principle. Now define a function $F : C \to \{1, \ldots, h-1\}$ by putting $F(x) := G(x)$ for $x \in C_1$ and $F(x) := G(x) - 1$ for $x \in C_2$. In view of Exercise 8, the function F is internal. Since $F : \{1, \ldots, k\} \to \{1, \ldots, h-1\}$ is one-to-one, Lemma 6.7 implies $k \le h - 1$, i.e. ${}^{\#}B \le {}^{\#}A - 1$.

3. In case ${}^{\#}B = \infty$, there is nothing to prove. Thus, assume ${}^{\#}B < \infty$. Since $B_0 := \operatorname{rng}(f)$ is an internal subset of B, we find by 1. that $h := {}^{\#}B_0 \le {}^{\#}B$. Hence, there is an internal bijection $g : \{1, \ldots, h\} \to B_0$. The bijection $f^{-1} \circ g : \{1, \ldots, h\} \to A$ is internal by Exercise 7 and Theorem 3.19. Hence, ${}^{\#}A = h \le {}^{\#}B$.

4. In case ${}^{\#}A = \infty$, there is nothing to prove. Thus, assume $h := {}^{\#}A < \infty$, i.e. there is an internal bijection $g : \{1, \ldots, h\} \to A$. Assume first that B is *-finite. Then $k := {}^{\#}B < \infty$, and there is an internal bijection $g_1 : \{1, \ldots, k\} \to B$. Since $g_1^{-1} \circ f \circ g : \{1, \ldots, h\} \to \{1, \ldots, k\}$ is an internal surjection, Lemma 6.7 implies $h \ge k$, i.e. ${}^{\#}A \ge {}^{\#}B$.

To see that B must be *-finite, we apply Theorem 6.11: There is some $\mathscr{A} \in \widehat{S}$ whose elements are finite entities such that $A \in {}^*\mathscr{A}$. Since B is internal by Theorem 3.19, we find some $\mathscr{C} \in \widehat{S}$ such that $B \in {}^*\mathscr{C}$ and such that $\mathscr{C}$ contains only entities (Corollary A.3). Let $\mathscr{B} \subseteq \mathscr{C}$ be the collection of all finite entities of $\mathscr{C}$. Putting $U := \bigcup \mathscr{A}$, $V := \bigcup \mathscr{C}$, we have $U, V \in \widehat{S}$ by Theorem 2.1. Let $\mathscr{F}$ denote the system of all functions f with $\operatorname{dom}(f) \subseteq U$ and $\operatorname{rng}(f) \subseteq V$. Since functions map finite sets into finite sets, we have

$$\forall \underline{x} \in \mathscr{F} : (\exists \underline{y} \in \mathscr{A} : \operatorname{dom}(\underline{x}) \subseteq \underline{y}) \implies (\exists \underline{z} \in \mathscr{B} : \operatorname{rng}(\underline{x}) = \underline{z}).$$

The reader should take care that the shortcuts $\operatorname{dom}(\underline{x})$ and $\operatorname{rng}(\underline{x})$ used here are not transitively bounded, but we may take the quantifiers over the sets U and V. Since ${}^*\mathscr{F}$ consists of all internal functions $\underline{x}$ with $\operatorname{dom}(\underline{x}) \subseteq {}^*U$ and $\operatorname{rng}(\underline{x}) \subseteq {}^*V$ (Exercise 83) and thus in particular $f \in {}^*\mathscr{F}$, we conclude from the $*$-transform of the above sentence that $B = \operatorname{rng}(f) \in {}^*\mathscr{B}$ (concerning U^* and V^* observe Theorem A.4). Applying the converse direction of Theorem 6.11, this shows that B is indeed *-finite, as claimed. □

Exercise 23. (Difficult). Prove the following generalization of Theorem 6.13:

1. If $f : A \to B$ is an internal surjection, then there exists an internal injection $g : B \to A$.
2. If there exist internal injections $f_1 : A \to B$ and $f_2 : B \to A$, then there exists an internal bijection $g : A \to B$.

Hint: Use that the above facts are right in the standard world, i.e. without the term "internal". In the second case, this fact (in the standard world) can be found in literature on set theory under the name *Schröder-Bernstein theorem.*

Theorem 6.14. *If A and B are *-finite, then $A \cup B$ and $A \times B$ are *-finite and satisfy*

$$^{\#}(A \cup B) = {}^{\#}A + {}^{\#}B \quad \text{if } A \cap B = \varnothing, \tag{6.1}$$

$$^{\#}(A \times B) = {}^{\#}A \cdot {}^{\#}B. \tag{6.2}$$

Also the system P_A of all internal subsets of A is *-finite, and

$$^{\#}P_A = 2^{^{\#}A}.$$

Proof. By Theorem 6.11, there are entities $\mathscr{A}, \mathscr{B} \in \widehat{S}$ consisting of finite entities such that $A \in {}^*\mathscr{A}$, $B \in {}^*\mathscr{B}$. Let $\mathscr{C}, \mathscr{D}, \mathscr{E} \in \widehat{S}$ denote the collection of all sets of the form $X \times Y$, $X \cup Y$ resp. $\mathcal{P}(X)$, with $X \in \mathscr{A}$, $Y \in \mathscr{B}$. Note that

$$A \times B \in {}^*\mathscr{C}, \quad A \cup B \in {}^*\mathscr{D}, \quad P_A \in {}^*\mathscr{E}, \tag{6.3}$$

as follows from Exercise 84, Corollary A.5, respectively Theorem A.8.

Put $C := \bigcup \mathscr{C}$, $D := \bigcup \mathscr{D}$, and $E := \bigcup \mathscr{E}$. Let $f_1 : \mathscr{A} \to \mathbb{N}$ and $f_2 : \mathscr{B} \to \mathbb{N}$ be the functions which associate to each set the number of elements. Then we have

$$\forall \underline{x} \in \mathscr{A}, \underline{y} \in \mathscr{B}, \underline{z} \in \mathscr{C} : \underline{z} = \underline{x} \times \underline{y} \implies {}^{\#}\underline{z} = f_1(\underline{x}) f_2(\underline{y})$$

where the last expression is a shortcut for

$$\exists \underline{w} \in C^{\mathbb{N}} : \text{“}\underline{w} \text{ maps } \{1, \ldots, f_1(\underline{x}) f_2(\underline{y})\} \text{ bijectively onto } \underline{z}\text{”}.$$

The transfer of this sentence implies in view of (6.3) that there is some $w \in {}^*(C^{\mathbb{N}})$ (i.e. some internal function $w : {}^*\mathbb{N} \to {}^*C$ by Theorem 3.21) such that $w : \{1, \ldots, {}^*f_1(A) \cdot {}^*f_2(B)\} \to z = A \times B$ is bijective. Since ${}^*f_1(A) = {}^{\#}A$ and ${}^*f_2(B) = {}^{\#}B$ by Theorem 6.11, this proves (6.2).

Analogously, the transfer of the sentence

$$\forall \underline{x} \in \mathscr{A}, \underline{y} \in \mathscr{B}, \underline{z} \in \mathscr{D} : \underline{z} = \underline{x} \cup \underline{y} \implies \alpha$$

where α is a shortcut for

$$\exists \underline{w} \in D^{\mathbb{N}}, \underline{v} \in \mathbb{N} : \text{“}\underline{w} \text{ maps } \{1, \ldots, \underline{v})\} \text{ bijectively onto } \underline{z}\text{”}$$

implies in view of (6.3) that $A \cup B$ is *-finite. Similarly, the transfer of the sentence

$$\forall \underline{x} \in \mathscr{A}, \underline{y} \in \mathscr{B}, \underline{z} \in \mathscr{D} : (\underline{z} = \underline{x} \cup \underline{y} \wedge \underline{x} \cap \underline{y} = \varnothing) \implies {}^{\#}vz = f_1(\underline{x}) + f_2(\underline{y})$$

where the last expression is a shortcut for

$$\exists \underline{w} \in D^{\mathbb{N}} : \text{“}\underline{w} \text{ maps } \{1, \ldots, f_1(\underline{x}) + f_2(\underline{y})\} \text{ bijectively onto } \underline{z}\text{”}$$

proves (6.1).

For the last statement, consider the sentence

$$\forall \underline{x} \in \mathscr{E}, \underline{y} \in \mathscr{A} : (\alpha(\underline{x}, \underline{y}) \implies \beta(\underline{x}, \underline{y}, E))$$

where $\alpha(\underline{x}, \underline{y})$ and $\beta(\underline{x}, \underline{y}, E)$ are shortcuts for

$$\forall \underline{z} \in \underline{x} : \underline{z} \subseteq \underline{y}$$

and

$$\exists \underline{w} \in E^{\mathbb{N}} : \text{“}\underline{w} \text{ maps } \{1, \ldots, 2^{f_1(\underline{y})}\} \text{ bijectively onto } \underline{x}\text{”}.$$

Note that for $x \in \mathscr{E}$ and $y \in \mathscr{A}$ the statement $\alpha(x, y)$ can be interpreted as $x = \mathcal{P}(y)$; now apply the $*$-transform of the above sentence with $\underline{x} = P_A \in {}^*\mathscr{E}$ and $\underline{y} = A \in {}^*\mathscr{A}$. □

Exercise 24. Prove that for any *-finite entities A, B the formula

$${}^{\#}(A \cup B) = {}^{\#}A + {}^{\#}B - {}^{\#}(A \cap B)$$

holds.

Given some *-finite sequence $x : \{1, \ldots, h\} \to X$, we denote by $\#_X(x)$ the number h.

Proposition 6.15. *Given some internal entity X, the system*

$$X^{<{}^*\mathbb{N}} = \{\underline{x} \mid \exists \underline{n} \in {}^*\mathbb{N} : \underline{x} : \{1, \ldots, \underline{n}\} \to X \textit{ is internal}\}$$

is internal. Moreover, the function $\#_X(\cdot)$ is internal.

Proof. By Exercise 82, the set $\mathscr{F}$ of all internal functions x with $\operatorname{dom}(x) \subseteq {}^*\mathbb{N}$ and $\operatorname{rng}(x) \subseteq X$ is internal. Hence,

$$X^{<{}^*\mathbb{N}} = \{\underline{x} \in \mathscr{F} \mid \exists \underline{n} \in {}^*\mathbb{N} : \underline{x} : \{1, \ldots, \underline{n}\} \to X \text{ is internal}\}$$

is internal by the internal definition principle. Similarly,

$$\varphi := \{\underline{z} \mid \exists \underline{x} \in \mathscr{F}, \underline{n} \in {}^*\mathbb{N} : (\underline{x} : \{1, \ldots, \underline{n}\} \to X \text{ is internal} \wedge \underline{z} = (\underline{x}, \underline{n}))\}$$

is internal; but $\varphi = \#_X(\cdot)$. □

We use the notation $\#_X(x)$ also for sequences in the standard world, i.e. if $x : \{1, \ldots, n\} \to X$, then $\#_X(x) = n$.

Exercise 25. Let $X^{<\mathbb{N}}$ denote the set of all functions of the form $x : \{1, \ldots, n\} \to X$ where $n \in \mathbb{N}$ may depend on f. Prove that ${}^*(X^{<\mathbb{N}}) = {}^*X^{<{}^*\mathbb{N}}$ and ${}^*(\#_X(\cdot)) = \#_{{}^*X}(\cdot)$.

Let $\sum : \mathbb{R}^{<\mathbb{N}} \to \mathbb{R}$ be the mapping which associates to each finite sequence its sum. Then ${}^*\sum : {}^*(\mathbb{R}^{<\mathbb{N}}) \to {}^*\mathbb{R}$. For any *-finite (internal) sequence $x \in {}^*\mathbb{R}^{<{}^*\mathbb{N}}$, we define

$$\sum_{n=1}^{h} x(n) := {}^*\sum(x).$$

We use the more intuitive notation $x_1, \ldots, x_h$ with $x_n = x(n)$ in place of x.

Corollary 6.16. *If x and y are *-finite sequences of length h in ${}^*\mathbb{R}$ with $x_n \le y_n$, then*

$$\sum_{n=1}^{h} x_n \le \sum_{n=1}^{h} y_n.$$

Similarly, $|x_n| \le y_n$ implies

$$\left|\sum_{n=1}^{h} x_n\right| \le \sum_{n=1}^{h} y_n.$$

Moreover,

$$\sum_{n=1}^{h} 1 = h.$$

Proof. The first statement follows in view of the previous results by the transfer of

$$\forall \underline{x}, \underline{y} \in \mathbb{R}^{<\mathbb{N}} : (\forall \underline{n} \in \mathbb{N} : (\underline{n} \le \#_{\mathbb{R}}(\underline{x}) \implies \underline{x}(\underline{n}) \le \underline{y}(\underline{n}))) \implies \sum(\underline{x}) \le \sum(\underline{y}).$$

The proof of the other statements is similar. □

Exercise 26. (Difficult). Let R be a totally ordered internal entity (the order relation also being internal), and $\mathscr{A}$ be an internal system of *-finite subsets of R. Prove that the function $\max : \mathscr{A} \to R$ (with the obvious meaning) is well-defined and internal.

Exercise 27. Let $\mathscr{A} \in \widehat{S}$ be a nonempty entity which contains no atoms, and $c : \mathscr{A} \to \mathbb{N} \cup \{\infty\}$ be the function which associates to each $A \in \mathscr{A}$ the number of its elements. We write $\infty := {}^*\infty$. Prove that ${}^*c : {}^*\mathscr{A} \to {}^*\mathbb{N} \cup \{\infty\}$ satisfies ${}^*c(B) = {}^{\#}B$ for each $B \in {}^*\mathscr{A}$. Moreover, prove that a standard entity $B = {}^*A$ is *-finite if and only if A is finite.

§7 Calculus

The basic calculus is very easily described by nonstandard methods. As remarked in Section 1.1, this was historically one of the main motivations of nonstandard analysis.

However, the use of nonstandard analysis has the drawback that even the simplest results make use of the axiom of choice: Recall that without the axiom of choice (more precisely: Without the existence of δ-free ultrafilters) we were not able to construct nonstandard embeddings. We will see later that this restriction is essential: Indeed, with nonstandard methods one can "construct" so-called Hahn-Banach limits and also nonmeasurable functions, as we will see; without the axiom of choice it is for fundamental reasons not possible to prove the existence of such objects.

The above observation is an essential disadvantage since this means in the author's opinion that nonstandard analysis is not a good model for "real-world" phenomena.

On the other hand, if one is particularly interested in such objects whose existence can only be proved by the axiom of choice, nonstandard analysis is a much more convenient tool than classical analysis. We will see this later in particular in our discussion of Hahn-Banach limits.

7.1 Sequences

We first discuss real sequences. Recall that a sequence is a mapping $x : \mathbb{N} \to \mathbb{R}$; as usual, we write x_n instead of $x(n)$. The essential point of nonstandard analysis here is that x, as a mapping, has a $*$-transform ${}^*x : {}^*\mathbb{N} \to \mathbb{R}$. We will also write ${}^*x_n$ in place of ${}^*x(n)$. For $n \in \mathbb{N}$, we have ${}^*x_{{}^*n} = {}^*(x_n)$ (Theorem 3.13), i.e. the sequence ${}^*x_n$ may be identified with the sequence x_n for standard numbers. However, for $h \in \mathbb{N}_\infty$, we get additional values of ${}^*x_h$ on "infinite" places. One will suspect that these values have something to do with the limit of the sequence x_n. This is indeed the case:

Theorem 7.1. *Let x_n be a real sequence. Then we have for $x \in \mathbb{R}$:*

1. *$x_n \to x$ if and only if ${}^*x_h \approx {}^*x$ for each infinite $h \in \mathbb{N}_\infty$.*
2. *x_n has the accumulation point x if and only if ${}^*x_h \approx {}^*x$ for some infinite $h \in \mathbb{N}_\infty$.*

Proof. 1. If $x_n \to x$, then for any $\varepsilon \in \mathbb{R}_+$, we have

$$\forall \underline{n} \in \mathbb{N} : \underline{n} \geq n_0 \implies |x_{\underline{n}} - x| < \varepsilon$$

for some $n_0 \in \mathbb{N}$. The transfer principle implies

$$\forall \underline{n} \in {}^*\mathbb{N} : \underline{n} \geq {}^*n_0 \implies |{}^*x_{\underline{n}} - {}^*x| < {}^*\varepsilon.$$

In particular, $|{}^*x_h - {}^*x| \leq {}^*\varepsilon$ for any $h \in \mathbb{N}_\infty$. Since the latter holds for any $\varepsilon \in \mathbb{R}_+$, we have ${}^*x_h \approx {}^*x$.

Conversely, if ${}^*x \approx {}^*x_h$ for each infinite $h \in \mathbb{N}_\infty$, then for any $\varepsilon \in \mathbb{R}_+$ the internal formula $|{}^*x_{\underline{n}} - {}^*x| < {}^*\varepsilon$ holds true for each infinite $\underline{n} \in \mathbb{N}_\infty$. The permanence principle implies that there is some ${}^*n_0 \in {}^\sigma\mathbb{N}$ such that $|{}^*x_{\underline{n}} - {}^*x| < {}^*\varepsilon$ holds for all $\underline{n} \in {}^*\mathbb{N}$ with $\underline{n} \geq {}^*n_0$, i.e.

$$\forall \underline{n} \in {}^*\mathbb{N} : \underline{n} \geq {}^*n_0 \implies |{}^*x_{\underline{n}} - {}^*x| < {}^*\varepsilon.$$

The reverse form of the transfer principle implies that $|x_n - x| < \varepsilon$ for all $n \in \mathbb{N}$ with $n \geq n_0$, and so $x_n \to x$.
2. If x_n has the accumulation point x, then the transfer principle immediately shows that

$$\forall \underline{\varepsilon} \in {}^*\mathbb{R}_+ : \forall \underline{n} \in {}^*\mathbb{N} : \exists \underline{m} \in {}^*\mathbb{N} : \underline{m} \geq \underline{n} \wedge |{}^*x_{\underline{m}} - {}^*x| < \underline{\varepsilon}.$$

Choosing $\underline{\varepsilon} \in \inf({}^*\mathbb{R})$ and $\underline{n} \in \mathbb{N}_\infty$, we thus find some $m \in \mathbb{N}_\infty$ with ${}^*x_m \approx x$.

Conversely, if ${}^*x \approx {}^*x_h$ for some infinite $h \in \mathbb{N}_\infty$, then we have for each $\varepsilon \in \mathbb{R}_+$ and each $n_0 \in \mathbb{N}$ that

$$\exists \underline{n} \in {}^*\mathbb{N} : (\underline{n} \geq {}^*n_0 \wedge |{}^*x_{\underline{n}} - {}^*x| < {}^*\varepsilon).$$

Applying the converse direction of the transfer principle, we find

$$\exists \underline{n} \in \mathbb{N} : (\underline{n} \geq n_0 \wedge |x_{\underline{n}} - x| < \varepsilon).$$

Since n_0 and ε were arbitrary, x is an accumulation point of x_n. □

Also boundedness is easily characterized:

Theorem 7.2. *Let x_n be a real sequence.*

1. *x_n is bounded if and only if ${}^*x_h$ is finite for each $h \in {}^*\mathbb{N}$ (or, equivalently, for each $h \in \mathbb{N}_\infty$).*
2. *$x_n \to \pm\infty$ if and only if ${}^*x_h$ is infinite and positive/negative for each infinite $h \in \mathbb{N}_\infty$.*

Proof. 1. If x_n is bounded, say $|x_n| \leq c \in \mathbb{R}$, then we have by the transfer principle

$$\forall \underline{n} \in {}^*\mathbb{N} : |{}^*x_{\underline{n}}| \leq {}^*c,$$

i.e. ${}^*x_h$ is finite for each $h \in {}^*\mathbb{N}$.

Conversely, if $^*x_h$ is finite for each $h \in {}^*\mathbb{N}$, then the internal predicate

$$\forall \underline{n} \in {}^*\mathbb{N} : |{}^*x_{\underline{n}}| \leq \underline{m}$$

holds for each infinite $\underline{m} \in \mathbb{N}_\infty$. The permanence principle implies that it also holds for some finite $\underline{m} = {}^*m \in {}^\sigma\mathbb{N}$. An application of the converse direction of the transfer principle implies that x_n is bounded by m.
2. If $x_n \to \infty$, then we have for any $N \in \mathbb{N}$ that there is some $n_0 \in \mathbb{N}$ such that (using the transfer principle)

$$\forall \underline{n} \in {}^*\mathbb{N} : \underline{n} \geq {}^*n_0 \implies x_{\underline{n}} > {}^*N.$$

In particular, $x_h > {}^*N$ for each infinite $h \in \mathbb{N}_\infty$. Since $x_h > {}^*N$ for any $N \in \mathbb{N}$, this means that x_h is infinite.

Conversely, if x_h is infinite and positive for any $h \in \mathbb{N}_\infty$, then for any $N \in \mathbb{N}$ the internal formula $^*x_{\underline{n}} > N$ holds true for each infinite $\underline{n} \in \mathbb{N}_\infty$. The permanence principle implies that there is some $n_0 \in {}^\sigma\mathbb{N}$ such that $x_{\underline{n}} > N$ holds for all $\underline{n} \in {}^*\mathbb{N}$ with $\underline{n} \geq n_0$; in particular $x_n > N$ for all sufficiently large $n \in \mathbb{N}$. □

It may be slightly astonishing to the reader that there is some relation between boundedness and finiteness. However, if one thinks of $^*x_h$ for infinite $h \in \mathbb{N}_\infty$ as "generalized accumulation points", this is not surprising. This interpretation indeed makes sense:

Corollary 7.3. *Let x_n be a real sequence. Then its set of accumulation points is*

$$\{\mathrm{st}({}^*x_h) : h \in \mathbb{N}_\infty, {}^*x_h \text{ finite}\}.$$

Proof. If $^*x_h$ is finite, then $\mathrm{st}({}^*x_h) \approx {}^*x_h$ and $\mathrm{st}({}^*x_h) \in \mathbb{R}$, and so Theorem 7.1 implies that $\mathrm{st}({}^*x_h)$ is an accumulation point of x_n. Conversely, if x is an accumulation point of x_n, then Theorem 7.1 implies $x \approx x_h \approx \mathrm{st}({}^*x_h)$ for some $h \in \mathbb{N}_\infty$. □

Now we can give a simple nonstandard proof for a standard fact:

Corollary 7.4. *If a bounded real sequence has at most one accumulation point, then it converges.*

Proof. If the sequence x_n is bounded, $^*x_h$ is finite for each $h \in \mathbb{N}_\infty$. Hence, $\mathrm{st}(x_h)$ is an accumulation point of x by Theorem 7.1. The assumption thus implies that $x = \mathrm{st}(x_h) \in \mathbb{R}$ is independent of h, i.e. $^*x_h \approx {}^*x$ for all $h \in \mathbb{N}_\infty$ which implies $x_n \to x$ by Theorem 7.1. □

The proof of the classical Bolzano-Weierstraß theorem is even simpler:

Corollary 7.5 (Bolzano-Weierstraß). *Any bounded real sequence has an accumulation point.*

Proof. If the sequence x_n is bounded, ${}^*x_h$ is finite for each $h \in \mathbb{N}_\infty$, in particular finite for *some* $h \in \mathbb{N}_\infty$. Then $\operatorname{st}(x_h)$ is an accumulation point of x_n. □

We emphasize once more that despite the simplicity of the above proof, the nonstandard method has in contrast to the standard method the disadvantage that it relies on the axiom of choice.

Also for the well-known limit rules we have simple proofs:

Corollary 7.6. *For real convergent sequences $x_n \to x$ and $y_n \to y$, we have $x_n \pm y_n \to x \pm y$, $x_n \cdot y_n \to x \cdot y$ and $x_n / y_n \to x/y$ (if $y_n, y \neq 0$).*

Proof. For each $h \in \mathbb{N}_\infty$, we have ${}^*x_h \approx {}^*x$ and ${}^*y_h \approx {}^*y$ by Theorem 7.1, and so ${}^*x_h \pm {}^*y_h \approx {}^*x \pm {}^*y$ by Proposition 5.17 which by Theorem 7.1 implies $x_n \pm y_n \to x \pm y$; the other statements are proved analogously. □

Exercise 28. Prove that for any bounded real sequence

$$\limsup x_n = \sup \{\operatorname{st}({}^*x_h) : h \in \mathbb{N}_\infty\}$$

and

$$\liminf x_n = \inf \{\operatorname{st}({}^*x_h) : h \in \mathbb{N}_\infty\}.$$

What can be said if the sequence is unbounded?

Exercise 29. Find a real sequence x_n with more than one accumulation point such that you are able to calculate ${}^*x_h$ for any $h \in \mathbb{N}_\infty$.

Interpret your example also for the map $*$ of Theorem 4.20.

Now we are in a position to prove that internal sets in nonstandard embeddings are either finite or have at least the cardinality of the continuum:

Proof of Theorem 3.23. Let $* : \widehat{S} \to \widehat{{}^*S}$ be a nonstandard embedding. If S is finite, then all sets in $\widehat{S}$ are finite, and $*$ is a bijection (recall the remarks following Corollary 3.11); in this case, all entities of $\widehat{{}^*S}$ are finite, and the claim is trivial.

Thus, assume that S is infinite. Then it is no loss of generality to assume that $\mathbb{N} \subseteq S$ (just rename the atoms). We prove first that $\{1, \ldots, h\}$ has the cardinality of the continuum for each $h \in \mathbb{N}_\infty$. To this end, consider the real sequence defined by $x_{2^n+k} := k/2^n$ $(k = 0, \ldots, 2^n - 1)$, i.e. $x_1 = x_2 = 0$, $x_3 = 1/2$, $x_4 = 0$, $x_5 = 1/4$, $x_6 = 2/4$, $x_7 = 3/4$, $x_8 = 0$, ... Then we have for any $x \in \mathbb{R}$, $0 \le x \le 1$, that

$$\forall \underline{m} \in \mathbb{N} : \exists \underline{n} \in \mathbb{N} : (\underline{n} \le 2^{\underline{m}+1} \wedge |x_{\underline{n}} - x| \le 2^{-\underline{m}}).$$

The transfer principle implies

$$\forall \underline{m} \in {}^*\mathbb{N} : \exists \underline{n} \in {}^*\mathbb{N} : (\underline{n} \le 2^{\underline{m}+1} \wedge |{}^*x_{\underline{n}} - {}^*x| \le 2^{-\underline{m}}). \tag{7.1}$$

Observe now that, by the transfer principle,

$$\forall \underline{n} \in {}^*\mathbb{N} : (\underline{n} \geq 4 \implies \exists \underline{m} \in {}^*\mathbb{N} : 2^{\underline{m}+1} \leq \underline{n} < 2^{\underline{m}+2}).$$

Hence, since $h \in \mathbb{N}_\infty$, we find some $m \in {}^*\mathbb{N}$ such that $2^{m+1} \leq h \leq 2^{m+2}$. We cannot have $m \in {}^\sigma\mathbb{N}$, since this would imply $h \in {}^\sigma\mathbb{N}$ in view of Proposition 5.9. Thus, $m \in \mathbb{N}_\infty$, and so $2^{-m} \approx 0$ (apply e.g. Theorem 7.1), i.e. $2^{-m} \in \inf({}^*\mathbb{R})$. By (7.1), we find for each $x \in \mathbb{R}$, $0 \leq x \leq 1$, some $n(x) \in {}^*\mathbb{N}$ with $n(x) \leq 2^{m+1} \leq h$ and $|{}^*x_{n(x)} - {}^*x| \leq 2^{-m} \in \inf({}^*\mathbb{R})$, and so ${}^*x_{n(x)} \approx {}^*x$. By the axiom of choice, we may assume that $n : [0,1] \to \{1, \dots, h\}$ is a map. Moreover, n is one-to-one, since for real numbers $x_1, x_2 \in [0,1]$ with $n(x_1) = n(x_2)$, we have ${}^*x_1 \approx {}^*x_{n(x_1)} = {}^*x_{n(x_2)} \approx {}^*x_2$, i.e. ${}^*x_1 \approx {}^*x_2$ which for real numbers implies $x_1 = x_2$. We thus have established for any $h \in \mathbb{N}_\infty$ an injection $x : [0,1] \to \{1, \dots, h\}$, and so any $\{1, \dots, h\}$ $(h \in \mathbb{N}_\infty)$ has at least the cardinality of the continuum.

Now let A be an internal entity. Theorem 6.12 implies that we either find an internal injection $f : {}^*\mathbb{N} \to A$ or an internal bijection $f : \{1, \dots, h\} \to A$ where $h \in {}^*\mathbb{N}$. Thus, we find either an internal bijection $f : \{1, \dots, n\} \to A$ for some $n \in {}^\sigma\mathbb{N}$ in which case A is finite, or an internal injection $f : \{1, \dots, h\} \to A$ for some $h \in \mathbb{N}_\infty$. The latter implies that A has at least the cardinality of the continuum by what we proved above. □

Recall that a sequence $x_n \in \mathbb{R}$ is called a *Cauchy sequence*, if for each $\varepsilon > 0$ there is some n_0 such that $|x_n - x_m| < \varepsilon$ for $n, m \geq n_0$.

Exercise 30. Prove, without using the fact that Cauchy sequences converge, that a real sequence x_n is a Cauchy sequence if and only if ${}^*x_h \approx {}^*x_k$ for each $h, k \in \mathbb{N}_\infty$.

With Exercise 30, we find another nonstandard proof of a well-known standard fact:

Corollary 7.7. *A real sequence converges if and only if it is a Cauchy sequence (i.e.,* $\mathbb{R}$ *is complete).*

Proof. If $x_n \to x$ converges, then ${}^*x_h \approx {}^*x \approx {}^*x_k$ for each $h, k \in \mathbb{N}_\infty$ by Theorem 7.1. Conversely, if x_n is a Cauchy sequence, then x_n is bounded, and so ${}^*x_h$ is finite for any $h \in \mathbb{N}_\infty$ by Theorem 7.2. Put $x := \operatorname{st}(x_h)$ for some $h \in \mathbb{N}_\infty$. Then Exercise 30 implies ${}^*x_k \approx {}^*x_h \approx {}^*x$ for each $k \in \mathbb{N}_\infty$, and it follows from Theorem 7.1 that $x_n \to x$. □

It is known that the completeness of $\mathbb{R}$ is equivalent to its Dedekind completeness: We used the Dedekind completeness of $\mathbb{R}$ in the previous proof implicitly when we made use of the function st (recall the proof of Theorem 5.19).

7.2 Sets

While the fact that boundedness (in the classical sense) and finiteness (in the nonstandard sense) are related for sequences is rather intuitive, the reader may be surprised to see the same relation for sets:

Theorem 7.8. *A set $A \subseteq \mathbb{R}$ is bounded if and only if ${}^*A$ contains only finite elements, i.e. if and only if*

$$ {}^*A \subseteq \operatorname{fin}({}^*\mathbb{R}). \tag{7.2}$$

More precisely, A is unbounded from above if and only if A contains an infinite positive element, and A is unbounded from below if and only if A contains an infinite negative element.

Proof. If $A \subseteq \mathbb{R}$ is bounded from above, we find some $c \in \mathbb{R}_+$ with

$$\forall \underline{x} \in A : \underline{x} \le c.$$

The transfer reads $\forall \underline{x} \in {}^*A : \underline{x} \le {}^*c$ which implies that all elements of ${}^*A$ are either finite or negative. Conversely, if each $x \in {}^*A$ is either finite or negative, then any sequence $x_n \in A$ is bounded from above. Indeed, otherwise there were some positive infinite ${}^*x_h$ by Theorem 7.2. The transfer of the sentence $\forall \underline{n} \in \mathbb{N} : x_{\underline{n}} \in A$ implies in particular that ${}^*x_h \in {}^*A$, i.e. A contains an infinite positive element. □

Thus the map $A \mapsto {}^*A$ reflects the boundedness of A by joining some positive resp. negative infinite elements to ${}^*A$ if A is unbounded. It is not very surprising that it also reflects the local properties of A. In particular, we have:

Theorem 7.9. *A set $A \subseteq \mathbb{R}$ is* closed *if and only if each finite point of ${}^*A$ is infinitely close to some (standard) point of ${}^\sigma A$, i.e. if and only if*

$$\operatorname{st}({}^*A \cap \operatorname{fin}({}^*\mathbb{R})) = A. \tag{7.3}$$

Proof. Let A be closed, and $x \in {}^*A$ be finite. We claim that $\operatorname{st}(x) \in A$. Indeed, put $y := \operatorname{st}(x)$. For each $\varepsilon \in \mathbb{R}_+$, we have

$$\exists \underline{z} \in {}^*A : |\underline{z} - {}^*y| < {}^*\varepsilon,$$

because ${}^*A \ni x \approx {}^*y$. The converse form of the transfer principle implies that we find some $z \in A$ with $|z - y| < \varepsilon$. Since $\varepsilon \in \mathbb{R}_+$ was arbitrary and A is closed, this implies $y \in A$, as claimed.

Conversely, if each finite point of ${}^*A$ is infinitely close to some point from ${}^\sigma A$, and $x_n \in A$ is a sequence with $x_n \to x$, we have $x \in A$: Indeed, the transfer principle implies ${}^*x_h \in {}^*A$ for each $h \in {}^*\mathbb{N}$, and by Theorem 7.1, we have ${}^*x_h = {}^*x$ for some $h \in \mathbb{N}_\infty$. Hence, ${}^*x \in {}^*A$. By assumption, ${}^*x$ is infinitely close to some standard point of ${}^\sigma A$. Since ${}^*x$ is itself standard, it must be that standard point of ${}^\sigma A$, i.e. ${}^*x \in {}^\sigma A$, and so $x \in A$. Hence, A is closed. □

We call a set $A \subseteq \mathbb{R}$ *compact* if A is closed and bounded. The previous results imply the following consequence:

Corollary 7.10. *A set $A \subseteq \mathbb{R}$ is compact if and only if each point from ${}^*A$ is infinitely close to some (standard) point of ${}^\sigma A$.*

Proof. If A is compact, then a combination of (7.2) and (7.3) shows that ${}^\sigma({}^*A) = {}^*A$. Conversely, if each $x \in {}^*A$ is infinitely close to some standard point, then A is closed by Theorem 7.9, and each $x \in {}^*A$ is finite, whence A is bounded by Theorem 7.8. □

We will see later another (deeper) reason why Corollary 7.10 is true. For later applications, Corollary 7.10 is one of the most essential tools. In fact, all results which are typically proved by the Heine-Borel compactness criterion (a set is compact if and only if each open covering has a finite subcovering) can usually more easily be proved by an application of Corollary 7.10.

Exercise 31. Prove that a point $x \in A$ is an interior point of A if and only if the relation $y \approx {}^*x$ implies $y \in {}^*A$, i.e. if $\mathrm{mon}(x) \in {}^*A$. Thus, A is *open* if and only if

$$\bigcup_{x\in A} \mathrm{mon}(x) \subseteq {}^*A.$$

Exercise 32. Give a standard characterization of those sets $A \subseteq \mathbb{R}$ with the property that the relations $x \in {}^*A$ and $y \approx x$ imply $y \in {}^*A$.

Why does the answer not contradict Exercise 31?

Exercise 33. Give a standard characterization of those sets $A \subseteq \mathbb{R}$ satisfying

$${}^*A = \bigcup_{x\in A} \mathrm{mon}(x).$$

The following exercises are easier to prove if one makes use of Theorem 7.12 below. However, the author recommends solving them *now* (without appealing to Theorem 7.12).

Exercise 34. Give a standard characterization of those sets $A \subseteq \mathbb{R}$ with the property that each finite point $x \in {}^*A$ satisfies $x = {}^*(\mathrm{st}(x))$.

Exercise 35. Recall that a set $A \subseteq \mathbb{R}$ is called *perfect*, if each point $x \in A$ is an accumulation point of $A \setminus \{x\}$. Give a nonstandard characterization of perfect sets.

Theorem 7.11. *A point $x \in \mathbb{R}$ belongs to the closure of some set $A \subseteq \mathbb{R}$ if and only if x is the standard part of some point from ${}^*A$, i.e. if and only if $x \in \mathrm{st}({}^*A)$.*

Proof. If x belongs to the closure of A, then there is a sequence $x_n \in A$ with $x_n \to x$. Then ${}^*x_h \approx {}^*x$ for some $h \in \mathbb{N}_\infty$ by Theorem 7.1. Since the permanence principle implies ${}^*x_h \in {}^*A$ and since $x = \mathrm{st}({}^*x_h)$, we have the required representation of x. Conversely, if ${}^*x = \mathrm{st}(y)$ for some $y \in {}^*A$, then $y \in {}^*\overline{A}$, since the permanence principle implies ${}^*A \subseteq {}^*\overline{A}$. By Theorem 7.9, y is infinitely close to some standard point of ${}^\sigma\overline{A}$. Since ${}^*x$ is such a standard point, we must have ${}^*x \in {}^\sigma\overline{A}$, i.e. $x \in \overline{A}$. □

Theorem 7.12. *Let $A \subseteq \mathbb{R}$. A point $x \in A$ is isolated if and only if ${}^*A$ contains no point $y \neq {}^*x$ with $y \approx {}^*x$.*

Proof. If x is isolated, there is some $\varepsilon \in \mathbb{R}_+$ such that

$$\forall \underline{y} \in A : (\underline{y} \neq x \implies |\underline{y} - x| > \varepsilon).$$

The transfer implies that all points $y \in {}^*A \setminus \{{}^*x\}$ satisfy $|y - {}^*x| > {}^*\varepsilon$, and so $y \not\approx {}^*x$.

Conversely, if x is not isolated, there is a sequence $x_n \in A$, $x_n \neq x$ with $x_n \to x$. We have ${}^*x_h \approx {}^*x$ for some $h \in \mathbb{N}_\infty$ by Theorem 7.1, and the transfer principle implies ${}^*x_h \in {}^*A$ and ${}^*x_h \neq x$. Hence, $y = {}^*x_h \in {}^*A$ satisfies $y \neq {}^*x$ and $y \approx {}^*x$. □

7.3 Functions

Throughout, we consider functions $f : D \to \mathbb{R}$ where $D \subseteq \mathbb{R}$. Then ${}^*f$ defines a function ${}^*f : {}^*D \to {}^*\mathbb{R}$ which extends f. One might expect that ${}^*f$ reflects properties like continuity of f in nonstandard terms. This is indeed true.

Theorem 7.13. *Let x_0 be an accumulation point of D. Then for $c \in \mathbb{R}$ the following statements are equivalent:*

1. $\lim\limits_{\substack{x \to x_0 \\ x \in D}} f(x) = c.$
2. *For any $x \in {}^*D$ with ${}^*x_0 \neq x \approx {}^*x_0$ we have ${}^*f(x) \approx {}^*c$.*

Proof. Let $f(x) \to c$ as $x \to x_0$. For any $\varepsilon \in \mathbb{R}_+$, we find some $\delta \in \mathbb{R}_+$ with

$$\forall \underline{x} \in D : (0 < |\underline{x} - x_0| < \delta \implies |f(\underline{x}) - c| < \varepsilon).$$

The transfer principle implies

$$\forall \underline{x} \in {}^*D : (0 < |\underline{x} - {}^*x_0| < {}^*\delta \implies |{}^*f(\underline{x}) - {}^*c| < {}^*\varepsilon).$$

In particular, $|{}^*f(x) - {}^*c| < {}^*\varepsilon$ whenever ${}^*x_0 \neq x \approx {}^*x_0$. Since this holds for all $\varepsilon \in \mathbb{R}_+$, we even have ${}^*f(x) \approx {}^*c$.

Conversely, if ${}^*x_0 \neq x \approx {}^*x_0$ implies ${}^*f(x) \approx {}^*c$, let $\varepsilon \in \mathbb{R}_+$ be given. Then the following internal predicate is true for all infinitesimal $d \in \inf({}^*\mathbb{R})$, $d > 0$:

$$\forall \underline{x} \in {}^*D : (0 < |\underline{x} - {}^*x_0| < d \implies |{}^*f(\underline{x}) - {}^*c| < {}^*\varepsilon).$$

By the permanence principle for ${}^*\mathbb{R}$ (Cauchy principle), the predicate then also holds for some $d = {}^*\delta$ where $\delta \in \mathbb{R}_+$. The inverse direction of the transfer principle implies

$$\forall \underline{x} \in D : (0 < |\underline{x} - x_0| < \delta \implies |f(\underline{x}) - c| < \varepsilon).$$

But this means that $f(x) \to c$ as $x \to x_0$. □

Corollary 7.14. *$f : D \to \mathbb{R}$ is continuous at $x_0 \in \overline{D}$ if and only if the relation $x \in {}^*D$, $x \approx {}^*x_0$, implies ${}^*f(x) \approx {}^*(f(x_0)) = {}^*f({}^*x_0)$.*

Proof. If $x_0 \in D$ is not isolated, then f is continuous at x_0 if and only if $\lim\limits_{\substack{x\to x_0\\ x\in D}} f(x) = f(x_0)$. Thus, the statement follows from Theorem 7.13.

If $x_0 \in D$ is isolated, any function $f : D \to \mathbb{R}$ is continuous at x_0. Moreover, by Theorem 7.12 the only point $x \in {}^*D$ which satisfies $x \approx {}^*x_0$ is $x = {}^*x_0$, and so ${}^*f(x) \approx {}^*f({}^*x_0)$ is always satisfied. □

As an application, let us give a simple proof of the following fact whose proof is much more complicated by standard methods (recall that we defined compact subsets of $\mathbb{R}$ simply as the closed and bounded subsets):

Corollary 7.15. *If $D \subseteq \mathbb{R}$ is compact and $f : D \to \mathbb{R}$ is continuous, then $f(D)$ is compact.*

Proof. Put $B := f(D)$. By Corollary 7.10, we have to prove that each point $y \in {}^*B$ is infinitely close to some standard point of ${}^\sigma B$. Thus, let $x \in {}^*B$ be given. Since ${}^*f : {}^*D \to {}^*B$ is onto (Theorem 3.13), we have $y = {}^*f(x)$ for some $x \in {}^*D$. Since D is compact, Corollary 7.10 implies that x is infinitely close to some point ${}^*x_0$ with $x_0 \in D$. Since f is continuous at x_0 and $x \approx {}^*x_0$, Corollary 7.14 implies $y = {}^*f(x) \approx {}^*f({}^*x_0) = {}^*(f(x_0)) \in {}^\sigma B$, as desired. □

As a further application, we prove that continuous functions map intervals into intervals:

Corollary 7.16 (Intermediate Value Theorem)**.** *If $f : [a, b] \to \mathbb{R}$ is continuous, then f attains all values between $f(a)$ and $f(b)$.*

Proof. Without loss of generality, let $f(a) < c < f(b)$, and we have to prove that $c \in B := f([a, b])$. Choose $h \in \mathbb{N}_\infty$, and let $x_n := a + n(b - a)/h$ $(n = 0, \ldots, h)$ be an infinite equidistant partition of $[a, b]$. Let $n_0 \in {}^*\mathbb{N}$ be the first index with ${}^*f(x_{n_0}) > {}^*c$, i.e. ${}^*f(x_{n_0-1}) \le {}^*c$. By Corollary 7.10, the point x_{n_0} is infinitely close to some standard point from ${}^\sigma[a, b]$, i.e. $x_{n_0} \approx {}^*x$ for some $x \in [a, b]$. Since ${}^*x \approx x_{n_0} \approx x_{n_0-1}$ and since f is continuous at x, we have ${}^*f({}^*x) \approx {}^*f(x_{n_0}) > {}^*c$ and ${}^*f({}^*x) \approx {}^*f(x_{n_0-1}) \le {}^*c$ which implies ${}^*f({}^*x) \approx {}^*c$ and so $f(x) = c$ (since all points are standard points). □

Exercise 36. In the previous proof, we used that there is a first index $n_0 \in {}^*\mathbb{N}$ with ${}^*f(x_{n_0}) > {}^*c$. Why does such an index exist?

Theorem 7.17. *A function $f : D \to \mathbb{R}$ is uniformly continuous if and only if the relations $x, y \in {}^*D$ and $x \approx y$ imply ${}^*f(x) \approx {}^*f(y)$.*

Proof. Let $f : D \to \mathbb{R}$ be uniformly continuous. For any $\varepsilon \in \mathbb{R}_+$, we find some $\delta \in \mathbb{R}_+$ such that, in view of the transfer principle,

$$\forall \underline{x}, \underline{y} \in {}^*D : (|\underline{x} - \underline{y}| < {}^*\delta \implies |{}^*f(\underline{x}) - {}^*f(\underline{y})| < {}^*\varepsilon).$$

In particular, the relation $x \approx y$ for hyperreal numbers $x, y \in {}^*D$ implies $|{}^*f(x) - {}^*f(y)| < {}^*\varepsilon$. Since this holds for any $\varepsilon \in \mathbb{R}_+$, we even have ${}^*f(x) \approx {}^*f(y)$.

Conversely, if $x \approx y$ implies ${}^*f(x) \approx {}^*f(y)$, then we have for any $\varepsilon \in \mathbb{R}_+$ that the internal predicate

$$\forall \underline{x}, \underline{y} \in {}^*D : (|\underline{x} - \underline{y}| < c \implies |{}^*f(\underline{x}) - {}^*f(\underline{y})| < {}^*\varepsilon)$$

holds for any infinitesimal $c \in \inf({}^*\mathbb{R})$, $c > 0$. By the Cauchy principle, this predicate holds also for some $c = {}^*\delta$ with $\delta \in \mathbb{R}_+$. The converse direction of the transfer principle now shows that the relation $|x - y| < \delta$ for $x, y \in D$ implies $|f(x) - f(y)| < \varepsilon$. Hence, f is uniformly continuous. □

Theorem 7.17 might appear strange at first glance, because it is not clear how the uniformity comes into play, compared to e.g. Corollary 7.14: The only difference to the characterization of continuous functions by Corollary 7.14 is that we want the relation ${}^*f(x) \approx {}^*f(y)$ for $x \approx y$ even if y is a nonstandard point. In this sense, Theorem 7.17 is in a certain sense a "local" (nonstandard) characterization of uniform continuity which is somewhat paradoxical.

Employing the above paradox, we get a simple proof for another well-known standard result:

Corollary 7.18. *If $D \subseteq \mathbb{R}$ is compact, then any continuous $f : D \to \mathbb{R}$ is uniformly continuous.*

Proof. Let $x, y \in {}^*D$ with $x \approx y$. By Theorem 7.17, we have to prove that ${}^*f(x) \approx {}^*f(y)$. But since D is compact, we find by Corollary 7.10 some $x_0 \in D$ with $x \approx {}^*x_0$. Since $x, y \approx {}^*x_0$, Corollary 7.14 implies ${}^*f(x) \approx {}^*(f(x_0)) \approx {}^*f(y)$, as claimed. □

We now come to the real calculus:

Theorem 7.19. *Let $x_0 \in D$ be an accumulation point of $D \subseteq \mathbb{R}$. Then $f : D \to \mathbb{R}$ is differentiable in x_0 with derivative $c \in \mathbb{R}$ if and only if for each $x \in {}^*D$ with $x \approx {}^*x_0$ and $x \neq x_0$ the relation*

$$\frac{{}^*f(x) - {}^*f({}^*x_0)}{x - {}^*x_0} \approx {}^*c$$

holds. Equivalently: For each $0 \neq dx \approx 0$ *with* ${}^*x_0 + dx \in {}^*D$ *the relation*

$$\frac{{}^*f({}^*x_0 + dx) - {}^*f({}^*x_0)}{dx} \approx {}^*c \tag{7.4}$$

holds.

Proof. Put $g(x) := (f(x) - f(x_0))/(x - x_0)$. Then we have ${}^*g(x) = ({}^*f(x) - {}^*(f(x_0)))/(x - {}^*x_0)$ (why?). Now the first statement follows by Theorem 7.13. The second statement follows by putting $dx := x - x_0$ resp. $x := x_0 + dx$. □

Theorem 7.19 implies that we can really do calculus with infinitesimals in the sense of Leibniz by just dropping infinitesimal terms:

Example 7.20. Let us determine the derivative of the function $f(x) = x^2$. For $x \in {}^*\mathbb{R}$ and $h \in \inf({}^*\mathbb{R})$, we have

$$\frac{{}^*f(x + dx) - f(x)}{dx} = \frac{x^2 + 2xdx + dx^2 - x^2}{dx} \approx 2x,$$

where the "loss" of the infinitesimal dx is justified, because we wrote "$\approx$" instead of "$=$". A comparison with (7.4) for $x = {}^*x_0$ now shows that f is differentiable in x_0 with derivative $2x_0$.

Corollary 7.21. *If f is differentiable at x_0, then f is continuous at x_0.*

Proof. For $x \approx {}^*x_0$, we have ${}^*f(x) - {}^*f({}^*x_0) \approx (x - {}^*x_0){}^*c \approx 0$, and so ${}^*f(x) \approx {}^*f({}^*x_0)$. □

Exercise 37. Prove by nonstandard methods that $f(x) = |x|$ is not differentiable at 0.

We get now a clear proof for the chain rule of the calculus. As usual, we write $f'(x_0)$ for the derivative of f at x_0 (if it exists).

Corollary 7.22 (Chain rule)**.** *We have for real functions f, g:*

$$(f \circ g)'(x_0) = f'(g(x_0))g'(x_0),$$

provided the derivatives on the right-hand side exist.

Proof. Put $F(x) := f(g(x))$. Given $dx \approx 0$ with $dx \neq 0$, we define $dg := {}^*g({}^*x_0 + dx) - {}^*g({}^*x_0)$ and $df := {}^*F({}^*x_0 + dx) - {}^*F({}^*x_0) = {}^*f({}^*g({}^*x_0) + dg) - {}^*f({}^*g({}^*x_0))$. We have by Theorem 7.19

$$\frac{dg}{dx} \approx {}^*(g'(x_0)),$$

in particular also $dg \approx g'(x_0)dx \approx 0$. Thus, in case $dg \neq 0$, Theorem 7.19 implies

$$\frac{df}{dg} \approx {}^*(f'(g(x_0))),$$

which in view of the above formula shows that

$$\frac{df}{dx} = \frac{df}{dg} \cdot \frac{dg}{dx} \approx {}^*(f'(g(x_0))g'(x_0)).$$

If $dg = 0$, we have $df = 0$, and so the previous formula holds also (because ${}^*(g'(x_0)) \approx dg/dx = 0$ and $df/dx = 0$). Now the statement follows by Theorem 7.19. □

Thus, essentially, the chain rule follows by just multiplying nominator and numerator of $\frac{df}{dx}$ by dg: The crucial point here is that we may in fact calculate with the infinitesimals df, dx, and dg as if they were real numbers. As for real numbers, one only has to take care of the special case $dg = 0$.

Exercise 38. Give a nonstandard proof (as intuitive as possible) of the product formula

$$(f \cdot g)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0).$$

Exercise 39. Let $f : [a, b] \to \mathbb{R}$ be differentiable on (a, b) with derivative f'. Derive from the mean value theorem that for each $x, y \in {}^*\mathbb{R}$ with ${}^*a \le x < y \le {}^*b$ there is some $\xi \in {}^*\mathbb{R}$, $x < \xi < y$ such that

$$\frac{{}^*f(x) - {}^*f(y)}{x - y} = {}^*f'(\xi).$$

We now turn to the integral:

Theorem 7.23. *If $f : [a, b] \to \mathbb{R}$ is* Riemann-Stieltjes integrable *with respect to some function $\varphi : [a, b] \to \mathbb{R}$, then the integral may be infinitely closely approximated by a Riemann-Stieltjes sum to an infinitely fine internal partition, i.e.*

$$^*\left(\int_a^b f(x)\, d\varphi(x)\right) \approx \sum_{\underline{n}=1}^{h} {}^*f(x_{\underline{n}-1})({}^*\varphi(x_{\underline{n}}) - {}^*\varphi(x_{\underline{n}-1})) \tag{7.5}$$

where $x_0 = {}^*a$, $x_h = {}^*b$, and $0 < x_n - x_{n-1} < \delta$ for $n = 1, \ldots, h$ where $0 < \delta \in \inf({}^*\mathbb{R})$, and the *-finite sequence x_n is internal. Conversely, if the right-hand side of (7.5) *is finite and has the same standard part for all infinitely fine internal partitions, then f is Riemann-Stieltjes integrable with respect to φ.*

Proof. If f is Riemann-Stieltjes integrable with integral c, then we find for any $\varepsilon \in \mathbb{R}_+$ some $\delta \in \mathbb{R}_+$ such that

$$\forall \underline{x} \in \mathbb{R}^{\mathbb{N}} : \exists \underline{n} \in \mathbb{N} : \alpha(\underline{x}, \underline{n}, \delta) \implies \left| c - \sum_{\underline{k}=1}^{\underline{n}} f(\underline{x}_{\underline{k}-1})(\varphi(\underline{x}_{\underline{k}}) - \varphi(\underline{x}_{\underline{k}-1})) \right| < \varepsilon,$$

where $\alpha(\underline{x}, \underline{n}, \delta)$ is a shortcut for

$$\underline{x}_0 = a \wedge \underline{x}_{\underline{k}} = b \wedge \forall \underline{k} \in \mathbb{N} : (0 < \underline{k} \leq \underline{n} \implies 0 < \underline{x}_{\underline{k}} - \underline{x}_{\underline{k}-1} < \delta).$$

Now we apply the transfer principle, observing that ${}^*(\mathbb{R}^{\mathbb{N}})$ consists of all internal sequences, and that any *-finite partition may be extended to such a sequence. We thus find that for all infinitely fine internal partitions $x_1, \ldots, x_h$, the relation

$$\left| {}^*c - \sum_{\underline{n}=1}^{h} {}^*f(x_{\underline{n}-1})({}^*\varphi(x_{\underline{n}}) - {}^*\varphi(x_{\underline{n}-1})) \right| < {}^*\varepsilon$$

holds. Since $\varepsilon \in \mathbb{R}_+$ was arbitrary, we have

$$\sum_{\underline{n}=1}^{h} {}^*f(x_{\underline{n}-1})({}^*\varphi(x_{\underline{n}}) - {}^*\varphi(x_{\underline{n}-1})) \approx {}^*c,$$

as claimed.

For the second statement, assume that the right-hand side of (7.5) is infinitely close to ${}^*c$ for some $c \in \mathbb{R}$ whenever x_n is an infinitely fine internal partition. Then we have for any $\varepsilon \in \mathbb{R}_+$ that the internal predicate

$$\forall \underline{x} \in {}^*(\mathbb{R}^{\mathbb{N}}) : \exists \underline{n} \in {}^*\mathbb{N} :$$
$${}^*\alpha(\underline{x}, \underline{n}, \underline{z}) \implies \left| {}^*c - \sum_{\underline{k}=1}^{\underline{n}} {}^*f(\underline{x}_{\underline{k}-1})({}^*\varphi(\underline{x}_{\underline{k}}) - {}^*\varphi(\underline{x}_{\underline{k}-1})) \right| < {}^*\varepsilon$$

holds for any $\underline{z} \in \inf({}^*\mathbb{R})$, $\underline{z} > 0$. By the permanence principle (Cauchy principle), the above internal predicate holds for some $\underline{z} = {}^*\delta$, $\delta \in \mathbb{R}_+$. Then the inverse direction of the permanence principle implies that the Riemann-Stieltjes sum for any finite δ-fine partition differs from c by less than ε. Hence, f is Riemann-Stieltjes integrable. □

Exercise 40. If $f : [a, b] \to \mathbb{R}$ is continuous and $\varphi : [a, b] \to \mathbb{R}$ is monotone, it is well-known that f is Riemann-Stieltjes integrable with respect to φ. Prove that in this case

$${}^*\left(\int_a^b f(x)\, d\varphi(x) \right) \approx \sum_{\underline{n}=1}^{h} {}^*f(y_{\underline{n}})({}^*\varphi(x_{\underline{n}}) - {}^*\varphi(x_{\underline{n}-1}))$$

whenever x_n is an infinitely fine internal partition and y_n is an arbitrary internal sequence with $x_{n-1} \le y_n \le x_n$.

Proof. Since ${}^*f$ is uniformly continuous, we have ${}^*f(y_{\underline{n}}) \approx {}^*f(x_{\underline{n}-1})$ and so $|{}^*f(y(\underline{n})) - {}^*f(x_{\underline{n}-1})| < \varepsilon$ for any $\underline{n} \in {}^*\mathbb{N}$ and and any $\varepsilon \in {}^\sigma\mathbb{R}_+$. Hence,

$$\begin{aligned} \left|\sum_{\underline{n}=1}^{h}({}^*f(y_{\underline{n}}) - {}^*f(x_{\underline{n}-1}))({}^*\varphi(x_{\underline{n}}) - {}^*\varphi(x_{\underline{n}-1}))\right| &\le \varepsilon \sum_{\underline{n}=1}^{h} |{}^*\varphi(x_{\underline{n}}) - {}^*\varphi(x_{\underline{n}-1})| \\ &= \varepsilon\, |\varphi(b) - \varphi(a)| \end{aligned}$$

for any $\varepsilon \in {}^\sigma\mathbb{R}_+$. Thus, the left-hand side is infinitesimal which means that

$$\sum_{\underline{n}=1}^{h} {}^*f(y_{\underline{n}})({}^*\varphi(x_{\underline{n}}) - {}^*\varphi(y_{\underline{n}})) \approx \sum_{\underline{n}=1}^{h} {}^*f(x_{\underline{n}-1})({}^*\varphi(x_{\underline{n}}) - {}^*\varphi(x_{\underline{n}-1})).$$

Now the claim follows from Theorem 7.23. □

As an application, we prove one part of the fundamental theorem of calculus:

Corollary 7.24. *If $f : [a,b] \to \mathbb{R}$ is continuously differentiable, then*

$$\int_a^b f'(x)\, dx = f(b) - f(a).$$

Proof. Let $x_0, \ldots, x_h$ be an infinitely fine partition. Since F is differentiable, we have by Exercise 39 that

$$\frac{{}^*f(x_{n+1}) - {}^*f(x_n)}{x_{n+1} - x_n} = {}^*f'(\xi_n)$$

where $x_n < \xi_n < x_{n+1}$. Multiplying by $(x_{n+1} - x_n)$ and summing up, we find

$${}^*f({}^*b) - {}^*f({}^*a) = \sum_{n=1}^{h} {}^*f'(\xi_n)(x_{n+1} - x_n).$$

By Exercise 40, the right-hand side is infinitely close to ${}^*\left(\int_a^b f'(x)\, dx\right)$, and the left-hand side is equal to ${}^*(f(b) - f(a))$. □

Concerning functions of more variables, we restrict ourselves to a nonstandard continuity criterion which is somewhat surprising, since it suffices to consider continuity in each variable separately. However, the point is that this continuity is even needed at nonstandard points (recall the remarks following Theorem 7.17).

Theorem 7.25. *Let $D = [a,b]\times[c,d]$. Then $f : D \to \mathbb{R}$ is continuous at $(x_0, y_0) \in D$ if and only if for each $(x,y) \in {}^*D$ with $x \approx {}^*x_0$ and $y \approx {}^*y_0$ we have ${}^*f(x,y) \approx {}^*f(x, {}^*y_0)$ and ${}^*f(x,y) \approx f({}^*x_0, y)$.*

Proof. Let f have the properties of the statement. Then we actually have ${}^*f(x,y) \approx {}^*f(x, {}^*y_0) \approx {}^*f({}^*x_0, {}^*y_0)$ whenever $(x,y) \in {}^*D$ satisfy $x \approx {}^*x_0$ and $y \approx {}^*y_0$. Hence, for each $\varepsilon \in \mathbb{R}_+$ the following predicate holds for each infinitesimal $\underline{z} > 0$:

$$\begin{gathered} \forall \underline{x} \in {}^*[a,b], \underline{y} \in {}^*[c,d] : \\ (|\underline{x} - x_0| < \underline{z} \wedge |\underline{y} - y_0| < \underline{z}) \implies |{}^*f(\underline{x}, \underline{y}) - {}^*f({}^*x_0, {}^*y_0)| < {}^*\varepsilon. \end{gathered} \tag{7.6}$$

By the permanence principle (Cauchy principle), the predicate holds also for some $\underline{z} = {}^*\delta$ with $\delta \in \mathbb{R}_+$. The inverse direction of the transfer principle shows that f is continuous at (x_0, y_0).

Conversely, let f be continuous at (x_0, y_0). Then we find for all $\varepsilon \in \mathbb{R}_+$ some $\delta \in \mathbb{R}_+$ such that by the transfer principle the sentence (7.6) is true for $\underline{z} = {}^*\delta$. In particular, if $x \approx x_0$ and $y \approx y_0$, then $|{}^*f(x,y) - {}^*f(x_0,y_0)| < {}^*\varepsilon$. Since this holds for all $\varepsilon \in \mathbb{R}_+$, we even have ${}^*f(x,y) \approx {}^*f(x_0,y_0)$. But since then ${}^*f({}^*x_0, y) \approx {}^*f({}^*x_0, {}^*y_0) \approx {}^*f(x, {}^*y_0)$, we have also ${}^*f(x,y) \approx {}^*f({}^*x_0, y) \approx {}^*f(x, {}^*y_0)$. □

We shall now use nonstandard analysis to define *explicitly* a nonmeasurable function. Since this is not possible without the (uncountable) axiom of choice (at least under the assumption that the existence of a so-called inaccessible cardinal is consistent [Sol70]), we already see that the (uncountable) axiom of choice cannot be avoided in the construction of the mapping $*$.

We need some preparation. At first, we make use of the following well-known generalization of the fundamental theorem of calculus for the Lebesgue integral (for a proof, see e.g. [Rud87, HS69]):

Proposition 7.26. *If $f : [a,b] \to \mathbb{R}$ is integrable in the sense of Lebesgue, then*

$$F(x) = \int_a^x f(t)\,dt$$

is differentiable at almost all points of $[a,b]$ (in the sense of Lebesgue) and satisfies there $F'(x) = f(x)$.

Using this fact, we can prove the following standard result:

Proposition 7.27. *Let $f : \mathbb{R} \to \mathbb{R}$ be measurable on some nontrivial interval and have arbitrarily small periods, i.e. there is a sequence $T_n \downarrow 0$ with $f(x+T_n) = f(x)$ $(x \in \mathbb{R})$. Then f is almost everywhere constant, i.e. there is some $c \in \mathbb{R}$ with $f(x) = c$ for almost all $x \in \mathbb{R}$ (in the sense of Lebesgue).*

Proof. Without loss of generality, we may assume that $|f(x)| \leq 1$: Indeed, if a constant c as in the claim does not exist, then there is a constant c such that $A := \{x : f(x) \geq c\}$ and its complement both have positive measure. Put $g(x) :=$ 1 for $x \in A$ and $g(x) := 0$ for $x \notin A$. Then g is measurable with arbitrarily small periods, and $|g(x)| \leq 1$. If we can prove that g is a.e. constant, we have a contradiction.

Put $F(x) := \int_0^x f(t)\,dt$. We claim that $f(x) = F(T_1)/T_1$ for all x for which $F'(x)$ exists with $F'(x) = f(x)$ (these are almost all x by Proposition 7.26). By the definition of F', it suffices to prove that

$$\frac{F(x+T_n) - F(x)}{T_n} = T_n^{-1} \int_x^{x+T_n} f(t)\,dt \to T_1^{-1} F(T_1)$$

for all those x. But since T_n is a full period of f, we may replace x in the integral by any other number, and thus have to prove that

$$T_n^{-1} F(T_n) \to T_1^{-1} F(T_1). \tag{7.7}$$

Choosing $k_n \in \mathbb{N}$ such that $k_n T_n \leq T_1 < (k_n + 1)T_n$, we find

$$|F(k_n T_n) - F(T_1)| \leq \int_{k_n T_n}^{T_1} |f(t)|\,dt \leq T_1 - k_n T_n < T_n \to 0.$$

Since $|k_n T_n - T_1| < T_n^{-1} \to 0$, we may conclude that

$$\left| \frac{F(k_n T_n)}{k_n T_n} - \frac{F(T_1)}{T_1} \right| \to 0.$$

But the periodicity of f implies $F(k_n T_n) = k_n F(T_n)$, and so (7.7) could be proved. □

For the moment, we define $[x]$ for $x \in \mathbb{R}$ as the largest natural number which is not larger than x. Then for each natural number n, the number

$$f(x) := |[2^n x] - 2[2^{n-1} x]|$$

might be interpreted as the n-th digit (after the colon) of the binary expansion of x.

Theorem 7.28. *Put $h \in \mathbb{N}_\infty$, and*

$$f(x) := \operatorname{st}(|{}^*[2^{h*} x] - 2{}^*[2^{h-1*} x]|) \qquad (x \in \mathbb{R}).$$

Then $f : \mathbb{R} \to \{0, 1\}$ is nonmeasurable (in the sense of Lebesgue) on each nontrivial interval.

Proof. Put $g(n,x) := |[2^n x] - 2[2^{n-1}x]|$, and note that $f(x) = \mathrm{st}({}^*g(h, {}^*x))$. Since $g : \mathbb{N} \times \mathbb{R} \to \{0,1\}$, we have ${}^*g : {}^*\mathbb{N} \times {}^*\mathbb{R} \to \{{}^*0, {}^*1\}$, and so $f : \mathbb{R} \to \{0,1\}$. The transfer of the statement

$$\forall \underline{k}, \underline{n} \in \mathbb{N} : (\underline{k} < \underline{n} \implies \forall \underline{x} \in \mathbb{R} : g(\underline{n}, \underline{x} + 2^{-\underline{k}}) = g(\underline{n}, \underline{x}))$$

implies that ${}^*g(h, \cdot)$ is periodic with 2^{-k} as a period for any $k < h$. In particular, f has arbitrarily small periods. Thus, if f were measurable, Proposition 7.27 would imply that we have either $f(x) = 0$ for almost all x, or $f(x) = 1$ for almost all x. In particular, one of the sets $A := \{x \in [0,1] : f(x) = 0\}$ and $B := \{x \in [0,1] : f(x) = 1\}$ has measure 0, and the other has measure 1. We prove that A and B have the same measure and thus find a contradiction.

To see this, let $\alpha(\underline{x})$ for $\underline{x} \in [0,1]$ be the predicate "$\underline{x}$ is not dyadic", i.e.

$$\forall \underline{n}, \underline{k} \in \mathbb{N} : \underline{x} \neq \underline{k} 2^{-\underline{n}}.$$

Since $g(n,x)$ is the n-th number of the binary expansion of x, we have

$$\forall \underline{x} \in [0,1], \underline{n} \in \mathbb{N} : \alpha(\underline{x}) \implies g(\underline{n}, 1 - \underline{x}) + g(\underline{n}, \underline{x}) = 1.$$

The transfer implies that $g(h,x) = 1 - g(h, 1-x)$ for all $x \in {}^*[0,1]$ with ${}^*\alpha(x)$. If $x \in [0,1]$ is not dyadic, i.e. $\alpha(x)$ holds, then the transfer principle implies ${}^*\alpha({}^*x)$, and so $f(x) = 1 - f(1-x)$. Thus, $x \in A$ if and only if $1 - x \in B$. Since the dyadic numbers are countable and thus form a null set, we may conclude that A and B have the same measure, as claimed. □

Theorem 7.28 is taken from [Lux73]. Another example of a nonmeasurable function is given by

$$f(x) = \mathrm{st}({}^*\sin(2\pi h x)) \qquad (x \in \mathbb{R}) \tag{7.8}$$

with an appropriate $h \in \mathbb{N}_\infty$, see e.g. [SL76, Example 8.4.45] (see also [Tay69]). The measurability of (7.8) in dependence of h is discussed in [BH02, Tay69].

Exercise 41. Let $\mathscr{U}$ be a free ultrafilter over $\mathbb{N}$. Apply Theorem 7.28 to prove that the set

$$\{x \in [0,1] : \lim_{n \to \mathscr{U}} ([2^n x] - 2[2^{n-1}x]) = 0\}$$

is nonmeasurable (in the sense of Lebesgue).

It should be noted that the statement of Exercise 41 holds without the (uncountable) axiom of choice. This is of some interest, because it means that the mere existence of a free ultrafilter over $\mathbb{N}$ (which is less restrictive than the axiom of choice) implies the existence of a nonmeasurable set (recall that the classical proofs on the existence of nonmeasurable sets require a more powerful form of the axiom of choice). This fact was first observed by Sierpinski [Sie38] (Sierpinski used standard arguments, of course).

Theorem 7.28 has another interesting consequence:

Recall that a map $\|\cdot\| : X \to [0,\infty)$ on a linear space (=vector space) X over $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$ is called a *norm*, if the following holds:

1. $\|x\| = 0$ if and only if $x = 0$.
2. $\|\lambda x\| = |\lambda| \, \|x\|$ for scalars $\lambda \in \mathbb{K}$.
3. $\|x + y\| \le \|x\| + \|y\|$ (the *triangle inequality*).

Theorem 7.29. *Let ℓ_∞ denote the space of all bounded sequences with the natural operations. There is a norm $\|\cdot\|$ on ℓ_∞ which is additionally monotone (i.e. if $|x_n| \le |y_n|$, then $\|(x_n)_n\| \le \|(y_n)_n\|$) and a measurable function $f : \mathbb{N} \times \mathbb{R} \to \mathbb{R}$ such that*

$$x \mapsto \|f(\cdot, x)\|$$

is nonmeasurable on any nontrivial interval.

Proof. Fix some $h \in \mathbb{N}_\infty$, and define the norm by the formula

$$\|(x_n)_n\| = \left(\sup_{n \in \mathbb{N}} |x_n| \right) + |\mathrm{st}({}^*x_h)| .$$

It is easily checked that this indeed provides a norm (the first term is only needed to have that $\|(x_n)_n\| = 0$ implies $x_n = 0$ for all n). For $f(n,x) := [2^n x] - 2[2^{n-1}x]$, we have

$$\|f(\cdot, x)\| = 1 + |\mathrm{st}([2^n x] - 2[2^{n-1}x])| = 1 + \mathrm{st}(|[2^n x] - 2[2^{n-1}x]|)$$

which is nonmeasurable by Theorem 7.28. □

Theorem 7.29 answers a problem in the theory of ideal spaces which was open for a long time. It was proved (by a slightly different argument but with the same idea) in [Lux63]. Using a similar argument as in Exercise 41, one can give a formula for a norm with the property of Theorem 7.29 which does not involve nonstandard expressions, namely

$$\|(x_n)_n\| = \left(\sup_{n \in \mathbb{N}} |x_n| \right) + |\lim_{n \to \mathscr{U}} x_n| ,$$

where $\mathscr{U}$ denotes a free ultrafilter over $\mathbb{N}$. (Note that in view of Theorem 5.30, the limit in this expression always exists when x_n is a bounded sequence).

Chapter 4

Enlargements and Saturated Models

§8 Enlargements, Saturation, and Concurrency

Throughout this section, let $* : \widehat{S} \to \widehat{{}^*S}$ be elementary.

Definition 8.1. Let $* : \widehat{S} \to \widehat{{}^*S}$ be elementary. Then $*$ is called:

1. κ-*enlargement* (with some set κ) if for any nonempty system $\mathscr{A}$ of entities $A \in \widehat{S}$ which has the finite intersection property and at most the cardinality of κ, we have $\bigcap {}^\sigma \mathscr{A} \neq \varnothing$, i.e.
$$\bigcap \{ {}^*A : A \in \mathscr{A} \} \neq \varnothing.$$
2. *enlargement* if it is a κ-enlargement for any κ (i.e. if the above condition holds without any assumption on the cardinality of $\mathscr{A}$).
3. κ-*saturated* (with some set κ) if for any nonempty system $\mathscr{B}$ of internal entities which has the finite intersection property and at most the cardinality of κ, we have
$$\bigcap \mathscr{B} \neq \varnothing.$$
4. *polysaturated* if it is $\widehat{S}$-saturated.

The reader should be warned that the definitions of κ-enlargement, κ-saturated and polysaturated differ slightly in literature: Usually, κ denotes a cardinal number, and one requires that the cardinality of $\mathscr{A}$ resp. $\mathscr{B}$ be strictly *less* than κ (e.g. in [CK90, SL76]). Thus, e.g. what we call $\mathbb{N}$-saturated, is in literature usually called $\aleph_1$-*saturated* where $\aleph_1$ denotes the first uncountable ordinal. Moreover, in e.g. [SL76] it is required that $\mathscr{B}$ itself be an *internal* entity. (We shall see, however, that it makes actually no difference if we would require that $\mathscr{A}$ and $\mathscr{B}$ be entities). The above definition of polysaturated maps is used in [LR94], where also enlargements are called “strong nonstandard embeddings”.

Proposition 8.2. *If $*$ is κ-saturated, it is a κ-enlargement.*

Proof. If $\mathscr{A}$ is as in Definition 8.1, then ${}^{\sigma}\mathscr{A}$ is a system of internal entities (Proposition 3.16) with the finite intersection property which has the same cardinality as $\mathscr{A}$. Hence $\bigcap {}^{\sigma}\mathscr{A} \neq \varnothing$. □

We shall see later that for any set S and any κ one can find κ-saturated maps (and thus also κ-enlargements).

It looks rather non-symmetric that for the definition of enlargements no restriction on the cardinality on $\mathscr{A}$ is made while for the definition of polysaturated maps a restriction is made. However, for enlargements, this restriction is implicit, since (see Lemma 8.7 below), one may assume that $\mathscr{A} \in \widehat{S}$. Hence, each $\widehat{S}$-enlargement is automatically an enlargement. In particular, each polysaturated map is an enlargement.

It would not make sense to drop the assumptions on the cardinality of $\mathscr{B}$ in the definition of polysaturated maps, since no such maps can exist:

Proposition 8.3. *If $A \in \widehat{S}$ is an infinite entity, then there is a nonempty system $\mathscr{B}$ of internal subsets of ${}^*A$ with the finite intersection property and $\bigcap \mathscr{B} = \varnothing$.*

Proof. Let $\mathscr{B}$ be the system of all sets of the form $B_b = \{\underline{x} \in {}^*A : \underline{x} \neq b\}$ ($b \in {}^*A$). Each B_b is internal by the internal definition principle (more precisely, by Corollary 3.18), and the other claims are evident. □

Corollary 8.4. *If S is infinite and $*$ is κ-saturated, then κ has at most the cardinality of ${}^*S$. In particular, there is no map $*$ which is κ-saturated for any κ.* □

This does of course not exclude that for any κ we can find a κ-saturated map $*$ (and, as remarked above, such maps indeed exist): Corollary 8.4 only implies that $*$ then must depend on κ.

We already see that it cannot be too easy to find saturated maps: If we use the construction of §4, then ${}^*S$ consists of the equivalence classes of maps $x : J \to S$. In particular, the cardinality of ${}^*S$ is at most the cardinality of S^J. We thus need that S^J has a larger cardinality than $\widehat{S} \supseteq S \cup \mathcal{P}(S) \cup \mathcal{P}(\mathcal{P}(S)) \cup \cdots$. Thus, the cardinality of J must be rather large (in particular, the choice $J := \mathbb{N}$ is never sufficient). Hence, there exists a large class of nonstandard maps which is not polysaturated.

Theorem 8.5. *Let S be infinite. Then $*$ is an $\mathbb{N}$-enlargement if and only if $*$ is a nonstandard embedding.*

In particular, each enlargement (and so each polysaturated map) is a nonstandard embedding.

Proof. Let $*$ be an $\mathbb{N}$-enlargement, and $B \subseteq S$ be infinite countable. To see that $*$ is a nonstandard embedding, it suffices by Theorem 3.22 to prove that ${}^\sigma B \neq {}^*B$. Let $\mathscr{A}$ be the system of all sets of the form $B \setminus \{b\}$ $(b \in B)$. Since B is infinite, $\mathscr{A}$ has the finite intersection property. Moreover, $\mathscr{A}$ is countable (since B is countable). Hence, $\bigcap\{{}^*A : A \in \mathscr{A}\} \neq \varnothing$. Since ${}^*(B \setminus \{b\}) = {}^*B \setminus \{{}^*b\}$, this means ${}^*B \setminus {}^\sigma B = \bigcap_{b \in B}({}^*B \setminus \{{}^*b\}) \neq \varnothing$, and so $*$ is a nonstandard embedding.

Conversely, assume that $*$ is a nonstandard embedding. We show that $*$ is an $\mathbb{N}$-enlargement. Thus, let $\mathscr{A}$ denote a nonempty countable system of entities $A \in \widehat{S}$ which has the finite intersection property. Let $A_1, A_2, \ldots$ be an enumeration of all elements of $\mathscr{A}$. By an obvious identification, we may assume that $\mathbb{N} \subseteq S$ (just rename the atoms). Define a function $f : \mathbb{N} \to \mathcal{P}(A_1)$ by $f(n) := A_1 \cap \cdots \cap A_n$. Note that Theorem 2.1 implies $\mathcal{P}(A_1) \in \widehat{S}$ and $f \in \widehat{S}$. Since $\mathscr{A}$ has the finite intersection property, we have $f(n) \neq \varnothing$ for each n, i.e. $\forall \underline{x} \in \mathbb{N} : f(\underline{x}) \neq \varnothing$. The transfer principle implies $\forall \underline{x} \in {}^*\mathbb{N} : {}^*f(\underline{x}) \neq \varnothing$. Since $*$ is a nonstandard embedding, there is some $h \in {}^*\mathbb{N} \setminus {}^\sigma\mathbb{N}$, and we have ${}^*f(h) \neq \varnothing$. If we can prove that ${}^*f(h)$ is contained in ${}^*A_n$ for each $n \in \mathbb{N}$, it follows that the intersection of the sets ${}^*A_n$ $(n \in \mathbb{N})$ is nonempty, and so that $*$ is an $\mathbb{N}$-enlargement. To prove ${}^*f(h) \subseteq {}^*A_n$, let $n \in \mathbb{N}$ be given. The transfer of the true sentence $\forall \underline{x} \in \mathbb{N} : (\underline{x} > n \implies f(\underline{x}) \subseteq A_n)$ reads $\forall \underline{x} \in {}^*\mathbb{N} : (\underline{x} > {}^*n \implies {}^*f(\underline{x}) \subseteq {}^*A_n)$. Now we note that $h > {}^*n$ by Proposition 5.9, and so ${}^*f(h) \subseteq {}^*A_n$, as claimed. $\square$

It is in general not true that each nonstandard embedding is $\mathbb{N}$-saturated. There exist even enlargements which are not $\mathbb{N}$-saturated, see [CK90, Exercise 4.4.29]. However, these examples are rather "exotic". All elementary embeddings that we discuss are "nicer": We will see that all embeddings arising from the ultrapower construction of §4 are $\mathbb{N}$-saturated. To see this, we need some auxiliary notions.

If $f : {}^\sigma A \to {}^*B$ is a function where $A \in \widehat{S}$ is an infinite entity, then f cannot be an internal function, since otherwise $\operatorname{dom}(f) = {}^\sigma A$ were internal by Theorem 3.19, contradicting Theorem 3.22. One may ask whether f is just the restriction of an internal function to the external set ${}^\sigma A$. If this is the case for *any* function, we call $*$ *comprehensive*:

Definition 8.6. An elementary map $* : \widehat{S} \to \widehat{{}^*S}$ is called *comprehensive*, if for each entities $A, B \in \widehat{S}$ and each function $f : {}^\sigma A \to {}^*B$, there is an internal function $F : {}^*A \to {}^*B$ such that $F(a) = f(a)$ for each $a \in {}^\sigma A$.

We will see in §9 that all embeddings arising from the ultrapower construction of §4 are comprehensive. We intend to prove now a relation of comprehensive and saturated maps. In particular, we will show that all comprehensive maps in turn are $\mathbb{N}$-saturated and also a partial converse. To this end, we need a result which is of independent interest:

Lemma 8.7. *To verify that a map $*$ is a κ-enlargement resp. κ-saturated, it suffices to consider systems $\mathscr{A}$ resp. $\mathscr{B}$ in Definition* 8.1 *which additionally are entities in $\widehat{S}$ resp. $\widehat{{}^*S}$. Moreover, it additionally suffices to consider systems $\mathscr{B}$ which are subsets of standard entities.*

Proof. Indeed, if $\mathscr{A}$ is as in Definition 8.1, fix some $A_0 \in \mathscr{A}$ and consider the set $\mathscr{A}_0 := \{A \cap A_0 : A \in \mathscr{A}\}$ in place of $\mathscr{A}$: Then $\mathscr{A}_0 \in \mathcal{P}(A_0)$ is an entity by Theorem 2.1. Moreover, $\bigcap^\sigma \mathscr{A}_0 \subseteq \bigcap^\sigma \mathscr{A}$, since ${}^*(A \cap A_0) \subseteq {}^*A$ for each $A \in \mathscr{A}$ (Lemma 3.5). It thus suffices to verify that $\bigcap^\sigma \mathscr{A}_0 \neq \varnothing$. Now the first statement follows if we observe that $\mathscr{A}_0$ has the finite intersection property and at most the cardinality of $\mathscr{A}$.

Concerning κ-saturated maps, we first argue similarly: If $\mathscr{B}$ is given as in Definition 8.1, fix some $B_0 \in \mathscr{B}$ and consider the set $\mathscr{B}_0 := \{B \cap B_0 : B \in \mathscr{B}\}$. Then $\mathscr{B}_0$ consists of internal entities (Theorem 3.19) with the finite intersection property whose cardinality is not larger than $\mathscr{B}$. Since $\bigcap \mathscr{B}_0 \subseteq \bigcap \mathscr{B}$, it suffices to verify that $\bigcap \mathscr{B}_0 \neq \varnothing$. Now observe that $\mathscr{B}_0$ consists only of internal subsets of B_0. The system P of *all* internal subsets of B_0 is internal (Exercise 81), and so $\mathscr{B}_0 \subseteq P$. Proposition 3.16 implies that there is some n with $P \in {}^*S_n$ and $P \subseteq {}^*S_n$. Consequently, $\mathscr{B}_0 \subseteq {}^*S_n$. Since ${}^*S_n$ is an entity of $\widehat{{}^*S}$, this inclusion implies also that $\mathscr{B}_0$ is an entity of $\widehat{{}^*S}$ (Theorem 2.1). □

Theorem 8.8. *Let S be infinite. If $*$ is a comprehensive nonstandard embedding, then $*$ is $\mathbb{N}$-saturated. Conversely, if $*$ is polysaturated, then it is comprehensive.*

More precisely, if $*$ is κ-saturated, then for any internal entities A, B and for each function $f : A_0 \to B$ with $A_0 \subseteq A$ where A_0 has at most the cardinality of κ, there is an internal function $F : A \to B$ such that $F(a) = f(a)$ for each $a \in A_0$.

Proof. We start to prove the last claim. Thus, let $*$ be κ-saturated, A, B be internal entities, and $f : A_0 \to B$ with $A_0 \subseteq A$ where A_0 has at most the cardinality of κ. Let $\mathscr{F}$ denote the system of all internal functions $g : {}^*A \to B$. Recall that $\mathscr{F}$ is internal by Exercise 82. Consider now the family $\mathscr{B}$ of sets

$$B_a = \{\underline{x} \in \mathscr{F} : \underline{x}(a) = f(a)\} \qquad (a \in A_0).$$

By the internal definition principle, each set B_a is internal. Moreover, the system $\mathscr{B}$ has the finite intersection property: Indeed, an element of $B_{a_1} \cap \cdots \cap B_{a_n}$ can be defined in view of Exercise 8 by redefining *some* function from $\mathscr{F}$ at the finitely many points $a_1, \ldots, a_n \in A_0$. Since $\mathscr{B}$ has at most the cardinality of A_0, we find some $F \in \bigcap \mathscr{B}$. Then $F \in \mathscr{F}$ satisfies $F(a) = f(a)$ for each $a \in A$.

If $*$ is polysaturated, and $A, B \in \widehat{S}$ and $f : {}^\sigma A \to {}^*B$ are given, then ${}^*A, {}^*B$ are internal by Proposition 3.16, and ${}^\sigma A \subseteq {}^*A$. Since ${}^\sigma A$ has the cardinality of A which has a strictly smaller cardinality than $\widehat{S}$, we may conclude by what we proved above that an internal function $F : {}^*A \to {}^*B$ exists which satisfies $F(x) = f(x)$ for each $x \in {}^\sigma A$. Hence, $*$ is comprehensive.

Conversely, let $*$ be comprehensive. Let $\mathscr{B}$ be a countable system of internal entities with the finite intersection property. By Lemma 8.7, we may assume that $\mathscr{B} \subseteq {}^*C$ for some entity $C \in \widehat{S}$. Let $B_1, B_2, \dots$ be an enumeration of all elements of $\mathscr{B}$. Since S is infinite, we may assume that $\mathbb{N} \subseteq S$ (rename the atoms), and define a map $f : {}^\sigma\mathbb{N} \to {}^*C$ by $f({}^*n) = B_n$. Since $*$ is comprehensive, we thus find an internal map $F : {}^*\mathbb{N} \to {}^*C$ with $F({}^*n) = B_n$ for each $n \in \mathbb{N}$. Note that $U := \bigcup {}^*C$ is internal by Theorem 3.19. The set

$$G := \{\underline{x} \in {}^*\mathbb{N} \mid \exists \underline{y} \in U : \forall \underline{z} \in {}^*\mathbb{N} : (\underline{z} \le \underline{x} \implies \underline{y} \in F(\underline{z}))\}$$

is internal by the internal definition principle. For any $n \in \mathbb{N}$, we have ${}^*n \in G$, since $\mathscr{B}$ has the finite intersection property and thus $F(1) \cap \dots \cap F({}^*n) \neq \varnothing$ (recall Proposition 5.9). Hence, by the permanence principle, G also contains some $h \in {}^*\mathbb{N} \setminus {}^\sigma N$. This means that there is some $y \in U$ such that

$$y \in \bigcap \{F(\underline{z}) : \underline{z} \in {}^*\mathbb{N} \wedge \underline{z} \le h\} \subseteq \bigcap \{F(\underline{z}) : \underline{z} \in {}^\sigma\mathbb{N}\} = \bigcap \mathscr{B}.$$

Thus, $\bigcap \mathscr{B} \neq \varnothing$, and $*$ is $\mathbb{N}$-saturated. □

Let us now come to the point why enlargements and polysaturated maps are of particular interest.

Definition 8.9. Let φ be a binary relation. We say that φ is *satisfied by* $b \in \operatorname{rng}(\varphi)$ on $A \subseteq \operatorname{dom}(\varphi)$, if $(a, b) \in \varphi$ for each $a \in A$.

We call φ *concurrent* on $A \subseteq \operatorname{dom}(\varphi)$, if for each finite subset $A_0 \subseteq A$ there is some $b \in \operatorname{rng}(\varphi)$ which satisfies φ on A_0.

In other words: φ is concurrent on A, if for each finitely many $a_1, \dots, a_n \in A$ there is some b with $(a_1, b), \dots, (a_n, b) \in \varphi$.

Let us first discuss the connection of enlargements and concurrent binary relations:

Theorem 8.10. *The following statements are equivalent for an elementary map $*$:*

1. *$*$ is a κ-enlargement.*
2. *For any binary relation $\varphi \in \widehat{S}$ for which* $\operatorname{dom}(\varphi)$ *has at most the cardinality of κ the following holds: If φ is concurrent on* $\operatorname{dom}(\varphi)$*, then there is some $b \in \operatorname{rng}({}^*\varphi)$ which satisfies ${}^*\varphi$ on the set ${}^\sigma\operatorname{dom}(\varphi)$.*

3. *For any entity $A \in \widehat{S}$ which has at most the cardinality of κ there is some *-finite entity B with ${}^\sigma A \subseteq B \subseteq {}^*A$.*

Proof. Let $*$ be a κ-enlargement, and φ be a binary standard relation such that $\mathrm{dom}(\varphi)$ has at most the cardinality of κ. Let $\mathscr{A}$ be the system of all sets of the form

$$A_d := \{\underline{y} \in \mathrm{rng}(\varphi) \mid (d, \underline{y}) \in \varphi\}$$

where $d \in \mathrm{dom}(\varphi)$. If φ is concurrent on $\mathrm{dom}(\varphi)$, then $\mathscr{A}$ has the finite intersection property. Since $\mathscr{A}$ has at most the cardinality of $\mathrm{dom}(\varphi)$, the set $\bigcap_{d \in \mathrm{dom}(\varphi)} {}^*A_d$ contains some b. By the standard definition principle, we have

$${}^*A_d = \{\underline{y} \in {}^*\mathrm{rng}(\varphi) \mid ({}^*d, \underline{y}) \in {}^*\varphi\},$$

i.e. $({}^*d, b) \in {}^*\varphi$ for any $d \in \mathrm{dom}(\varphi)$, i.e. b satisfies ${}^*\varphi$ on ${}^\sigma\mathrm{dom}(\varphi)$. Thus 2. holds.

To prove that 2. implies 3., let $A \in \widehat{S}$ be an entity which has at most the cardinality of κ. Consider the relation

$$\varphi := \{(\underline{x}, \underline{y}) \in A \times \mathcal{P}(A) \mid \underline{x} \in \underline{y} \wedge \text{“}\underline{y} \text{ is finite”}\}.$$

Then φ is concurrent on A. Thus, 2. implies that we find some $B \in \mathrm{rng}({}^*\varphi)$ such that $({}^*a, B) \in {}^*\varphi$ for any $a \in A$. Note that, by the standard definition principle for relations,

$${}^*\varphi = \{(\underline{x}, \underline{y}) \in {}^*A \times {}^*\mathcal{P}(A) \mid \underline{x} \in \underline{y} \wedge \text{“}\underline{y} \text{ is } {}^*\text{-finite”}\}.$$

Hence, B is a *-finite element of ${}^*\mathcal{P}(A)$ (i.e. an internal subset of ${}^*A$) which contains ${}^\sigma A$.

Assume now that 3. holds. Then $*$ is a κ-enlargement. Indeed, let $\mathscr{A}$ be a nonempty system of entities $A \in \widehat{S}$ which has the finite intersection property and at most the cardinality of κ. By Lemma 8.7, we may assume that $\mathscr{A}$ is an entity. Then 3. implies that there is some *-finite entity $\mathscr{B}$ with ${}^\sigma\mathscr{A} \subseteq \mathscr{B} \subseteq {}^*\mathscr{A}$. Since $\mathscr{A}$ has the finite intersection property, the transitively bounded sentence

$$\forall \underline{x} \in \mathcal{P}(\mathscr{A}) : (\text{“}\underline{x} \text{ is finite”} \implies \bigcap \underline{x} \neq \varnothing)$$

is true. Its $*$-transform reads

$$\forall \underline{x} \in {}^*\mathcal{P}(\mathscr{A}) : (\text{“}\underline{x} \text{ is } {}^*\text{-finite”} \implies \bigcap \underline{x} \neq \varnothing).$$

Since $\mathscr{B} \in {}^*\mathcal{P}(\mathscr{A})$ is *-finite, we have $\bigcap \mathscr{B} \neq \varnothing$. Hence, ${}^\sigma\mathscr{A} \subseteq \mathscr{B}$ implies $\bigcap {}^\sigma\mathscr{A} \supseteq \bigcap \mathscr{B} \neq \varnothing$. $\square$

Thus, enlargements mean, roughly speaking, that whenever we can satisfy in the standard universe each finite number of conditions, we can in the nonstandard universe satisfy *all* conditions simultaneously.

The reader should observe the analogy with the permanence principle: If something holds for all finite sets, then it also holds for all *-finite sets, and thus in particular for an infinite set (actually, this is what we have used in the proof of Theorems 8.5 and 8.8). The difference between the permanence principle and enlargements is that for enlargements we are not forced to the special role of $\mathbb{N}$ but may instead use any other (standard) set as the index set.

Exercise 42. Prove directly that the property 2. of Theorem 8.10 implies that $*$ is a κ-enlargement.

The following consequence is in a sense a generalization of Corollary 8.4 for enlargements.

Proposition 8.11. *If $*$ is a κ-enlargement, then $^*\mathbb{N}$ has at least the cardinality of either $\widehat{S}$ or κ (whichever is smaller).*

Proof. If $A \in \widehat{S}$ is an entity which has at most the cardinality of κ, then there is some *-finite $B \subseteq {}^*A$ with ${}^\sigma A \subseteq B$. Since B is in one-to-one correspondence with some set $\{1, \ldots, h\} \subseteq {}^*\mathbb{N}$, it follows that $^*\mathbb{N}$ has at least the cardinality of A.

Now we distinguish two cases: If there is some n such that S_n has at least the cardinality of κ, choose some $A \subseteq S_n$ which has precisely the cardinality of κ (axiom of choice!), and by what we just proved, $|{}^*\mathbb{N}| \geq |\kappa|$. If there is no such entity, then we have $|{}^*\mathbb{N}| \geq |S_n|$ for each n which implies in view of $S_0 \subseteq S_1 \subseteq \cdots$ that $|{}^*\mathbb{N}| \geq |\bigcup S_n| = |\widehat{S}|$. □

Proposition 8.11 was first observed in [Zak69].

Exercise 43. Let S be infinite. Prove that the ultrapower model of §4 with $J = \mathbb{N}$ does not provide an enlargement. Show more precisely that if κ has a strictly larger cardinality than $\mathcal{P}(\mathbb{N})$, then $*$ is not a κ-enlargement.

For applications in topology, the following compactness property of enlargements is crucial:

Exercise 44. Let $*$ be a κ-enlargement. Then for any system $\mathscr{A}$ of entities $A \in \widehat{S}$ which has at most the cardinality of κ and any standard ${}^*A_0 \subseteq \bigcup {}^\sigma\mathscr{A}$ there is a finite $\mathscr{A}_0 \subseteq \mathscr{A}$ with

$$A_0 \subseteq \bigcup \mathscr{A}_0.$$

Exercise 45. Let $\mathbb{R} \in \widehat{S}$ be an entity, and $* : \widehat{S} \to \widehat{{}^*S}$ be an $\mathbb{R}$-enlargement. Prove that there is a number $h \in {}^*\mathbb{N}$ such that

$$^*\sin(\pi h x) \approx 0 \qquad (x \in {}^\sigma\mathbb{R}).$$

In particular, there is a number $h \in {}^*\mathbb{N}$ such that (7.8) is constant and thus measurable on $\mathbb{R}$.

Hint: Use without proof that for each finitely many $x_1, \ldots, x_n \in \mathbb{R}$ and any $\varepsilon \in \mathbb{R}_+$ there is some $h \in \mathbb{N}$ such that the distance of hx_k to an integer is at most ε (for any k).

Actually, the statement of Exercise 45 holds if $*$ is just an arbitrary nonstandard map [Tay69], but the proof is harder for this case.

The fact that ${}^*\varphi$ is not satisfied on $\mathrm{dom}({}^*\varphi) = {}^*\mathrm{dom}(\varphi)$ in Theorem 8.10 but only on ${}^\sigma\mathrm{dom}(\varphi)$ is rather disappointing. If we consider instead of κ-enlargements even κ-saturated maps, we do not have this restriction. Moreover, φ can even be an internal relation:

Theorem 8.12. *The following statements are equivalent for an elementary map* $*$:

1. $*$ *is* κ*-saturated.*
2. *For any (not necessarily internal) binary relation* φ *and any (not necessarily internal)* $A \subseteq \mathrm{dom}(\varphi)$ *which has at most the cardinality of* κ *and for which each of the sets*

$$\varphi(a) := \{\underline{y} \in \mathrm{rng}(\varphi) : (a, \underline{y}) \in \varphi\} \qquad (a \in A)$$

is internal, the following holds: If φ *is concurrent on* A*, then it is satisfied on* A*.*

Proof. Let $*$ be κ-saturated, φ be a binary relation, and $A \subseteq \mathrm{dom}(\varphi)$ have at most the cardinality of κ. Let $\mathscr{B} := \{\varphi(a) : a \in A\}$. If φ is concurrent on A, then $\mathscr{B}$ has the finite intersection property. Since $*$ is κ-saturated and $\mathscr{B}$ has at most the cardinality of A, the set $\bigcap \mathscr{B}$ contains some b, if $\mathscr{B}$ consists only of internal sets. This means that $(a, b) \in \varphi$ for any $a \in A$, i.e. b satisfies φ on A.

Conversely, let 2. be satisfied, and $\mathscr{B}$ be a nonempty system of internal entities which has the finite intersection property and at most the cardinality of κ. Put

$$\varphi := \{(\underline{x}, \underline{y}) : \underline{y} \in \underline{x} \in \mathscr{B}\}.$$

Since $\mathscr{B}$ has the finite intersection property, φ is concurrent on $\mathscr{B}$. Moreover for any $B \in \mathrm{dom}(\varphi) = \mathscr{B}$, the set

$$\{\underline{y} \in \mathrm{rng}(\varphi) : (B, \underline{y}) \in \varphi\} = B$$

is internal. Hence, the assumption implies that φ is satisfied on $\mathscr{B}$, i.e. there is some b with $(B, b) \in \varphi$ for any $B \in \mathscr{B}$, i.e. $b \in \bigcap \mathscr{B}$. $\square$

Corollary 8.13. *If* $*$ *is* κ*-saturated, then for any internal binary relation* φ *for which* $\operatorname{dom}(\varphi)$ *has at most the cardinality of* κ*, we have: If* φ *is concurrent on* $\operatorname{dom}(\varphi)$*, then it is satisfied on* $\operatorname{dom}(\varphi)$*.*

Proof. The sets $\varphi(b)$ in Theorem 8.12 are internal by the internal definition principle. □

The difference between Theorem 8.12 and Corollary 8.13 corresponds to the different definitions of κ-saturated maps which can be found in literature (e.g. in [SL76]):

Exercise 46. Show that for elementary maps $*$ the property of Corollary 8.13 is equivalent to the fact that $*$ is "κ-saturated" in the sense that for any nonempty *internal* system $\mathscr{B}$ of entities which has the finite intersection property and at most the cardinality of κ, we have $\bigcap \mathscr{B} \neq \varnothing$.

Roughly speaking that a map $*$ is polysaturated means: Whenever it appears possible that an internal relation can be satisfied (because there are not finitely many elements witnessing the contrary), then it actually *is* satisfied (if the cardinality of the domain is not too large).

The following result is a generalization of the compactness property of enlargements (Exercise 44):

Exercise 47. Let $*$ be κ-saturated. Then for any system $\mathscr{A}$ of internal entities and any internal $A_0 \subseteq \bigcup \mathscr{A}$ there is some finite $\mathscr{A}_0 \subseteq \mathscr{A}$ with

$$A_0 \subseteq \bigcup \mathscr{A}_0.$$

There is a concept which is in between κ-enlargements and κ-saturated maps:

Definition 8.14. Let $* : \widehat{S} \to \widehat{{}^*S}$ be elementary. Then $*$ is called a *compact* κ*-enlargement* (with some set κ) if for any internal binary relation φ the following holds:

If φ is concurrent on a (not necessarily internal) entity A which has at most the cardinality of κ and the property that at most finitely many elements of A are nonstandard, then φ is satisfied on A.

The embedding $*$ is called a *compact enlargement* if it is a compact κ-enlargement for any κ (i.e. if the above condition holds without any assumption on the cardinality of A).

Proposition 8.15. *We have for each set* κ*:*

1. *Each* κ*-saturated map is a compact* κ*-enlargement, and each compact* κ*-enlargement is a* κ*-enlargement.*
2. *Each polysaturated map is a compact enlargement, and each compact enlargement is an enlargement.*

Proof. 1. The first statement follows from Theorem 8.12 by restricting $\operatorname{dom}(\varphi)$ to the set A. For the second statement, we apply Theorem 8.10: Let $\varphi \in \widehat{S}$ be a binary relation which is satisfied on $A := \operatorname{dom}(\varphi)$ where A has at most the cardinality of κ. Assume that φ is satisfied on A. Then ${}^*\varphi$ is a standard relation which is concurrent on ${}^\sigma A$. If $*$ is a compact κ-enlargement, then ${}^*\varphi$ is satisfied on ${}^\sigma A$ (because ${}^\sigma A$ contains only standard elements and has the same cardinality as A). Hence, Theorem 8.10 implies that $*$ is a κ-enlargement.
2. Let $*$ be polysaturated, and φ be an internal binary relation which is concurrent on an entity A which has at most finitely many nonstandard elements. Since we may assume that A is infinite (otherwise φ is trivially satisfied on A), also $\widehat{S}$ is infinite, and so A has at most the cardinality of $\widehat{S}$. By Theorem 8.12, it follows that φ is satisfied, and so $*$ is a compact enlargement. The second statement follows immediately from 1. □

There exist enlargements which are not compact enlargements (even ultrafilter models with this property exist [Lux69a]). For most practical purposes, compact enlargements are sufficient, and as we shall see, they are considerably easier to construct than saturated maps.

Compact enlargements also have a "finite subcovering property" which is "in between" the corresponding property for enlargements and for saturated maps (Exercise 44 resp. Exercise 47): In contrast to enlargements, the covered set A_0 may be internal (and need not be standard). But in contrast to saturated maps, the covering family must consist of standard sets (and not of internal sets).

Theorem 8.16. *Let $*$ be a compact κ-enlargement. Then for any system $\mathscr{A}$ of entities $A \in \widehat{S}$ which has at most the cardinality of κ and any internal $A_0 \subseteq \bigcup {}^\sigma\mathscr{A}$ there is a finite $\mathscr{A}_0 \subseteq \mathscr{A}$ with*

$$A_0 \subseteq \bigcup {}^\sigma\mathscr{A}_0.$$

Proof. Consider the internal relation

$$\varphi := \{(\underline{x}, \underline{y}) \in {}^*\mathscr{A} \times A_0 \mid \underline{y} \notin \underline{x}\}.$$

If there is no finite $\mathscr{A}_0 \subseteq \mathscr{A}$ with $A_0 \subseteq \bigcup {}^\sigma\mathscr{A}_0$, then φ is concurrent on ${}^\sigma\mathscr{A}$. Since ${}^\sigma\mathscr{A}$ has at most the cardinality of κ, it follows that φ is satisfied by some $a \in A_0$ on ${}^\sigma\mathscr{A}$, i.e. $a \notin \bigcup {}^\sigma\mathscr{A}$, a contradiction to our assumption $A_0 \subseteq \bigcup {}^\sigma\mathscr{A}$. □

§9 Saturated Models

In this section, we shall "construct" for any κ and any set S a model which provides a κ-saturated map $*$. (This of course implies that there exist polysaturated maps for any S). To this aim, we first construct enlargements and then consider a countable chain of enlargements (a so-called direct limit) which provides a compact enlargement. If we consider even an uncountable chain, we obtain a κ-saturated map. The model for the enlargement is a special ultrapower model. In contrast, the direct limit models are not ultrapowers.

The way we proceed is not the only possible construction of polysaturated maps. In fact, it is even possible to obtain polysaturated models as ultrapowers, see e.g. [LR94, Lux69a]. However, the construction of the latter depends on the existence of a certain type of ultrafilters which is extremely hard to prove. For this reason, we chose the direct limit construction. The latter is essentially due to W. A. J. Luxemburg [Lux69a, SL76].

However, for many applications, it suffices to have $\mathbb{N}$-saturated maps. For these applications, we need no additional considerations at all, since *each* nonstandard map $*$ as constructed in §4 by ultrapowers has this property:

Theorem 9.1. *Let $*$ be a map arising from the ultrapower construction of §4. Then $*$ is comprehensive. In particular, if S is infinite and $\mathscr{U}$ is δ-incomplete (i.e. $*$ is nonstandard), then $*$ is $\mathbb{N}$-saturated.*

Proof. The second statement follows from the first by Theorem 8.12. Let entities $A, B \in \widehat{S}$ and a function $f : {}^{\sigma}A \to {}^{*}B$ be given. Consider first the abstract model $\mathscr{S}$ from Section 4.2. Then ${}^{*}A$ corresponds to the set of all equivalence classes of maps $x : J \to A$, and ${}^{\sigma}A$ corresponds to the subset of classes which have a constant function as their representative (we write in abuse of notation that $[a]$ is the equivalence class of the constant function with value a). Similarly, ${}^{*}B$ corresponds to the set of all classes of functions $x : J \to B$. Hence, for each $a \in A$, the value $f([a])$ corresponds to the equivalence class of some function $f_a : J \to B$. Consider now the function $F_0 : J \to B^A$, defined by $(F_0(j))(a) := f_a(j)$. By Proposition 4.19, the equivalence class $[F_0]$ is mapped by φ into an internal element F. By Proposition 4.19, we have $F \in {}^{*}(B^A)$, i.e. F is an internal function from ${}^{*}A$ into ${}^{*}B$ (Theorem 3.21). To prove that F is an extension of f, we have to prove that $({}^{*}a, f({}^{*}a)) \in F$ is a true sentence for any given $a \in A$. By Theorem 4.18, this sentence is true if and only if $([a], [f_a]) \in_{\mathscr{U}} [F_0]$ where the shortcut for pairing is to be understood in the interpretation of Theorem 4.18. By the Łoś/Luxemburg Theorem 4.14, this is equivalent to $(a, f_a(j)) \in F_0(j)$ for almost all j which in turn is equivalent to $(F_0(j)) = f_a(j)$ for almost all j. The latter is true by construction, and so we have indeed $F({}^{*}a) = f({}^{*}a)$ for any given $a \in A$. $\square$

9.1 Models for Enlargements

We already know (Exercise 43) that not any ultrapower provides an enlargement. However, for certain choices of J and an appropriate ultrafilter $\mathscr{U}$ over J, we obtain enlargments. More, precisely, we get enlargements if we choose a so-called λ-adequate filter:

Definition 9.2. Let λ be a set. A filter $\mathscr{F}$ over J is called λ-*adequate*, if for each nonempty family $\mathscr{A}$ of subsets of λ with the finite intersection property there is a map $f : J \to \lambda$ such that for each $A \in \mathscr{A}$ there is some $F \in \mathscr{F}$ with $f(F) \subseteq A$.

Theorem 9.3. *For each set λ there exists a λ-adequate ultrafilter $\mathscr{U}$ on an appropriate set J.*

Proof. Let J be the system of all finite collections of subsets from λ, i.e. J consists of all finite subsets of $\mathcal{P}(\lambda)$. For each $A \subseteq \lambda$ define a subset of J by

$$F_A := \{j \in J : A \in j\}.$$

The collection $\mathscr{F}_0$ of all sets F_A (i.e. $\mathscr{F}_0 := \{F_A : A \subseteq J\}$), has the finite intersection property. Indeed, if $A_1, \ldots, A_n \subseteq \lambda$, then $j := \{A_1, \ldots, A_n\}$ belongs to J, and the intersection $F_{A_1} \cap \cdots \cap F_{A_n}$ contains j. Hence, $\mathscr{F}_0$ generates a filter $\mathscr{F}$ (recall Proposition 4.5). By Theorem 4.9 (axiom of choice!), there exists an ultrafilter $\mathscr{U}$ on J containing $\mathscr{F}$ and thus $\mathscr{F}_0$.

We claim that $\mathscr{U}$ is λ-adequate. Thus, let $\mathscr{A}$ be a nonempty family of subsets of λ with the finite intersection property. Recall that each $j \in J$ is a finite collection of subsets of λ. Let J_0 be the subset of those collections which contain only elements of $\mathscr{A}$. Since $\mathscr{A}$ has the finite intersection property, each of the sets

$$B_j := \bigcap_{A \in j} A \qquad (j \in J_0)$$

is nonempty, i.e. it contains some $f(j)$. For $j \notin J_0$, we define $f(j)$ arbitrarily.

Considering f as a function (axiom of choice!), we claim that $f : J \to \lambda$ has the required property: If $A \in \mathscr{A}$, then $F_A \in \mathscr{F}_0 \subseteq \mathscr{F} \subseteq \mathscr{U}$. We claim that $f(F_A) \subseteq A$. Indeed, for any $j \in F_A$ we have $A \in j$, and so $f(j) \in B_j \subseteq A$. $\square$

Theorem 9.4. *If $\mathscr{U}$ is an $\widehat{S}$-adequate filter, then the ultrafilter model of §4 provides an enlargement.*

Proof. Let $*$ be the corresponding map. Let $\mathscr{A}$ be a nonempty system of entities $A \in \widehat{S}$ which has the finite intersection property. We have to prove that there is some internal element a which is contained in each of the sets ${}^*A$ where $A \in \mathscr{A}$.

Recall that we have $x \in {}^*A$ if and only if $x = \varphi([f])$ for some $f : J \to A$ (Proposition 4.19). We thus have to prove that there exists a function $f : J \to \widehat{S}$

such that for any $A \in \mathscr{A}$ the relation $f(j) \in A$ holds for almost all j. But since $\mathscr{U}$ is $\widehat{S}$-adequate, we find a function $f : J \to \widehat{S}$ such that for each $A \in \mathscr{A}$ there is some $F \in \mathscr{U}$ with $f(F) \subseteq A$. This is the required function, since $f(F) \subseteq A$ means $f(j) \in A$ for all $j \in F$, i.e. (since $F \in \mathscr{U}$) for almost all j. □

9.2 Compact Enlargements

By the results of Section 9.1, we are now able to define for any set S an enlargement $* : \widehat{S} \to \widehat{{}^*S}$. The idea to define a compact enlargement is to iterate this process.

Given a superstructure $\widehat{S}$, we put S_0 and start the induction with an enlargement of $\widehat{S}_0$, say $*_0 : \widehat{S}_0 \to \widehat{S}_1$. Next, we choose an enlargement of $\widehat{S}_1$, i.e. $*_1 : \widehat{S}_1 \to \widehat{S}_2$, and so on, i.e. $*_n : \widehat{S}_n \to \widehat{S}_{n+1}$ $(n = 0, 1, 2, \ldots)$ is an enlargement. We define for $n < m$ the composition $*_n^m := *_{m-1} \circ \cdots \circ *_n$, i.e. $*_n^m$ is the map which sends $\widehat{S}_n$ into $\widehat{S}_m$.

Now the reader should note that each $\widehat{S}_n$ has a natural language $\mathscr{L}_n$ where the set of constants $\operatorname{cns}(\mathscr{L}_n)$ is in a one-to-one correspondence to $\widehat{S}_n$ (the correspondence being the interpretation map I_n). For $n > 0$, some of these constants represent standard elements, i.e. elements which correspond to the image of some element of $\widehat{S}_{n-1}$ under the map $*_{n-1}$. Thus, each $*_n$ induces an injection $i_n : \operatorname{cns}(\mathscr{L}_n) \to \operatorname{cns}(\mathscr{L}_{n+1})$, and $*_n^m$ induces the injection $i_n^m := i_{m-1} \circ \cdots \circ i_n$ from $\operatorname{cns}(\mathscr{L}_n)$ into $\operatorname{cns}(\mathscr{L}_m)$. We also let $*_n^n$ and i_n^n be the identity maps of $\widehat{S}_n$ resp. $\operatorname{cns}(\mathscr{L}_n)$, and then have the crucial identities

$$*_n^m = *_k^m *_n^k \quad \text{and} \quad i_n^m = i_k^m i_n^k \qquad (n \le k \le m). \tag{9.1}$$

A sentence α in the language $\mathscr{L}_n$ is transformed by $*_n^m$ into a sentence ${}^{i_n^m}\alpha$ in the language $\mathscr{L}_m$ which arises from α by replacing any occurrence of a constant c in α by the constant $i_n^m(c)$.

The crucial point now is that a transitively bounded sentence α in the language $\mathscr{L}_n$ is true (interpreted by I_n in $\widehat{S}_n$) if and only if the sentence ${}^{i_n^m}\alpha$ in the language $\mathscr{L}_m$ is true (interpreted by I_m in $\widehat{S}_m$): This is a reformulation of the statement that each $*_n$ is elementary (after a trivial induction). Hence, we have:

Lemma 9.5. *Each of the maps $*_n^m : \widehat{S}_n \to \widehat{S}_m$ is elementary.*

Proof. It only remains to verify the second condition of Definition 3.1, i.e. that ${}^{*_n^m}S_n = S_m$. But since ${}^{*_n}S_n = S_{n+1}$ (because $*_n$ is elementary), this follows by a trivial induction on m. □

To simplify notation, we assume that all of the sets $\widehat{S}_n$ are pairwise disjoint (which can be arranged by choosing the atom sets S_n pairwise disjoint). This means that for each $x \in \bigcup_n \widehat{S}_n$ there is a unique n with $x \in \widehat{S}_n$. We denote this n by n_x.

To construct an abstract limit model, we introduce an equivalence relation on the set $\bigcup_n \widehat{S}_n$: We call two elements $x \in \widehat{S}_n$, $y \in \widehat{S}_m$ equivalent if $y = *^m_n(x)$ resp. $x = *^n_m(y)$ (depending on whether $n = n_x$ or $m = m_y$ is larger).

This is indeed an equivalence relation: By the convention $*^n_n(x) = x$, we have $x \sim x$. Moreover, $x \sim y$ means $y \sim x$, since the definition is symmetric. To see the transitivity (i.e. that $x \sim y$ and $y \sim z$ implies $x \sim z$), one has to distinguish six cases. In all cases, the transitivity follows from (9.1).

The abstract limit model $\mathscr{S}_\omega$ consists of all equivalence classes on the set $\bigcup_n \widehat{S}_n$. The relations $\in_\omega$ and $=_\omega$ are defined by:

$$[x] =_\omega [y] \iff *^k_n(x) = *^k_m(y) \text{ for all } k \geq n, m \text{ where } n = n_x,\ m = m_y. \quad (9.2)$$

$$[x] \in_\omega [y] \iff *^k_n(x) \in *^k_m(y) \text{ for all } k \geq n, m \text{ where } n = n_x,\ m = m_y. \quad (9.3)$$

Actually, $=_\omega$ is the usual equality of equivalence classes:

Lemma 9.6. *The relations $=_\omega$ and $\in_\omega$ are well-defined on $\mathscr{S}_\omega$, and moreover:*

$$[x] =_\omega [y] \iff *^k_n(x) = *^k_m(y) \text{ for some } k \geq n, m \text{ where } n = n_x,\ m = m_y. \quad (9.4)$$

$$[x] \in_\omega [y] \iff *^k_n(x) \in *^k_m(y) \text{ for some } k \geq n, m \text{ where } n = n_x,\ m = m_y. \quad (9.5)$$

Proof. The fact that the right-hand side of (9.2) resp. of (9.3) is equivalent to the right-hand of (9.4) resp. of (9.5) follows by observing that for any $K \geq k \geq n, m$, we have $*^K_n = *^K_k *^k_n$ and $*^K_m = *^K_k *^k_m$ and that $*^K_k$ preserves the equality and element relation: The latter follows from Lemma 3.5, since $*^K_k$ is elementary by Lemma 9.5. Once we know this equivalence, it follows straightforwardly that $=_\omega$ and $\in_\omega$ are well-defined. □

A sentence α in the language $\mathscr{L}_n$ will now be interpreted in the abstract model $\mathscr{S}_\omega$ by mapping any constant $c \in \text{cns}(\mathscr{L}_n)$ corresponding to some $x \in \widehat{S}_n$ to the equivalence class $[x]$.

Theorem 9.7. *A transitively bounded sentence α in the language $\mathscr{L}_n$ is true in the abstract model $\mathscr{S}_\omega$ if and only if it is true under the interpretation map I_n.*

Proof. It suffices to prove that whenever α is true under the interpretation map I_n, it is also true in the abstract model $\mathscr{S}_\omega$: Indeed, if we have proved this, and α is false under the interpretation map I_n, then the transitively bounded sentence $\beta = \neg\alpha$ is true. But then, by assumption, β is true in $\mathscr{S}_\omega$, and so α is false in $\mathscr{S}_\omega$.

Thus, without loss of generality, assume that α is true under the interpretation map I_n. We may equivalently rewrite α in *prenex normal form*, i.e. in the form

$$Q_1\underline{x}_1 : Q_2\underline{x}_2 : \ldots Q_k\underline{x}_k : \beta$$

where Q_j stands for a quantifier $\forall$ or $\exists$, and where β contains no further quantifiers.

For the proof, we make use of so-called *Herbrand-Skolem functors*. These are defined as follows: Let $j_1 < j_2 < \cdots < j_p$ be those indices for which Q_{j_i} is the symbol $\exists$ ($p = 0$ is not excluded). In particular, α has the form

$$\forall \underline{x}_1 : \ldots \underline{x}_{j_1-1} : \exists \underline{x}_{j_1} : Q_{j_1+1}\underline{x}_{j_1+1} : \ldots Q_k\underline{x}_k : \beta$$

($j_1 = 1$ is not excluded). Note that I_n is onto, i.e. each possible value of $\underline{x}_j$ is actually represented by some constant in the language $\mathrm{cns}(\mathscr{L}_n)$. Then the statement that α is true under the interpretation I_n means that for each possible value of $\underline{x}_1, \ldots, \underline{x}_{j_1-1}$, we find a constant $c_1^n(\underline{x}_1, \ldots, \underline{x}_{j_1} - 1)$ such that the sentence

$$Q_{j_1+1}\underline{x}_{j_1+1} : \ldots Q_k\underline{x}_k : \beta(c_1^n)$$

is true, where $\beta(c_1^n)$ arises from β by replacing all free occurrences of $\underline{x}_{j_1}$ by the constant $c_1^n(\underline{x}_1, \ldots, \underline{x}_{j_1-1})$. By the axiom of choice, we may assume that c_1^n is a function. Note that conversely, the existence of such a function c_1^n implies that α is true under the interpretation map I_n. (However, the reader should be aware that c_1^n is not a function in the sense of the language $\mathscr{L}_n$). The function c_1^n is called the *Herbrand-Skolem functor* for the existential quantifier Q_{j_1}. By an induction, we thus can eliminate all quantifiers, and find that α is true if and only if

$$\beta(c_1^n, \ldots, c_p^n)$$

holds for any values of the free variables $\underline{x}_j$ ($j \neq j_1, \ldots, j_p$) where c_k^n is a Herbrand-Skolem functor depending only on the choice of the values of $\underline{x}_j$ ($j \neq j_1, \ldots, j_p$). Note now that α is transitively bounded, and so it remains true under the interpretation I_m if $m \geq n$, provided we replace all occurrences of constants c by $i_n^m(c)$. Hence, we find for each $m \geq n$ Herbrand-Skolem functors $c_1^m, \ldots, c_p^m$ such that

$$^{i_n^m}\beta(c_1^m, \ldots, c_p^m) \tag{9.6}$$

holds for all values of the free variables $\underline{x}_j$ ($j \neq j_1, \ldots, j_p$) in $\widehat{S}_m$, where $^{i_n^m}\beta$ arises from β by replacing all occurrences of constants c by $i_n^m(c)$. We claim that there are even Herbrand-Skolem functors $c_1^\omega, \ldots, c_p^\omega$ such that

$$^{\omega}\beta(c_1^\omega, \ldots, c_p^\omega) \tag{9.7}$$

is true for all values of the free variables $\underline{x}_j$ ($j \neq j_1, \ldots, j_p$) in $\mathscr{S}_\omega$. Here, $^\omega\beta$ arises from β by interpreting all constants c which represent some $x \in \widehat{S}_n$ by the corresponding equivalence class $[x] \in \mathscr{S}_\omega$. The existence of such Herbrand-Skolem functors then means in turn that α is true when interpreted in the model $\mathscr{S}$.

Thus, let $\underline{x}_j$ ($j \neq j_1, \ldots, j_p$) have fixed values in $\mathscr{S}$, i.e. $\underline{x}_j = [a_j]$ where $[a_j] \in \mathscr{S}$. Note that we have $a_j \in \widehat{S}_{n_j}$ for some n_j. Fix $m \geq n$ such that $m \geq n_j$

for all j, and put $b_j := {}^{*}{}^{m}_{n_j} a_j$. Then $b_j \in \widehat{S}_m$, and $[a_j] = [b_j]$. In particular, $\underline{x}_j = [b_j]$. Now let c_k^ω be the equivalence class containing the interpretation of c_k^m under the values $\underline{x}_j = b_j$. For the choice $\underline{x}_j = b_j$ the formula (9.6) is true in $\widehat{S}_m$, and it follows that also (9.7) is true for $\underline{x}_j = [b_j]$ in $\mathscr{S}_\omega$, as claimed:

Indeed, since β contains no quantifiers, it consists only of logical connectives and elementary formulas $a = b$ and $a \in b$. If we can prove that the elementary formula in (9.6) is true if and only if the corresponding formula in (9.7) is true, also the complete formulas in (9.6) resp. (9.7) must have the same truth value.

It thus remains to prove that for any $a, b \in \widehat{S}_m$ we have $a = b$ if and only if $[a] =_\omega [b]$, and $a \in b$ if and only if $[a] \in_\omega [b]$. But it follows from (9.2) that $[a] =_\omega [b]$ implies $*_m^m(a) = *_m^m(b)$ and so $a = b$ (since we assume $a, b \in \widehat{S}_m$); analogously, $[a] \in_\omega [b]$ implies by (9.3) that $*_m^m(a) \in *_m^m(b)$, and so $a \in b$. For the converse implication, observe that $a = b$ implies $*_m^m(a) = *_m^m(b)$, and so $[a] =_\omega [b]$ by (9.4); analogously, $a \in b$ implies $*_m^m(a) \in *_m^m(b)$, and so $[a] \in_\omega [b]$ by (9.4). $\square$

It follows from the proof, that Theorem 9.7 holds not only for transitively bounded sentences but even for a larger class of sentences, if the corresponding embeddings $*_n$ preserve the truth for the corresponding class of sentences.

Exercise 48. Prove that any sentence α can equivalently be rewritten in prenex normal form.

As in Section 4.3, we now want to replace the abstract model $\mathscr{S}_\omega$ by some superstructure $\widehat{S}_\omega$. Of course, this replacement should be done in such a way that transitively bounded sentences α in any of the languages $\mathscr{L}_n$ should keep their truth value.

If we want to proceed analogously to Section 4.3, we have to define sets $\mathscr{I}_k \subseteq \mathscr{S}_\omega$ which represent "internal objects of level at most k", in particular $\mathscr{I}_0$ will represent the atoms.

The definition of $\mathscr{I}_k$ is rather straightforward: Recall that each superstructure $\widehat{S}_n$ is built of level sets $S_{n,k}$ (i.e. $S_{n,0}$ represents the atoms of the superstructure, and $S_{n,k+1} = S_{n,0} \cup \mathcal{P}(S_{n,k})$). Then we just let

$$\mathscr{I}_k := \{[x] : x \in S_{n,k} \text{ for some } n\}.$$

It follows from the definition of the superstructures $\widehat{S}_n$ that $\mathscr{I}_0 \subseteq \mathscr{I}_1 \subseteq \cdots$ and $\bigcup \mathscr{I}_k = \mathscr{S}_\omega$.

As already announced, we put now $S_\omega := \mathscr{I}_0$, i.e. S_ω consists of all classes $[x]$ where x is an atom in some of the superstructures $\widehat{S}_n$.

To continue as in Section 4.3, we need an analogue to Lemma 4.16 and Lemma 4.17:

Lemma 9.8. *We have*

$$\mathscr{I}_k = \{[x] \in \mathscr{S}_\omega : \textit{There is some } [y] \in \mathscr{I}_{k+1} \textit{ with } [x] \in_\omega [y]\}.$$

Proof. If $[x] \in \mathscr{I}_k$ is given, we have $x \in S_{n,k}$ for some n. Then $y := S_{n,k}$ is an element of $S_{n,k+1}$, and so $[y] \in \mathscr{I}_{k+1}$. By (9.5), we have $[x] \in_\omega [y]$.

Conversely, assume that $[y] \in \mathscr{I}_{k+1}$ and $[x] \in_\omega [y]$. Then $x \in S_{n,i}$, $y \in S_{m,k+1}$ for some n, m and some i, and we find some $K \geq n, m$ with $*_n^K(x) \in *_m^K(y)$. Since $*_m^K$ is elementary (Lemma 9.5), we have ${}^{*_m^K}S_{m,k+1} \subseteq S_{K,k+1}$ (Theorem A.1), and so $*_m^K(y) \in S_{K,k+1} = S_{K,0} \cup \mathcal{P}(S_{K,k})$. In view of $*_n^K(x) \in *_m^K(y)$, this implies $*_n^K(x) \in S_{K,k}$. Hence, $[x] = [*_n^K(x)] \in \mathscr{I}_k$. □

Lemma 9.9. *If elements* $[x], [y] \in \mathscr{S}_\omega \setminus \mathscr{I}_0$ *satisfy*

$$\{[z] \in \mathscr{S}_\omega : [z] \in_\omega [x]\} = \{[z] \in \mathscr{S}_\omega : [z] \in_\omega [y]\}, \tag{9.8}$$

then $[x] = [y]$.

Proof. Let $x \in \widehat{S}_m$ and $y \in \widehat{S}_n$, without loss of generality, $n \geq m$: Putting $x_0 = *_m^n(x)$, we have $[x_0] = [x]$ and $x_0 \in \widehat{S}_n$. Hence, replacing x by x_0 if necessary, it is no loss of generality to assume that x and y belong to the *same* superstructure $\widehat{S}_n$. Then we have $[x] = [y]$ if and only if $x = y$. Moreover, x and y are not atoms of the superstructure $\widehat{S}_n$, because $[x], [y] \notin \mathscr{I}_0$. Consequently, if $[x] \neq [y]$, then the sets x and y are different, and so we find an element which belongs to one of this set but not to the other. Without loss of generality, $z \in x \setminus y$. In view of Theorem 9.7, we then have $[z] \in_\omega [x]$ and $\neg[z] \in_\omega [y]$, a contradiction to (9.8). □

Now we may proceed analogously to Section 4.3: As in Section 4.3, we may define an injection $\varphi_\omega : \mathscr{S}_\omega \to \widehat{S}_\omega$:

$$\varphi_\omega([y]) = \begin{cases} [y] & \text{if } [y] \in \mathscr{I}_0 = S_\omega, \\ \{\varphi_\omega([x]) : [x] \in_\omega [y]\} & \text{if } [y] \notin \mathscr{I}_0 = S_\omega. \end{cases} \tag{9.9}$$

Indeed, for the construction of φ_ω in Section 4.3, no particular properties of the abstract model $\mathscr{S}$ have been used except for Lemma 4.16 and Lemma 4.17 for which we just have proved corresponding replacements.

Also an analogue of Theorem 4.18 can be proved (with essentially the same proof): In this connection the reader should note that the set $\mathscr{I}'$ from Section 4.3 corresponds in our case to the whole abstract model $\mathscr{S}_\omega$. We only formulate the analogue of the second part of Theorem 4.18:

Theorem 9.10. *A transitively bounded sentence in the language whose constants are taken from* $\mathscr{S}_\omega$ *is true (interpreted by the identity map) if and only if it is true under the interpretation map* φ_ω. □

We emphasize that the proof of Theorem 9.10 makes use of the fact that the sentence is transitively bounded, but this restriction could be relaxed in the same way in which this restriction in Theorem 4.18 could have been relaxed.

Now we define a new language $\mathscr{L}_\omega$ by taking the set of its constants, $\operatorname{cns}(\mathscr{L}_\omega)$ in a one-to-one correspondence with $\widehat{S}_\omega$ (the correspondence being the interpretation map I_ω). Now we can define maps $*_n^\omega : \widehat{S}_n \to \widehat{S}_\omega$ by

$$*_n^\omega(x) = \varphi_\omega([x]),$$

and also $i_n^\omega : \operatorname{cns}(\mathscr{L}_n) \to \operatorname{cns}(\mathscr{L}_\omega)$ by $i_n^\omega = I_\infty^{-1} \circ *_n^\omega \circ I_n$, i.e. if $c \in \operatorname{cns}(\mathscr{L}_n)$ corresponds to the element $x \in \widehat{S}_n$, then $i_n^\omega(c)$ corresponds to the element $*_n^\omega(x)$.

For later reference, we observe by the way that the value $*_n^\omega(x)$ depends by definition only on the equivalence class of x, and so it follows that

$$*_n^\omega = *_k^\omega *_n^k \qquad (n \le k).$$

Hence, the relation (9.1) holds also for $m = \omega$ (and even for $k = \omega$ or $n = \omega$ if we define $*_\omega^\omega(x) = x$).

Theorem 9.11. *Each of the maps* $*_n^\omega : \widehat{S}_n \to \widehat{S}_\omega$ *is elementary.*

Moreover, if an element $x \in \widehat{S}_\omega$ *is internal under this map, then there is some* $m < \omega$ *such that* x *is a standard entity under the map* $*_m^\omega : \widehat{S}_m \to \widehat{S}_\omega$.

Proof. Let α be a transitively bounded sentence in the language $\mathscr{L}_n$ which is true under the interpretation map I_n. By Theorem 9.7, α is true under the interpretation map $J \circ I_n$ where $J : \widehat{S}_n \to \mathscr{S}_\omega$ is the map sending each element $x \in \widehat{S}_n$ onto the equivalence class $[x]$. Hence, Theorem 9.7 implies that α is true under the interpretation map $I_n' = \varphi_\omega \circ J \circ I_n = *_n^\omega \circ I_n$.

Hence, condition 1. of Definition 3.1 is satisfied. The other condition reads in our case ${}^{*_n^\omega} S_n = S_\omega$. To prove this, note that (9.9) implies

$${}^{*_n^\omega} S_n = \varphi_\omega([S_n]) = \{\varphi_\omega([x]) : [x] \in_\omega [S_n]\}. \tag{9.10}$$

By Lemma 9.8, the relation $[x] \in_\omega [S_n]$ implies $[x] \in \mathscr{I}_0$. But also the converse holds: If $[x] \in \mathscr{I}_0$, then $[x] \in_\omega [S_n]$.

To see the latter, let $[x] \in \mathscr{I}_0$ and note that we may by definition of $\mathscr{I}_0$ choose some representative x which is an atom $x \in S_m$ for some m. If $m \le n$, then ${}^{*_m^n} S_m = S_n$, because $*_m^n$ is elementary (Lemma 9.5), and so $x_0 = *_m^n(x) \in S_n$. Since $x_0 \sim x$, we have $[x] = [x_0] \in_\omega [S_n]$. If $m \ge n$, then the relation ${}^{*_n^m} S_n = S_m$ (because $*_n^m$ is elementary) implies that $[x] \in_\omega [{}^{*_n^m} S_n] = [S_n]$. Hence, in both cases $[x] \in_\omega [S_n]$, as claimed.

It thus follows from (9.10) that

$${}^{*_n^\omega} S_n = \{\varphi_\omega([x]) : [x] \in \mathscr{I}_0\} = \mathscr{I}_0 = S_\omega,$$

where we made use of (9.9). This completes the proof that $*_n^\omega$ is elementary.

If x is internal under the map $*_n^\omega$, then $x \in {}^{*_n^\omega}S_{n,k}$ for some k. By (9.9), we have

$$ {}^{*_n^\omega}S_{n,k} = \varphi_\omega([S_{n,k}]) = \{\varphi_\omega([y]) : [y] \in_\omega [S_n]\}. $$

In particular, $x = \varphi_\omega([y])$ for some $[y] \in \mathscr{S}_\omega$. We must have $y \in \widehat{S}_m$ for some m, and so $x = *_m^\omega(y)$. □

We thus have constructed a true limit model $\widehat{S}_\omega$ for the sequence $\widehat{S}_n$. Since each of the models $\widehat{S}_n$ becomes successively "more saturated", and since any transitively bounded sentence in the model $\widehat{S}_n$ is also true in the limit model $\widehat{S}_\omega$, one might expect that the limit model is "enormously saturated". This is indeed the case:

Theorem 9.12. *The map $*_0^\omega : \widehat{S} \to \widehat{S}_\omega$ is a compact enlargement.*

Proof. By Theorem 9.11, the map $* = *_0^\omega$ is elementary. Let φ be an internal binary relation which is concurrent on an entity A which contains at most finitely many nonstandard elements.

By Theorem 9.11, we find some index m_φ such that φ is standard under the map $*_{m_\varphi}^\omega : \widehat{S}_{m_\varphi} \to \widehat{S}_\omega$. This means that there is some $\varphi_0 \in \widehat{S}_{m_\varphi}$ such that $\varphi = {}^{*_{m_0}^\omega}\varphi_0$.

Similarly, we find for each $a \in A$ some m_a such that $a = {}^{*_{m_a}^\omega}c_a$ for some $c_a \in \widehat{S}_{m_a}$; moreover, we may assume that $m_a = 0$ for all except at most finitely many $a \in A$ (because A contains only finitely many nonstandard elements).

Hence, $m := \max(\{m_a : a \in A\} \cup \{m_0\})$ exists. Consider in the model $\widehat{S}_m$ the relation $\psi := {}^{*_{m_0}^m}\varphi_0$ and the set B consisting of the elements $b_a := {}^{*_{m_a}^m}c_a$.

We claim that ψ is concurrent on B: Indeed, if $a_1, \ldots, a_n \in A$, then there is some $y \in \operatorname{rng}(\varphi)$ such that $(a_k, y) \in \varphi$ for $k = 1, \ldots, n$, because φ is concurrent on A. Now note that $\operatorname{rng}(\varphi) = \operatorname{rng}({}^{*_m^\omega}\psi) = {}^{*_m^\omega}\operatorname{rng}(\psi)$ and $a_k = {}^{*_m^\omega}b_{a_k}$. We thus have in $\widehat{S}_\omega$,

$$ \exists \underline{y} \in {}^{*_m^\omega}\operatorname{rng}(\psi) : ({}^{*_m^\omega}b_{a_1}, \underline{y}), \ldots, ({}^{*_m^\omega}b_{a_n}, \underline{y}) \in {}^{*_m^\omega}\psi $$

Applying the inverse form of the transfer principle (for the map $*_m^\omega$), we find that there is some $y \in \operatorname{rng}(\psi)$ such that $(b_{a_1}, y), \ldots, (b_{a_n}, y) \in \psi$, i.e. ψ is concurrent on B, as claimed.

Since $*_m^{m+1} : \widehat{S}_m \to \widehat{S}_{m+1}$ is an enlargement, we may conclude that there is some $y \in \widehat{S}_{m+1}$ such that $({}^{*_m^{m+1}}b_a, y) \in {}^{*_m^{m+1}}\psi$ holds for each $b_a \in B$. Hence, $({}^{*_{m+1}^\omega}({}^{*_m^{m+1}}b_a), {}^{*_{m+1}^\omega}y) \in {}^{*_{m+1}^\omega}({}^{*_m^{m+1}}\psi)$ for each $a \in A$ which means $(a, {}^{*_{m+1}^\omega}y) \in \varphi$ for each $a \in A$, i.e. φ is satisfied on A, as desired. □

9.3 Polysaturated Models

To prove the existence of a polysaturated model, we have to work even more: We have to repeat the construction of Section 9.2 a transfinite number of times.

More precisely, we construct a transfinite sequence of enlargements: By a transfinite induction, we associate to each ordinal number α a model $\widehat{S}_\alpha$ and a family of maps $*^\alpha_\beta : \widehat{S}_\beta \to \widehat{S}_\alpha$ ($\beta \le \alpha$) such that the following holds:

1. For all ordinals $\alpha_1 \le \alpha_2 \le \alpha_3 \le \alpha$ we have
$$*^{\alpha_3}_{\alpha_1} = *^{\alpha_3}_{\alpha_2} *^{\alpha_2}_{\alpha_1},$$
and $*^\alpha_\alpha$ is the identity.
2. For each $\beta < \alpha$ the map $*^\alpha_\beta : \widehat{S}_\beta \to \widehat{S}_\alpha$ is elementary.
3. If α is a successor ordinal, i.e. $\alpha = \beta + 1$ for some ordinal β, then $*^\alpha_\beta : \widehat{S}_\beta \to \widehat{S}_\alpha$ is even an enlargement.
4. If α is a limit ordinal (i.e. not a successor ordinal), then each element which is internal under some map $*^\alpha_\beta : \widehat{S}_\beta \to \widehat{S}_\alpha$ with $\beta < \alpha$ is standard under another such map.

Theorem 9.13. *For each set S there exists a transfinite sequence as described above such that $\widehat{S}_0 = \widehat{S}$.*

Proof. The induction start is clear: Put $S_0 := S$, and let $*^0_0 : \widehat{S}_0 \to \widehat{S}_0$ be the identity. If $\alpha = \beta + 1$, let $*^\alpha_\beta : \widehat{S}_\beta \to \widehat{S}_\alpha$ be an enlargement, and for ordinals $\gamma < \beta$ define $*^\alpha_\gamma := *^\alpha_\beta *^\beta_\gamma$.

Thus, the only case which needs some care is if α is a limit ordinal. For $\alpha = \omega$, we have proved the existence of a limit model. However, for the general case, the proof is analogous: Actually, we have at no place used the fact that we determined the limit of a countable sequence (only the relation (9.1) has been used). For clearer representation we had used in the above proof languages $\mathscr{L}_\beta$, interpretation maps $I_\beta : \mathrm{cns}(\mathscr{L}_\beta) \to \widehat{S}_\beta$, and maps $i^{\beta_2}_{\beta_1} : \mathrm{cns}(\mathscr{L}_{\beta_1}) \to \mathrm{cns}(\mathscr{L}_{\beta_2})$: One may of course just put $\mathrm{cns}(\mathscr{L}_\beta) := \widehat{S}_\beta$, and let I_β and $i^{\beta_2}_{\beta_1}$ be the identity map. Then, with the same proof as above, we can construct a limit model $\widehat{S}_\alpha$ such that each of the embeddings $*^\alpha_\beta : \widehat{S}_\beta \to \widehat{S}_\alpha$ is elementary. □

The crucial point of the above construction is that we now indeed obtain κ-saturated models. The proof is in most parts analogous to Theorem 9.12, but one has to take care, since there need not exist a maximum for infinite sets of ordinals. To ensure that the supremum is actually strictly smaller than the ordinal number α of the highest model, we implicitly make use of a property of successor cardinals α which is called "regularity" in literature on set theory:

Theorem 9.14. *Let S and κ be arbitrary sets. Then in the transfinite sequence from Theorem 9.13 the elementary map $*_0^\alpha : \widehat{S} \to \widehat{S}_\alpha$ is κ-saturated, if α is the first ordinal with a strictly larger cardinality than κ.*

Proof. We may assume that κ is infinite, since otherwise each elementary embedding is κ-saturated.

Let $\mathscr{B}$ be a nonempty system of internal entities which has the finite intersection property and at most the cardinality of κ. We have to prove that $\bigcap \mathscr{B} \neq \varnothing$. Since each $B \in \mathscr{B}$ is an internal set, we find some index $\beta_B < \alpha$ such that B is a standard set under the embedding $*_{\beta_B}^\alpha : \widehat{S}_{\beta_B} \to \widehat{S}_\alpha$, i.e. there is some $C_B \in \widehat{S}_{\beta_B}$ such that $B = {}^{*_{\beta_B}^\alpha} C_B$.

Put $\beta := \bigcup \{\beta_B : B \in \mathscr{B}\}$. Then β is an ordinal number. Since $\beta_B < \alpha$, the set β_B has at most the cardinality of κ. Hence, β has at most the cardinality of the set $\kappa \times \mathscr{B}$. By assumption, $\mathscr{B}$ has at most the cardinality of κ. Hence, β has at most the cardinality of $\kappa \times \kappa$ which is the cardinality of κ, since κ is infinite. Summarizing, β has at most the cardinality of κ. By our choice of α, we may conclude that $\beta < \alpha$.

Consider in the model $\widehat{S}_\beta$ the sets $A_B := {}^{*_{\beta_B}^\beta} C_B$. Then the system $\mathscr{A} := \{A_B : B \in \mathscr{B}\}$ has the finite intersection property: Indeed, if $B_1, \ldots, B_n \in \mathscr{B}$, then $B_1 \cap \cdots \cap B_n \neq \varnothing$, since $\mathscr{B}$ has the finite intersection property by assumption. Now note that ${}^{*_\beta^\alpha} A_{B_i} = {}^{*_{\beta_B}^\alpha} C_{B_i} = B_i$, and so

$$
{}^{*_\beta^\alpha}(A_{B_1} \cap \cdots \cap A_{B_n}) = {}^{*_\beta^\alpha} A_{B_1} \cap \cdots \cap {}^{*_\beta^\alpha} A_{B_n} = B_1 \cap \cdots \cap B_n \neq \varnothing,
$$

because $*_\beta^\alpha$ is a superstructure monomorphism. Hence, $A_{B_1} \cap \cdots \cap A_{B_n} \neq \varnothing$, and $\mathscr{A}$ has the finite intersection property, as claimed.

Since $*_\beta^{\beta+1} : \widehat{S}_\beta \to \widehat{S}_{\beta+1}$ is an enlargement, we may conclude that there is some $d \in \widehat{S}_{\beta+1}$ which is contained in each of the sets $D_B := {}^{*_\beta^{\beta+1}} A_B$ ($B \in \mathscr{B}$). Since $D_B = {}^{*_{\beta_B}^{\beta+1}} C_B$, we find by Lemma 3.5 that $b := {}^{*_{\beta+1}^\alpha} d \in {}^{*_{\beta+1}^\alpha} D_B = {}^{*_{\beta_B}^\alpha} C_B = B$ ($B \in \mathscr{B}$). In particular, $b \in \bigcap \mathscr{B}$, and so $*_0^\alpha$ is κ-saturated. □

Corollary 9.15. *For each set S there exists a polysaturated map $* : \widehat{S} \to \widehat{{}^*S}$.* □

Our above construction of the transfinite sequence is similar to the construction of ultralimits in [SL76, Section 7.5] (see also [Lux69a, Section 1.7]): In these books, a transfinite sequence of *abstract* (ultrapower) models is considered. It is not discussed there, how the limit model might be embedded into a superstructure (and it appears that this is a rather difficult task).

However, our construction is different: We consider at each step of the transfinite sequence already the embedding into some superstructure. This has the advantage that one is not forced to use ultrapower models to construct the enlargements needed for the induction step $\alpha \mapsto \alpha + 1$. (As mentioned after Theorem 8.5, there exist enlargements which are not $\mathbb{N}$-saturated. By Theorem 9.1, these enlargements cannot be described by ultrapower models). However, even if one uses ultrapowers, it is not clear whether we end up with the same model as in [SL76], because the embedding of the abstract model into the superstructure in each step does not preserve all sentences but only the transitively bounded sentences.

Chapter 5

Functionals, Generalized Limits, and Additive Measures

Throughout this chapter, we assume that $\mathbb{R} \in \widehat{S}$ is an entity.

§10 Normed Spaces

10.1 Linear Functionals and Operators

Recall that a map $\|\cdot\| : X \to \mathbb{R}_+$ on a linear space (=vector space) X over $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$ is called a *norm*, if the following holds:

1. $\|x\| = 0$ if and only if $x = 0$.
2. $\|\lambda x\| = |\lambda| \, \|x\|$ for scalars $\lambda \in \mathbb{K}$.
3. $\|x + y\| \leq \|x\| + \|y\|$ (the *triangle inequality*).

Each norm induces a metric by means of the formula $d(x, y) = \|x - y\|$. A normed linear space is called *Banach space*, if it is *complete* with respect to the metric induced by the norm (i.e. if any Cauchy sequence converges). A map $f : X \to Y$ in normed spaces is called *linear* if $f(x + y) = f(x) + f(y)$ and $f(\lambda x) = \lambda f(x)$. If $Y = \mathbb{R}$, then f is called a *(linear) functional*. If X and Y are normed spaces, then a linear map $f : X \to Y$ is called *bounded*, if there is a constant $L \in \mathbb{R}_+$ with

$$\|f(x)\| \leq L \, \|x\| \qquad (x \in X),$$

or, equivalently

$$\sup_{\|x\|=1} \|f(x)\| = \sup_{\|x\| \leq 1} \|f(x)\| = \sup_{x \neq 0} \frac{\|f(x)\|}{\|x\|} < \infty.$$

The minimum of all such constants L (which is the above supremum) is called the *norm*. In particular, a linear functional is *bounded* if and only if

$$\|f\| = \sup_{\|x\|=1} |f(x)| < \infty.$$

The space of all bounded linear functionals on X is usually denoted by X^*. It is endowed with the canonical vector operations (i.e. $(f+g)(x) := f(x) + g(x)$ and $(\lambda f)(x) := \lambda f(x)$). It is well-known and can be easily checked that X^* is always a Banach space with respect to the norm introduced above; it is called the *dual space* to X.

A classical result of functional analysis is the theorem of Hahn-Banach for which we give now a nonstandard proof:

Theorem 10.1 (Hahn-Banach)**.** *Let X_0 be a linear subspace of a normed linear space X. Let $f \in X_0^*$. Then f may be extended to a bounded linear functional $F \in X^*$ without increasing the norm, i.e. $f(x) = F(x)$ for $x \in X_0$ and $\|f\| = \|F\|$.*

The nonstandard proof of Theorem 10.1 is reduced to the following special case in the standard world. This part of the proof is classical:

Lemma 10.2. *Let f, X_0, X be as in Theorem 10.1 with $\mathbb{K} = \mathbb{R}$. Then for each finite number of elements $x_1, \ldots, x_n \in X$ one may extend f to a subspace $U \subseteq X$ which contains $x_1, \ldots, x_n$ without increasing the norm.*

Proof. Evidently, it suffices to consider $n = 1$ (because in the general case, we may first extend f to a subspace which contains x_1, then to a subspace which also contains x_2, etc.). If $x_1 \in X_0$, the statement is trivial. Otherwise, put $U := \operatorname{span}(X_0 \cup \{x_1\})$. Note that x_1 is linearly independent of X_0, i.e. each $u \in U$ may uniquely be written in the form $u = x + \lambda x_1$ where $x \in X_0$ and $\lambda \in \mathbb{R}$. The only way to extend f linearly to U is by putting

$$F(x + \lambda x_1) := f(x) + \lambda c$$

with some $c \in \mathbb{R}$. We have to prove that we may choose c in such a way that $\|F\| \leq \|f\|$, i.e. $|F(u)| \leq \|f\| \, \|u\|$ for all $u \in U$. The latter is equivalent to

$$|f(x_0) + \lambda c| \leq \|f\| \; \|x_0 + \lambda x_1\| \qquad (x_0 \in X_0, \ \lambda \in \mathbb{R}).$$

After dividing by $|\lambda|$ (the case $\lambda = 0$ is trivial), we find, putting $x := x_0/\lambda \in X$, that the above inequality is equivalent to

$$|f(x) + c| \leq \|f\| \; \|x + x_1\| \qquad (x \in X_0).$$

Since $\mathbb{K} = \mathbb{R}$, we thus have to find some $c \in \mathbb{R}$ such that

$$-\|f\| \; \|x + x_1\| - f(x) \leq c \leq \|f\| \; \|x + x_1\| - f(x).$$

Such a constant c exists, since for each $x, y \in X_0$ the estimate

$$\begin{aligned} f(x) - f(y) = f(x-y) &\le \|f\| \, \|x-y\| = \|f\| \, \|(x+x_1) - (y+x_1)\| \\ &\le \|f\| \, (\|x+x_1\| + \|y+x_1\|) \end{aligned}$$

holds, which implies that

$$\sup_{y \in X_0} (- \|f\| \, \|y + x_1\| - f(y)) \le \inf_{x \in X_0} (\|f\| \, \|x + x_1\| - f(x)).$$

Thus, we may just choose some c in between these two quantities. □

Proof of Theorem 10.1. First, assume that $\mathbb{K} = \mathbb{R}$. Choose some S such that $\mathbb{R}, X \in \widehat{S}$ are entities. Let $* : \widehat{S} \to \widehat{{}^*S}$ be an enlargement. Consider the relation

$$\varphi := \{(\underline{x}, \underline{y}) \in X \times \mathscr{F}(X, \mathbb{R}) \mid \alpha(\underline{x}, \underline{y})\},$$

where $\mathscr{F}(X, \mathbb{R})$ denotes the set of all functions $y : X \to \mathbb{R}$ with $\operatorname{dom}(y) \subseteq X$ and $\operatorname{rng}(y) \subseteq \mathbb{R}$, and $\alpha(\underline{x}, \underline{y})$ is a transitively bounded sentence with the meaning "$\underline{y}$ is linear, defined on a subspace $U \supseteq X_0$ with $\underline{x} \in U$, and its norm does not exceed $\|f\|$". By Lemma 10.2, the relation φ is concurrent on X. Since $*$ is an enlargement, Theorem 8.10 implies that there is some $y \in {}^*\mathscr{F}(X, \mathbb{R})$ which satisfies ${}^*\varphi$ on ${}^\sigma\operatorname{dom}(\varphi) = {}^\sigma X$, i.e. $({}^*x, y) \in {}^*\varphi$ for any $x \in X$. To determine ${}^*\varphi$, we may use the standard definition principle for relations. In view of Exercise 83, y is an internal linear functional, defined on an internal subspace $U \subseteq {}^*X$ with ${}^*x \in U$ for each $x \in X$ and such that $|y(u)| \le {}^*(\|f\|)^* \|u\|$ for $u \in U$. Define now $F : X \to \mathbb{R}$ by $F(x) := \operatorname{st}(y({}^*x))$. Since st is linear (Theorem 5.21), also F is linear. Moreover, since

$$|y({}^*x)| \le {}^*\|y\|^* \|{}^*x\| \le {}^*(\|f\|)^* \|{}^*x\| = {}^*(\|f\| \, \|x\|),$$

the function F is indeed defined on X (because $y({}^*x)$ is finite), and $|F(x)| \le \|f\| \, \|x\|$, i.e. $\|F\| \le \|f\|$.

The case $\mathbb{K} = \mathbb{C}$ is reduced to the case $\mathbb{K} = \mathbb{R}$: Indeed, by considering only multiplication with real scalars, we find that the function $f_{\mathbb{R}}(x) := \operatorname{Re}(f(x))$ defines a linear functional over the (real) vector space X which satisfies $\|f_{\mathbb{R}}\| \le \|f\|$. By what we proved above, we may extend $f_{\mathbb{R}}$ to some (real) linear $F_{\mathbb{R}}$ on X without increasing the norm. Define now $F(x) := F_{\mathbb{R}}(x) - iF_{\mathbb{R}}(ix)$. Then F is linear in the complex sense, since it is linear in the real sense and $F(ix) = F_{\mathbb{R}}(ix) - iF_{\mathbb{R}}(-x) = F_{\mathbb{R}}(ix) + iF_{\mathbb{R}}(x) = iF(x)$. Moreover,

$$\|F\| = \sup_{\|x\|=1} \sup_{|\lambda|=1} \operatorname{Re}(\lambda F(x)) = \sup_{\|x\|=1} \sup_{|\lambda|=1} \operatorname{Re}(F(\lambda x)) \le \sup_{\|x\|=1} |F_{\mathbb{R}}(x)| = \|F_{\mathbb{R}}\|,$$

and so $\|F\| \le \|F_{\mathbb{R}}\| = \|f_{\mathbb{R}}\| \le \|f\|$. □

Some remarks are appropriate: The previous nonstandard proof is essentially due to W. A. J. Luxemburg [Lux62]. The main advantage compared to the classical (standard) proof is that a careful analysis of the underlying model shows that not the full power of the axiom of choice is needed but only the existence of a certain ultrafilter which is slightly less restrictive from the logical point of view [Pin73]. Actually, by a refinement of the method, it is proved in [Lux69c] that one does not even need an ultrafilter but only an additive measure which is even less restrictive [PS77]. Anyway, in the author's opinion these weakenings of the axiom of choice are not too interesting, since already the theorem of Hahn-Banach implies the most counterintuitive consequences of the axiom of choice (for example, the existence of nonmeasurable sets [FW91] or the Banach-Tarski paradox [Paw91] can be proved by Hahn-Banach's extension theorem without the axiom of choice).

Exercise 49. Prove the following form of the Hahn-Banach extension theorem: Let X be a real linear space, and $p : X \to \mathbb{R}$ *sublinear*, i.e. $p(x+y) \le p(x)+p(y)$ and $p(\lambda x) = \lambda p(x)$ for $\lambda > 0$. If $X_0 \subseteq X$ is a subspace and $f : X_0 \to \mathbb{R}$ is linear with $f(x) \le p(x)$ on X_0, then f may be extended to a linear $F : X \to \mathbb{R}$ with $F(x) \le p(x)$ on X. Why is Theorem 10.1 for $\mathbb{K} = \mathbb{R}$ a special case?

Exercise 50. A normed space X is called *separable*, if there is a countable set $C \subseteq X$ with $\overline{C} = X$. Give a standard proof of Theorem 10.1 for the case that X is separable, using only Lemma 10.2 and a countable (recursive) form of the axiom of choice.

Hint: Extend f first to the linear hull U of $C \cup X_0$.

It is a well-known result of functional analysis that on any infinite-dimensional normed space there exist linear functionals which are not bounded. However, in the nonstandard world, it suffices to consider bounded (internal) functionals:

Theorem 10.3. *Let X be a normed linear space with dual space X^*. Let $*$ be an X-enlargement. Then any linear functional f on X (not necessarily bounded) may be written in the form*

$$^*(f(x)) = g(^*x) \qquad (x \in X)$$

with some $g \in {}^*(X^*)$. In particular,

$$f(x) = \operatorname{st}(g(^*x)) \qquad (x \in X).$$

Proof. Consider the binary relation

$$\varphi := \{(\underline{x}, \underline{y}) \in X \times X^* \mid \underline{y}(\underline{x}) = f(\underline{x})\}.$$

Then φ is concurrent on X: Indeed, if $x_1, \ldots, x_n \in X$ are given, let X_0 be the linear hull of $x_1, \ldots, x_n$. Since X_0 has finite dimension, it is isomorphic to some $\mathbb{K}^n$, and so the restriction of f to X_0 is bounded. By the Hahn-Banach extension theorem, this restriction has an extension to some $F \in X^*$. Then $(x_1, F), \ldots, (x_n, F) \in \varphi$, and so φ is concurrent. Since $*$ is an X-enlargement, φ is satisfied on ${}^\sigma X$, i.e. we find some $g \in {}^*(X^*)$ such that $({}^*x, g) \in {}^*\varphi$ for any $x \in X$. The standard definition for relations implies that for any $x \in X$ the equality $g({}^*x) = {}^*f({}^*x) = {}^*(f(x))$ holds. □

Recall that if X has finite dimension N, then any linear map $A : X \to Y$ may be written in matrix form

$$A(x) = \sum_{n=1}^{N} f_n(x) y_n \tag{10.1}$$

where f_n are linear functionals and $y_n \in Y$. In nonstandard analysis, all operators have matrix form:

Exercise 51. Let X be a normed linear space, Y be a linear space, and $A : X \to Y$ be linear. If $*$ is an X-enlargement, prove that there are internal *-finite sequences $y_1, \ldots, y_h \in {}^*Y$ and $f_1, \ldots, f_h \in {}^*(X^*)$ such that

$${}^*(A(x)) = \sum_{n=1}^{h} f_n({}^*x) y_n \qquad (x \in X).$$

10.2 Hahn-Banach and Banach-Mazur Limits

The space ℓ_∞ is the set of all bounded real sequences $x = (\xi_n)_n$, equipped with the norm

$$\|x\|_\infty = \sup_n |\xi_n| .$$

It is well-known (and easily checked) that ℓ_∞ is a Banach space. By ℓ_p $(1 \leq p < \infty)$, we denote the set of all real sequences $x = (\xi_n)_n$ with finite norm

$$\|x\|_p = \left(\sum_{n=1}^{\infty} |\xi_n|^p \right)^{1/p} .$$

Also ℓ_p is a Banach space. It is well-known that in case $1 \leq p < \infty$ the space $\ell_{p'}$ with $\frac{1}{p} + \frac{1}{p'} = 1$ is dual to ℓ_p in the sense that each element $y = (\eta_n)_n \in \ell_{p'}$ defines a bounded linear functional f_y on ℓ_p by means of the formula

$$f_y(x) = \sum_{n=1}^{\infty} \eta_n \xi_n, \tag{10.2}$$

and conversely all bounded linear functionals have this form.

For ℓ_∞ the situation is different: Any $y \in \ell_1$ defines a bounded linear functional on ℓ_∞, but (if we assume the axiom of choice) there exist also bounded linear functionals on ℓ_∞ which do not have this form, as we shall see soon. Although much is known about the structure of the space ℓ_∞^* of bounded linear functionals on ℓ_∞, there are still some open problems. Nonstandard analysis is the most convenient tool to study ℓ_∞^*. In fact, by nonstandard methods the space ℓ_∞^* is easily described, as was first observed in [Rob64]. To give the main result of [Rob64], we need the following result of the standard world:

Lemma 10.4. *Let $f \in \ell_\infty^*$. Then for any finite number of elements $x_1, \ldots, x_K \in \ell_\infty$, $x_k = (\xi_{k,n})_n$ and any $\varepsilon > 0$ there exist real numbers $\eta_1, \ldots, \eta_N \in \mathbb{R}$ such that*

$$f(x_k) = \sum_{n=1}^{N} \eta_n \xi_{k,n} \qquad (k = 1, \ldots, K) \tag{10.3}$$

and

$$\sum_{n=1}^{N} |\eta_n| \leq (1+\varepsilon) \|f\| .$$

Proof. If one of the vectors $x_1, \ldots, x_K$, say x_K, is a linear combination of the others, it suffices to find the numbers $\eta_1, \ldots, \eta_N$ corresponding to $x_1, \ldots, x_{K-1}$, since (10.3) then holds also for x_K by the linearity. Successively eliminating such vectors, we may assume without loss of generality that $x_1, \ldots, x_K$ are linearly independent. Following [Ban87, p. 42/43], we prove under this assumption that there is some N such that

$$\|\lambda_1 x_1 + \cdots + \lambda_K x_K\|_\infty \leq (1+\varepsilon) \max_{n=1,\ldots,N} |\lambda_1 \xi_{1,n} + \cdots + \lambda_K \xi_{K,n}| \tag{10.4}$$

for any choice $\lambda_1, \ldots, \lambda_K \in \mathbb{R}$. Indeed, if this is not true, we find for each N corresponding numbers $\lambda_{k,N}$ such that

$$\|\lambda_{1,N} x_1 + \cdots + \lambda_{K,N} x_K\|_\infty > (1+\varepsilon) \max_{n=1,\ldots,N} |\lambda_{1,N} \xi_{1,n} + \cdots + \lambda_{K,N} \xi_{K,n}| . \tag{10.5}$$

Dividing (10.5) by $\max\{|\lambda_{1,N}|, \ldots, |\lambda_{K,N}|\}$ if necessary, we may assume that $|\lambda_{k,N}| \leq 1$. Successively passing to subsequences, we find a subsequence N_j such that $\lambda_{k,N_j} \to \lambda_k$ converges as $j \to \infty$ for any $k = 1, \ldots, K$. Observe now that

$$\|(\lambda_{1,N_j} x_1 + \cdots + \lambda_{K,N_j} x_K) - (\lambda_1 x_1 + \cdots + \lambda_K x_K)\|_\infty \to 0.$$

Hence, passing to the limit in (10.5) we find, for each N, that

$$\|\lambda_1 x_1 + \cdots + \lambda_K x_K\|_\infty \geq (1+\varepsilon) \max_{n=1,\ldots,N} |\lambda_1 \xi_{1,n} + \cdots + \lambda_K \xi_{K,n}| .$$

But this contradicts the definition of the norm in ℓ_∞ (note that the right-hand side does not vanish for sufficiently large N, since we assumed that the vectors $x_1, \dots, x_N$ are linearly independent). This contradiction shows that we find indeed some N satisfying (10.4).

Together with $x_1, \dots, x_K$, we consider the truncated vectors $y_1, \dots, y_K \in \mathbb{R}^N$ where $y_k := (\xi_{k,1}, \dots, \xi_{k,N})$. If we equip $\mathbb{R}^N$ with the max-norm, we may read (10.4) as

$$\|\lambda_1 x_1 + \cdots + \lambda_K x_K\|_\infty \leq (1+\varepsilon)\ \|\lambda_1 y_1 + \cdots + \lambda_K y_K\|. \tag{10.6}$$

Recalling that $x_1, \dots, x_K$ are linearly independent, (10.6) implies that also the vectors $y_1, \dots, y_K$ are linearly independent. On the subspace of $\mathbb{R}^N$ spanned by $y_1, \dots, y_K$ we define a functional g by

$$g(\lambda_1 y_1 + \cdots + \lambda_K y_K) := f(\lambda_1 x_1 + \cdots + \lambda_K x_K).$$

Since $y_1, \dots, y_K$ are linearly independent, g is well-defined. Moreover, using (10.6), we find

$$\begin{aligned}\|g(\lambda_1 y_1 + \cdots + \lambda_K y_K)\| &\leq \|f\|\,\|\lambda_1 x_1 + \cdots + \lambda_K x_K\|_\infty \\ &\leq \|f\|\,(1+\varepsilon)\,\|\lambda_1 y_1 + \cdots + \lambda_K y_K\|.\end{aligned}$$

Hence $\|g\| \leq (1+\varepsilon)\,\|f\|$. By the Hahn-Banach extension theorem (actually Lemma 10.2 suffices), we may extend g to a linear functional on $\mathbb{R}^N$ without increasing its norm. Let $e_1, \dots, e_N$ be the canonical base of $\mathbb{R}^N$. Put now $\eta_n := g(e_n)$. Since g is linear, we have

$$f(x_k) = g(y_k) = g\left(\sum_{n=1}^{N} \xi_{k,n} e_n\right) = \sum_{n=1}^{N} \eta_n \xi_{k,n},$$

and so (10.3) holds. In order to prove the norm estimate, we consider the vector $x := (\operatorname{sgn}(\eta_1), \dots, \operatorname{sgn}(\eta_N))$. Then

$$\sum_{n=1}^{N} |\eta_n| = \sum_{n=1}^{N} \operatorname{sgn}(\eta_n) g(e_n) = g(x) \leq \|g\|\,\|x\| \leq (1+\varepsilon)\,\|f\|.$$

□

Using Lemma 10.4, we obtain now:

Theorem 10.5. *Let $* : \widehat{S} \to \widehat{{}^*S}$ be a nonstandard embedding. If $*$ is even an enlargement, then for any $f \in \ell_\infty^*$ there exists an internal *-finite sequence $\eta_1, \dots, \eta_h \in {}^*\mathbb{R}$ with*

$${}^*(f(x)) = \sum_{n=1}^{h} \eta_n {}^*\xi_n \qquad (x = (\xi_n)_n \in \ell_\infty), \tag{10.7}$$

where

$$\|f\| = \operatorname{st}\left(\sum_{n=1}^{h} |\eta_n|\right). \tag{10.8}$$

Conversely, each **-finite internal sequence for which the right-hand side of* (10.8) *is finite gives rise to a functional* $f \in \ell_\infty^*$ *defined by*

$$f(x) := \operatorname{st}\left(\sum_{n=1}^{h} \eta_n {}^*\xi_n\right) \qquad (x = (\xi_n)_n \in \ell_\infty), \tag{10.9}$$

which satisfies

$$\|f\| \le \operatorname{st}\left(\sum_{n=1}^{h} |\eta_n|\right).$$

Proof. Let $f \in \ell_\infty^*$ be given. Let $\varphi \in \widehat{S}$ be the following binary relation from $\mathbb{R}_+ \times \ell_\infty$ into $\mathbb{R}^{<\mathbb{N}}$:

$$\varphi = \{(\underline{\varepsilon}, \underline{x}, \underline{y}) \in \mathbb{R}_+ \times \ell_\infty \times \mathbb{R}^{<\mathbb{N}} \mid f(\underline{x}) = \sum_{\underline{n}=1}^{\#_{\mathbb{R}}(\underline{y})} \underline{y}(\underline{n})\underline{x}(\underline{n}) \wedge \sum_{\underline{n}=1}^{\#_{\mathbb{R}}(\underline{y})} |\underline{y}(\underline{n})| \le (1+\underline{\varepsilon})\,\|f\|\}$$

(here, we consider sequences as functions). Lemma 10.4 implies that φ is concurrent on $\mathbb{R}_+ \times \ell_\infty$, i.e. for each finitely many $(\varepsilon_k, x_k) \in \mathbb{R}_+ \times \ell_\infty$ $(k = 1, \ldots, N)$ there is some $y \in \mathbb{R}^{<\mathbb{N}}$ such that $(\varepsilon_k, x_k, y) \in \varphi$ $(k = 1, \ldots, N)$. Since $*$ is an enlargement, Theorem 8.10 implies that there is some $y \in {}^*\mathbb{R}^{<\mathbb{N}}$ which satisfies ${}^*\varphi$ on ${}^\sigma\mathrm{dom}(\varphi)$, i.e. $({}^*\varepsilon, {}^*x, y) \in {}^*\varphi$ for any $\varepsilon \in \mathbb{R}_+$ and any $x \in \ell_\infty$. By Exercise 25, y is a *-finite internal sequence $\eta_1, \ldots, \eta_h$ where $h := ({}^*\#_{\mathbb{R}}(\cdot))(y)$. We thus have (10.7), and

$$\sum_{n=1}^{h} |\eta_n| \le (1 + {}^*\varepsilon)^*(\|f\|) \qquad (\varepsilon \in \mathbb{R}_+).$$

Conversely, if $\eta_1, \ldots, \eta_h$ is a *-finite internal sequence and f is given by (10.9), then we have for each $x = (\xi_n)_n \in \ell_\infty$ that $|{}^*\xi_n| \le {}^*(\|x\|_\infty)$ $(n \in {}^*\mathbb{N})$ by the transfer principle, and so

$${}^*(|f(x)|) \le {}^*\varepsilon + \sum_{n=1}^{h} |\eta_n|\,|{}^*\xi_n| \le {}^*\varepsilon + \sum_{n=1}^{h} |\eta_n|\,{}^*(\,\|x\|_\infty),$$

for each $\varepsilon \in \mathbb{R}_+$. Moreover (10.9) implies that f is linear, since st is linear (Theorem 5.21). Hence, $f \in \ell_\infty^*$ and

$${}^*(\|f\|) \le {}^*\varepsilon + \sum_{n=1}^{h} |\eta_n| \qquad (\varepsilon \in \mathbb{R}_+).$$

□

Exercise 52. Let X be a linear space (not necessarily normed) of real sequences. Prove that any linear functional f on X can be written in the form

$$^*(f(x)) = \sum_{n=1}^{h} \eta_n \xi_n \qquad (x = (\xi_n)_n \in X),$$

where $\eta_1, \ldots, \eta_h$ is a *-finite internal sequence.

A functional $f \in \ell_\infty^*$ is called a *Hahn-Banach limit*, if for any convergent sequence $x = (\xi_n)_n$ the relation

$$f(x) = \lim_{n\to\infty} \xi_n$$

holds. Thus, Hahn-Banach limits are generalizations of the lim-operator which associate a "limit" to *any* bounded sequence.

Exercise 53. Prove that Hahn-Banach limits may not be written in the form (10.2) with some sequence $y = (\eta_n)_n$. Prove however, that Hahn-Banach limits exist (and so ℓ_∞^* indeed contains elements which cannot be written in the form (10.2)).

Theorem 10.6. *Let $* : \widehat{S} \to \widehat{^*S}$ be a nonstandard embedding. If $*$ is even an enlargement, then for any Hahn-Banach limit f there exist $h_0, h \in \mathbb{N}_\infty$, $h_0 < h$, and an internal sequence $\eta_{h_0}, \ldots, \eta_h \in {}^*\mathbb{R}$ such that*

$$^*(f(x)) = \sum_{n=h_0}^{h} \eta_n {}^*\xi_n \qquad (x = (\xi_n)_n \in \ell_\infty),$$

$$\sum_{n=h_0}^{h} \eta_n = 1,$$

and

$$\|f\| = \mathrm{st}\left(\sum_{n=h_0}^{h} |\eta_n|\right).$$

Conversely, each internal *-finite sequence $\eta_1, \ldots, \eta_h \in {}^*\mathbb{R}$ which satisfies $\eta_n \approx 0$ for any $n \in {}^\sigma\mathbb{N}$, $\eta_1 + \cdots + \eta_h \approx 1$, and for which $|\eta_1| + \cdots + |\eta_h|$ is finite defines a Hahn-Banach limit f by means of the formula

$$f(x) := \mathrm{st}\left(\sum_{n=1}^{h} \eta_n {}^*\xi_n\right) \qquad (x = (\xi_n)_n \in \ell_\infty).$$

Proof. Let f be a Hahn-Banach limit, and let $\eta_1, \ldots, \eta_h$ be the internal *-finite sequence of Theorem 10.5. The first statement follows if we can prove that there is some $h_0 \in \mathbb{N}_\infty$ with $\eta_n = 0$ for $n < h_0$, and $\eta_1 + \cdots + \eta_h = 1$. To see the latter, consider the constant sequence $x := (1)$ (i.e. $\xi_n := 1$). Since f is a Hahn-Banach limit, we must have $f(x) = 1$. But since ${}^*\xi_n = 1$ for each $n \in {}^*\mathbb{N}$ (this follows by the transfer principle or by Theorem 7.1), the formula (10.7) implies $\eta_1 + \cdots + \eta_h = 1$, as claimed. To see that $\eta_{{}^*n} = 0$ for $n \in \mathbb{N}$, we consider the particular sequence x defined by $\xi_k := 0$ for $k \neq n$ and $\xi_n := 1$. The transfer principle implies ${}^*\xi_k = 0$ for $k \neq {}^*n$ and ${}^*\xi_{{}^*n} = 1$. Since $f(x) = 0$, the formula (10.7) implies $\eta_{{}^*n} = 0$, as claimed. We thus have proved that the internal formula $\eta_{\underline{n}} = 0$ holds for each $\underline{n} \in {}^\sigma\mathbb{N}$. By the permanence principle there is some $h_1 \in \mathbb{N}_\infty$ such that $\eta_{\underline{n}} = 0$ holds for all $\underline{n} \leq h_1$. Thus, the first statement follows with $h_0 := h_1 + 1$.

For the second statement, let $\eta_1, \ldots, \eta_h$ and f be given as in the formulation of the theorem. Theorem 10.5 implies that $f \in \ell_\infty^*$. It remains to prove that if $x = (\xi_n)_n$ converges to some $l \in \mathbb{R}$, that $f(x) = l$, i.e. ${}^*l \approx \sum \eta_n {}^*\xi_n$. Let $\varepsilon \in \mathbb{R}_+$ be given. Since $\eta_n {}^*\xi_n \approx 0$ for each $n \in {}^\sigma\mathbb{N}$, the internal predicate

$$\sum_{\underline{n}=1}^{\underline{x}} |\eta_{\underline{n}}| < {}^*\varepsilon$$

is true for any $\underline{x} \in {}^\sigma\mathbb{N}$ and by the permanence principle thus also for some $\underline{x} = h_0 \in \mathbb{N}_\infty$. For $n > h_0$, we have ${}^*\xi_n \approx {}^*l$ by Theorem 7.1, in particular $|{}^*\xi_n - {}^*l| < {}^*\varepsilon$. Putting $c := \eta_1 + \cdots + \eta_n - 1$, we have by assumption $c \approx 0$ and thus also $|c| < {}^*\varepsilon$. Now we may calculate

$$\begin{aligned} \left|\sum_{n=1}^{h} \eta_n {}^*\xi_n - {}^*l\right| &= \left|\sum_{n=1}^{h} \eta_n {}^*\xi_n - \left(\sum_{n=1}^{h} \eta_n - c\right) {}^*l\right| \leq \left|\sum_{n=1}^{h} \eta_n ({}^*\xi_n - {}^*l)\right| + c\,|{}^*l| \\ &\leq \sum_{n=1}^{h_0} |\eta_n| \left(|{}^*\xi_n| + |{}^*l|\right) + \sum_{n=h_0+1}^{h} |\eta_n|\, {}^*\varepsilon + {}^*\varepsilon\, |{}^*l|\,. \end{aligned}$$

Now observe that the transfer principle implies $|{}^*\xi_n| \leq {}^*(\|x\|_\infty)$ for all $n \in {}^*\mathbb{N}$ and that by assumption $|\eta_1| + \cdots + |\eta_h| \leq {}^*M$ for some $M \in \mathbb{R}_+$. Hence, we have proved

$$\left|\sum_{n=1}^{h} \eta_n {}^*\xi_n - {}^*l\right| \leq {}^*\varepsilon({}^*(\|x\|_\infty) + |{}^*l|) + {}^*M {}^*\varepsilon + {}^*\varepsilon\, |{}^*l|\,.$$

Since this estimate holds for any $\varepsilon \in \mathbb{R}_+$, it follows that $\sum \eta_n {}^*\xi_n \approx {}^*l$, which we had to prove. □

Theorem 10.6 slightly generalizes [Lux92, Theorem 4.4], using a refinement of the technique from [Rob64].

Exercise 54. Does there exist a Hahn-Banach limit f such that for each $x \in \ell_\infty$ the point $f(x)$ is an accumulation point of the sequence x?

Another generalization of limits is the following: A linear functional f on ℓ_∞ is called a *Banach-Mazur limit*, if it has the following properties:

1. If $x = (\xi_n)_n$ is the constant sequence $\xi_n = c$, then $f(x) = c$.
2. f is positive, i.e. if $x = (\xi_n)_n \in \ell_\infty$ satisfies $\xi_n \geq 0$ for all n, then $f(x) \geq 0$.
3. f is shift invariant, i.e. $f((\xi_n)_n) = f((\xi_{n+1})_n)$.

Actually, we could have required the first property even for all convergent sequences $\xi_n \to c$ (which is apparently a more restrictive requirement):

Exercise 55. Prove that any Banach-Mazur limit f is a Hahn-Banach limit. More precisely, show that $f \in \ell_\infty^*$ with $\|f\| = 1$ and that for any $x = (\xi_n)_n \in \ell_\infty$ the estimate

$$\liminf_{n\to\infty} \xi_n \leq f(x) \leq \limsup_{n\to\infty} \xi_n$$

holds.

Exercise 56. Let f be a Banach-Mazur limit. Calculate $f(x)$ for the sequence $x = (\xi_n)_n$ which is given by $\xi_n := (-1)^n$.

Does there exist a Banach-Mazur limit which has the additional property from Exercise 54, i.e. such that $f(x)$ is always an accumulation point of the sequence x?

The standard proofs for the existence of Banach-Mazur limits are not very constructive. We just mention one of the simplest standard approaches from [Rud90, Chapter 3, Exercise 4]:

Exercise 57. Given a sequence $x = (\xi_n)_n \in \ell_\infty$, put $\zeta_n := (\xi_1 + \cdots + \xi_n)/n$. Apply Exercise 49 for $p(x) := \limsup \zeta_n$ and $f(x) := \lim \zeta_n$ (if ζ_n converges) to prove the existence of a Banach-Mazur limit.

We now present a class of Banach-Mazur limits which can easily be characterized by nonstandard methods. This result is taken from [Rob64]:

Theorem 10.7. *Let* $* : \widehat{S} \to \widehat{{}^*S}$ *be a nonstandard embedding, and* $\eta_1, \ldots, \eta_h \in {}^*\mathbb{R}$ *be an internal* **-finite sequence. Assume:*

1. $\eta_n \geq 0$ *for* $n = 1, \ldots, h$.
2. $\eta_1 + \cdots + \eta_h \approx 1$.
3. $\sum_{n=1}^{h} |\eta_n - \eta_{n-1}| \approx 0$ *(put* $\eta_0 := 0$*).*

Then a Banach-Mazur limit f *is given by the formula*

$$f(x) := \operatorname{st}\left(\sum_{n=1}^{h} \eta_n {}^*\xi_n\right) \qquad (x = (\xi_n)_n \in \ell_\infty).$$

Proof. An induction implies that $\eta_n \approx 0$ for each $n \in {}^\sigma\mathbb{N}$: Indeed, if $\eta_{n-1} \approx 0$, then $|\eta_n - \eta_{n-1}| \approx 0$ (by assumption) shows that also $\eta_n \approx 0$. Hence, by Theorem 10.6,

f is a Banach-Mazur limit. Moreover, positivity of f is trivial, since $\eta_n \geq 0$. It thus only remains to prove that f is shift-invariant. This can be seen as follows: If $x = (\xi_n)_n \in \ell_\infty$ is given, we have $|{}^*\xi_n| \leq {}^*(\|x\|_\infty)$ by the transfer principle. Putting $y := (\xi_{n+1})_n$, we have

$$\begin{aligned} {}^*(|f(x) - f(y)|) &\approx \left| \sum_{n=1}^{h} \eta_n {}^*\xi_n - \sum_{n=1}^{h} \eta_n {}^*\xi_{n+1} \right| = \left| \eta_1 {}^*\xi_1 + \sum_{n=2}^{h} (\eta_n - \eta_{n-1})\xi_n \right| \\ &\leq |\eta_1 {}^*\xi_1| + \sum_{n=2}^{h} |\eta_n - \eta_{n-1}| \, {}^*(\|x\|_\infty) \approx 0, \end{aligned}$$

and so $f(x) = f(y)$, as desired. □

Example 10.8. Let $* : \widehat{S} \to \widehat{{}^*S}$ be a nonstandard embedding. Fix some $h \in \mathbb{N}_\infty$. Then

$$f(x) := \mathrm{st}\left(\frac{1}{h} \sum_{n=1}^{h} {}^*\xi_n \right) \qquad (x = (\xi_n)_n \in \ell_\infty)$$

defines a Banach-Mazur limit.

We remark that Theorem 10.7 has the following powerful converse:

Theorem 10.9. *Let $* : \widehat{S} \to \widehat{{}^*S}$ be polysaturated. Then for any Banach-Mazur limit f there exist $h_0, h \in \mathbb{N}_\infty$, $h_0 < h$, and an internal sequence $\eta_{h_0}, \ldots, \eta_h \in {}^*\mathbb{R}$ such that*

$$f(x) = \mathrm{st}\left(\sum_{n=h_0}^{h} \eta_n {}^*\xi_n \right) \qquad (x = (\xi_n)_n \in \ell_\infty),$$

$$\eta_n \geq 0,$$

$$\sum_{n=h_0}^{h} \eta_n \approx 1,$$

and

$$\sum_{n=h_0+1}^{h} |\eta_n - \eta_{n-1}| \approx 0.$$ □

The proof of Theorem 10.9 needs deeper facts about the geometry of Banach spaces which are beyond the scope of this book. The proof can be found in [Lux92].

Theorem 10.9 suggests that the following Hahn-Banach limits are particularly "natural":

Example 10.10. Let $* : \widehat{S} \to \widehat{{}^*S}$ be a nonstandard embedding. For any $h_0, h \in \mathbb{N}_\infty$ with $h_0 \leq h$, the formula

$$f(x) := \mathrm{st}\left(\frac{1}{h - h_0 + 1} \sum_{n=h_0}^{h} {}^*\xi_n \right) \qquad (x = (\xi_n)_n \in \ell_\infty)$$

defines a Banach-Mazur limit (by Theorem 10.7).

The Banach-Mazur limits of Example 10.10 are said to be *of Cesàro type*.

For certain sequences $x \in \ell_\infty$, the value $f(x)$ is independent of the choice of the Banach-Mazur limit f. For example, if x converges to c, we must always have $f(x) = c$. One may wonder whether this property holds for a larger class of sequences x. This is indeed true: The sequences with this property have been characterized by Lorentz [Lor48], but the (standard) proof is rather cumbersome. A brief nonstandard proof of this result was given in [Lux92]. We present a variant of this proof now. The essential observation is that it suffices to consider Banach-Mazur limits of Cesàro type:

Proposition 10.11. *Let* $* : \widehat{S} \to \widehat{{}^*S}$ *be an enlargement. Let* $x_0 \in \ell_\infty$ *be such that* $f(x_0) = c$ *for any Banach-Mazur limit* f *of Cesàro type. Then* $f(x_0) = c$ *for any shift-invariant Hahn-Banach limit* f. *In particular,* $f(x_0) = c$ *for any Banach-Mazur limit* f.

Proof. Let $S : \mathbb{N} \times \ell_\infty \to \ell_\infty$ be defined by

$$S(n, x) := (\xi_{n+1}, \xi_{n+2}, \ldots) \qquad (x = (\xi_k)_k \in \ell_\infty).$$

Let f be a shift-invariant Hahn-Banach limit. Then the value $c_0 := f(S(n, x_0))$ is independent of $n \in \mathbb{N}$. The transfer principle implies

$$\forall \underline{n} \in {}^*\mathbb{N} : {}^*f({}^*S(\underline{n}, {}^*x_0)) = {}^*c_0.$$

By Theorem 10.6, we find an internal sequence $\eta_{h_0}, \ldots, \eta_h \in {}^*\mathbb{R}$ (with $h_0, h \in \mathbb{N}_\infty$) with $\sum \eta_k = 1$ and $\sum |\eta_k| \le {}^*(\|f\|) + 1$ such that

$${}^*f({}^*x) = \sum_{k=h_0}^{h} \eta_k {}^*\xi_k \qquad (x = (\xi_n)_n \in \ell_\infty).$$

Since ${}^*f({}^*S(n, {}^*x_0)) = {}^*c_0$ for $n = h_0, \ldots, h$, we find

$$(h - h_0 + 1){}^*c_0 = \sum_{n=h_0}^{h} {}^*f({}^*S(n, {}^*x_0)) = \sum_{n=h_0}^{h} \sum_{k=h_0}^{h} \eta_k {}^*\xi_{n+k},$$

and so

$${}^*c_0 = \sum_{k=h_0}^{h} \eta_k \left(\frac{1}{h - h_0 + 1} \sum_{n=h_0}^{h} {}^*\xi_{n+k} \right) = \sum_{k=h_0}^{h} \eta_k ({}^*c + \varepsilon_k)$$

where $\varepsilon_k \in \inf({}^*\mathbb{R})$. Since $\sum \eta_k = 1$, we find for any $\varepsilon \in \mathbb{R}_+$ that

$$|{}^*c_0 - {}^*c| = \left| \sum_{k=h_0}^{h} \eta_k \varepsilon_k \right| \le \sum_{k=h_0}^{h} |\eta_k| \, {}^*\varepsilon \le {}^*(\|f\| + 1){}^*\varepsilon,$$

and so ${}^*c_0 \approx {}^*c$ which implies $f(x) = c_0 = c$. □

A deeper reason why Proposition 10.11 is true is revealed in [KM92] (see also the remarks in [Lux92]). However, we will not go into further detail here.

Definition 10.12. A sequence $x = (\xi_n)_n$ is *almost convergent* to c if we have

$$\lim_{n\to\infty} \frac{1}{n}\sum_{m=1}^{n} \xi_{m+k} = c$$

uniformly in $k \in \mathbb{N}$.

It can be proved by standard methods that any almost convergent sequence is bounded. However, with nonstandard methods an easier proof can be given [Lux92]:

Exercise 58. Let $*$ be a nonstandard embedding. Then $x = (\xi_n)_n$ is almost convergent to c if and only if

$$\frac{1}{h}\sum_{n=1}^{h} {}^*\xi_{n+k} \approx {}^*c \qquad (h \in \mathbb{N}_\infty,\ k \in {}^*\mathbb{N}). \tag{10.10}$$

Moreover, in this case $x \in \ell_\infty$.

Hint: For the second statement, prove that ${}^*x \in {}^*\ell_\infty$.

Now we can prove the announced result:

Theorem 10.13. *A sequence $x = (\xi_n)_n$ is almost convergent to c if and only if $x \in \ell_\infty$ and $f(x) = c$ for any Banach-Mazur limit f.*

Proof. Let $* : \widehat{S} \to \widehat{{}^*S}$ be an enlargement, and x be almost convergent to c. Then $x \in \ell_\infty$ by Exercise 58. By Proposition 10.11, it suffices to prove that $f(x) = c$ for any Banach-Mazur limit f of Cesàro type, i.e. we may assume

$$f(x) = \mathrm{st}\left(\frac{1}{h-h_0+1}\sum_{m=h_0}^{h} {}^*\xi_m\right) \tag{10.11}$$

for some $h_0, h \in \mathbb{N}_\infty$ with $h_0 \le h$. Putting $n := h - h_0 + 1$ and $k := h_0 - 1$, we have by (10.10)

$$\frac{1}{h-h_0+1}\sum_{m=h_0}^{h} {}^*\xi_m = \frac{1}{n}\sum_{m=1}^{n} {}^*\xi_{m+k} \approx {}^*c,$$

and so $f(x) = c$, as desired.

Conversely, let $x \in \ell_\infty$ and $f(x) = c$ for any Banach-Mazur limit f. Let $* : \widehat{S} \to \widehat{{}^*S}$ be a nonstandard map. Given $h \in \mathbb{N}_\infty$ and $k \in {}^*\mathbb{N}$, we find by considering the Banach-Mazur limit

$$f((\zeta_n)_n) := \mathrm{st}\left(\frac{1}{h}\sum_{m=k+1}^{h+k} {}^*\zeta_m\right)$$

(Theorem 10.7) that

$$^*c \approx \frac{1}{h}\sum_{m=k+1}^{h+k} \xi_m = \frac{1}{h}\sum_{m=1}^{h} \xi_{m+k},$$

and so x is almost convergent to c by Exercise 58. □

Corollary 10.14. *If x is convergent to c, then x is almost convergent to c.*

The converse of Corollary 10.14 does not hold: For example, the sequence $x_n := (-1)^n$ is almost convergent to 0 but not convergent. One may ask for additional conditions which ensure that a sequence which is "almost convergent" (e.g. in our sense) is actually convergent. Theorems giving such conditions are called *Tauberian theorems*. In our case, it turns out that there is a simple Tauberian theorem giving a condition which is even necessary and sufficient. This was already observed in [Lor48]; the following easier nonstandard proof is taken from [Lux92]:

Theorem 10.15. *The sequence $x = (\xi_n)_n$ is convergent to c if and only if it is almost convergent to c and $\xi_n - \xi_{n+1} \to 0$.*

In particular, a series $\sum a_n$ converges if and only if the partial sums are almost convergent, and $a_n \to 0$.

Proof. One implication follows from Corollary 10.14. For the converse implication, let x be almost convergent to c and $\xi_{n+1} - \xi_{n+1} \to 0$. Let $* : \widehat{S} \to \widehat{^*S}$ be a nonstandard map. Given $h \in \mathbb{N}_\infty$, consider the internal sequence given by

$$\eta_n := {}^*\xi_h - {}^*\xi_{n+h} = \sum_{m=1}^{n} ({}^*\xi_{m+h-1} - {}^*\xi_{m+h}).$$

Since ${}^*\xi_{m+h-1} - {}^*\xi_{m+h} \approx 0$ by Theorem 7.1, we have $\eta_n \approx 0$ for any finite $n \in {}^\sigma\mathbb{N}$. Robinson's sequential lemma (Exercise 22) implies that there is some $h_0 \in \mathbb{N}_\infty$ with $\eta_n \approx 0$ for all $n \in {}^*\mathbb{N}$ with $n \leq h_0$. In particular, we have for any $\varepsilon \in \mathbb{R}_+$ that

$$\frac{1}{h_0}\left|\sum_{n=1}^{h_0} \eta_n\right| \leq \frac{1}{h_0} h_0 {}^*\varepsilon.$$

In view of (10.10), this implies

$$0 + {}^*c \approx \frac{1}{h_0}\sum_{n=1}^{h_0} \eta_n + \frac{1}{h_0}\sum_{n=1}^{h_0} {}^*\xi_{n+h} = {}^*\xi_h,$$

and so $\xi_n \to c$ by Theorem 7.1. □

§11 Additive Measures

Let S_0 be a set, and Σ be a *set algebra* over S_0, i.e. Σ is a system of subsets of S_0 with the property that $S_0 \in \Sigma$ and that $A, B \in \Sigma$ implies $A \cup B \in \Sigma$ and $S_0 \setminus A \in \Sigma$. A function $\mu : \Sigma \to [0, \infty]$ is called an *additive measure* if $\mu(A \cup B) = \mu(A) + \mu(B)$ for any $A, B \in \Sigma$ with $A \cap B = \varnothing$. If $\mu(S_0) = 1$, then μ is called an *additive probability measure*. If Σ is even a *σ-algebra*, i.e. additionally $\bigcup A_n \in \Sigma$ for countably many $A_n \in \Sigma$ and μ is even *σ-additive*, i.e. $\mu(\bigcup A_n) = \sum \mu(A_n)$ whenever $A_n \in \Sigma$ are pairwise disjoint, then μ is called a *measure* resp. a *probability measure*.

An additive measure μ is called *singular*, if $\mu(A) = 0$ for any finite set $A \in \Sigma$.

At the moment, we are interested in additive probability measures on $\mathbb{N}$ where $\Sigma = \mathcal{P}(\mathbb{N})$. Such a measure μ is singular if and only if $\mu(\{n\}) = 0$ for any $n \in \mathbb{N}$. It cannot be proved without the axiom of choice that such measures exist (even a rather powerful countable version of the axiom of choice is not sufficient [PS77]). In particular, it is not possible by standard methods to *construct* singular additive measures on $\mathbb{N}$. However, many Hahn-Banach limits f provide such a measure: Given $A \subseteq \mathbb{N}$, we let χ_A denote the sequence $a_n \in \mathbb{R}$ defined by

$$a_n := \begin{cases} 1 & \text{if } n \in A, \\ 0 & \text{if } n \notin A. \end{cases}$$

Then $\mu(A) := f(\chi_A)$ is additive, because if $A \cap B = \varnothing$, we have

$$\mu(A \cup B) = f(\chi_{A \cup B}) = f(\chi_A + \chi_B) = f(\chi_A) + f(\chi_B) = \mu(A) + \mu(B).$$

Moreover, $\mu(\{n\}) = 0$, because $\chi_{\{n\}}$ is a null sequence. The only property which is not necessarily satisfied is that $\mu(A) \geq 0$. However, this holds if we choose some f which has a representation as in Theorem 10.5 with $\eta_n \geq 0$. Hence, we found a rather large class of singular additive measures on $\mathbb{N}$.

In particular, if f is a Banach-Mazur limit, then $\mu(A) = f(\chi_A)$ defines a singular measure which additionally is *translation invariant*, i.e. $\mu(\{n : n + 1 \in A\}) = \mu(A)$.

By an appropriate choice, we can satisfy certain other additional properties. Let us give a sample application:

A set $A \subseteq \mathbb{N}$ is said to have a *density* d, if

$$d := \lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} a_n \qquad ((a_n)_n = \chi_A)$$

exists. The density (if it exists) may be considered as a "relative frequency" of the occurrences of 1 in the sequence $(a_n)_n$. A singular measure μ with the property that $\mu(A) = d$ whenever A has the density d may be considered as some sort of "Laplace measure" on $\mathbb{N}$ (i.e. each number has in a certain sense the same "weight" for the calculation of the probability).

Theorem 11.1. *There is a singular additive translation invariant measure μ on $\mathbb{N}$ with the additional property that $\mu(A) = d$ whenever A has the density d. Moreover, for any $A \subseteq X$ the sequence $\chi_A = (a_n)_n$ satisfies*

$$\liminf_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} a_k \le \mu(A) \le \limsup_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} a_k.$$

Proof. Let $* : \widehat{S} \to \widehat{{}^*S}$ be a nonstandard embedding. Fix $h_0 \in \mathbb{N}_\infty$ and choose some $h \in \mathbb{N}_\infty$ such that h/h_0 is infinite (put e.g. $h := h_0^2$). Consider the Banach-Mazur limit

$$f(x) := \operatorname{st}\left(\frac{1}{h-h_0}\sum_{k=h_0}^{h} {}^*\xi_k\right) \qquad (x = (\xi_k)_k)$$

of Cesàro type, and put $\mu(A) := f(\chi_A)$. Since f is a Banach-Mazur limit, μ is a singular translation invariant measure.

To see that the additional property holds, let $\chi_A = (a_n)_n$, and consider the sequence

$$b_n := \frac{1}{n}\sum_{k=1}^{n} a_k.$$

Exercise 28 implies that

$$\liminf_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} a_k \le \operatorname{st}({}^*b_h) \le \limsup_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} a_k.$$

Hence, it suffices to prove that $f(\chi_A) = \operatorname{st}({}^*b_h)$. But we have

$${}^*(f(\chi_A)) \approx \frac{1}{h-h_0}\sum_{k=h_0}^{h} {}^*a_k = \frac{h}{h-h_0}{}^*b_h - \frac{1}{h-h_0}\sum_{k=1}^{h_0-1} {}^*a_k.$$

Note now that $h/(h-h_0) = \frac{h/h_0}{h/h_0-1} \approx 1$, and

$$\left|\frac{1}{h-h_0}\sum_{k=1}^{h_0-1} {}^*a_k\right| \le \frac{h_0-1}{h-h_0} = \frac{h_0/h - 1/h}{1-h_0/h} \approx 0,$$

and so ${}^*f(\chi_A) \approx {}^*b_h$, as desired. □

Exercise 59. In the proof of Theorem 11.1, we have chosen a particular Banach-Mazur limit f of Cesàro type. Does the conclusion $\mu(A) = d$ also hold for any Banach-Mazur limit f of Cesàro type or even for any Banach-Mazur limit f?

Hint: Apply Theorem 10.13.

It turns out that in the nonstandard world any singular measure is a Laplace measure in another sense. This was first proved in [Hen72b]. We present a slightly modified version:

Theorem 11.2. *Let* $* : \widehat{S} \to \widehat{{}^*S}$ *be a nonstandard embedding,* $S_0 \in \widehat{S}$ *be an entity, and* Σ *be an algebra over* S_0*. Then for any nonempty* **-finite* $B \subseteq {}^*S_0$ *the function*

$$\mu(A) := \operatorname{st}\left(\frac{{}^{\#}({}^*A \cap B)}{{}^{\#}B}\right) \qquad (A \subseteq S_0) \tag{11.1}$$

defines an additive probability measure (even for $\Sigma = \mathcal{P}(S_0)$*). Moreover, if* B *is infinite, then* μ *is singular.*

Conversely, if $*$ *is even a* Σ*-enlargement, then any singular additive probability measure* μ *can be written in the above form.*

Proof. Clearly, $\mu(A) \geq 0$, and $\mu(S_0) = 1$. If $A_1, A_2 \in \Sigma$ are disjoint, then also $B_1 := {}^*A_1 \cap B$ and $B_2 := {}^*A_2 \cap B$ are disjoint, and Theorem 6.14 implies ${}^{\#}(B_1 \cup B_2) = {}^{\#}B_1 + {}^{\#}B_2$. Hence,

$$\frac{{}^{\#}({}^*(A_1 \cup A_2) \cap B)}{{}^{\#}B} = \frac{{}^{\#}({}^*A_1 \cap B)}{{}^{\#}B} + \frac{{}^{\#}({}^*A_1 \cap B)}{{}^{\#}B},$$

and the additivity of st implies that μ is additive. If A is finite and B is infinite, then ${}^*A$ and thus ${}^{\#}({}^*A \cap B) \in {}^{\sigma}\mathbb{N}$ is finite and ${}^{\#}B \in \mathbb{N}_\infty$ is infinite, and so $\mu(A) = 0$.

For the second statement, consider the relation

$$\varphi := \{(\underline{x}, \underline{y}) \in \Sigma \times \mathcal{P}(\Sigma) \mid \text{“}\underline{y} \text{ is finite”} \wedge \alpha(\underline{y}) \wedge \beta(\underline{x})\}$$

where $\alpha(\underline{y})$ is a shortcut for “each two different elements of $\underline{y}$ are disjoint”, and $\beta(\underline{y})$ is a shortcut for

$$\exists \underline{z} \in \mathcal{P}(\Sigma) : (\underline{z} \subseteq \underline{y} \wedge \underline{x} = \bigcup \underline{z}).$$

Then φ is concurrent: Given $A_1, \ldots, A_n \in \Sigma$, let Σ_1 denote the system of all finite intersections $A_{n_1} \cap \cdots \cap A_{n_k}$, and eliminate successively all elements which can be written as the union of other elements. The resulting set $\underline{y} = \Sigma_2$ satisfies $(A_1, \Sigma_2), \ldots, (A_n, \Sigma_2) \in \varphi$. Theorem 8.10 implies that ${}^*\varphi$ is satisfied on ${}^{\sigma}\Sigma$. In view of the standard definition principle for relations and Theorem 3.21, this means that we find some *-finite subset $\Sigma_0 := \underline{y} \subseteq {}^*\Sigma$ of pairwise disjoint sets such that for each $A \in \Sigma$, we find an internal $\Sigma_A \subseteq \Sigma_0$ with ${}^*A = \bigcup \Sigma_A$.

Let $c : \mathcal{P}(S_0) \to \mathbb{N} \cup \{\infty\}$ be the function which associates to each $A \subseteq S_0$ its number of elements. By Exercise 27, we have ${}^*c(A_0) = {}^{\#}(A_0)$ for any $A_0 \in {}^*\mathcal{P}(S_0)$.

To each $n \in \mathbb{N}$ and each infinite set $z \in \Sigma$, we can associate a subset $x(z) \subseteq z$ whose number of elements is the smallest integer which is at least $n\mu(z)$. Actually, this holds also if $z \in \Sigma$ is finite, because in this case $\mu(z) = 0$. Hence,

$$\forall \underline{n} \in \mathbb{N}, \underline{y} \in \mathcal{P}(\Sigma) : \exists \underline{x} \in \mathcal{P}(S_0)^{\Sigma} : \gamma$$

where γ is a shortcut for

$$\forall \underline{z} \in \Sigma : (\underline{x}(\underline{z}) \subseteq \underline{z} \wedge c(\underline{y}(\underline{z})) \leq \underline{n}\mu(\underline{z}) < c(\underline{y}(\underline{z})) + 1).$$

The transfer principle implies for the choice $\underline{y} := \Sigma_0$ in view of Theorem 3.21 that we find for any $h \in {}^*N$ some internal function $f : {}^*\Sigma \to {}^*\mathcal{P}(S_0)$ such that for any $A_0 \in \Sigma_0$ the set $f(A_0)$ is a *-finite subset of A_0 with

$$^{\#}(f(A_0)) \leq h^*\mu(A_0) < {}^{\#}(f(A_0)) + 1. \tag{11.2}$$

Fix some $h \in \mathbb{N}_\infty$ such that $h/{}^{\#}\Sigma_0$ is infinite, and let f denote the corresponding function. We claim that

$$B := \bigcup \{f(A_0) : A_0 \in \Sigma_0\}$$

has the required properties. Indeed, given $A \in \Sigma$, we find by our construction of Σ_0 an internal $\Sigma_A \subseteq \Sigma_0$ with ${}^*A = \bigcup \Sigma_A$. Theorem 6.13 implies that Σ_A is *-finite with ${}^{\#}\Sigma_A \leq {}^{\#}\Sigma_0$. Since the transfer principle implies that ${}^*\mu$ is additive on *-finite subsets of ${}^*\Sigma$ and since the elements of Σ_A are pairwise disjoint, we find

$$^*(\mu(A)) = {}^*\mu({}^*A) = {}^*\mu(\bigcup \Sigma_A) = \sum_{A_0 \in \Sigma_A} {}^*\mu(A_0).$$

Since (11.2) implies

$$\left| \sum_{A_0 \in \Sigma_A} ({}^{\#}(f(A_0)) - h^*\mu(A_0)) \right| \leq {}^{\#}\Sigma_A \leq {}^{\#}\Sigma_0,$$

we find

$$\left| h^*(\mu(A)) - \sum_{A_0 \in \Sigma_A} {}^{\#}(f(A_0)) \right| \leq {}^{\#}\Sigma_0.$$

Since $f(A_0) \subseteq A_0$ and since the sets $A_0 \in \Sigma_0$ are pairwise disjoint and ${}^*A = \bigcup \Sigma_A$, we find in view of the definition of B that ${}^*A \cap B = \bigcup \{f(A_0) : A_0 \in \Sigma_A\}$ where the union is pairwise disjoint. In particular,

$$^{\#}({}^*A \cap B) = \sum_{A_0 \in \Sigma_A} {}^{\#}(f(A_0)).$$

Summarizing,

$$\left|h^*(\mu(A)) - {}^{\#}({}^*A \cap B)\right| \leq {}^{\#}\Sigma_0 \qquad (A \in \Sigma). \tag{11.3}$$

Applying (11.3) for $A = S$, we find in view of $\mu(S) = 1$ that

$$\left|h - {}^{\#}B\right| \leq {}^{\#}\Sigma_0.$$

Dividing this estimate by ${}^{\#}\Sigma_0$, we find that ${}^{\#}B/{}^{\#}\Sigma_0 \geq (h/{}^{\#}\Sigma_0) - 1$ is infinite. (In particular, B is actually infinite).

Moreover, for any $A \in \Sigma$ we have by (11.3) in view of ${}^*\mu(A) \leq 1$ that

$$\left|\frac{{}^{\#}({}^*A \cap B)}{{}^{\#}B} - {}^*(\mu(A))\right| = \frac{\left|({}^{\#}({}^*A \cap B) - h^*\mu(A)) + (h - {}^{\#}B)^*\mu(A)\right|}{{}^{\#}B}$$
$$\leq \frac{2{}^{\#}\Sigma_0}{{}^{\#}B} \approx 0.$$

We thus have the required representation. □

Corollary 11.3 (Measure Extension Theorem)**.** *If μ is a singular additive probability measure defined on an algebra $\Sigma \subseteq S_0$, then μ may be extended to a singular additive probability measure on $\mathcal{P}(S_0)$.*

Proof. Let $*$ be a Σ-enlargement. Then we may write μ in the form (11.1), and the latter defines even a singular additive probability measure on $\mathcal{P}(S_0)$. □

In particular, the Lebesgue measure (defined on the measurable subsets of $[0,1]$) has an extension to an additive measure which is defined on all subsets of $[0,1]$. A famous theorem of Banach states that the Lebesgue measure on $\mathbb{R}$ has an extension to a translation invariant additive measure μ which is defined on all subsets of $\mathbb{R}$ (translation invariance means $\mu(A+x) = \mu(A)$ for each $A \subseteq \mathbb{R}$).

We intend to give a nonstandard proof for this result. By calculating modulo 1, it obviously suffices to extend the Lebesgue measure on $[0,1)$ to a measure μ which has the property that $\mu(A) = \mu(A \oplus x)$ for any $A \subseteq [0,1]$ and any $x \in \mathbb{R}$ (here, $A \oplus x := \{a \oplus x : a \in A\}$ and $a \oplus x := a + x + z$ where $z \in \mathbb{Z}$ is chosen such that $a \oplus x \in [0,1)$).

This result follows rather immediately from the following result of the standard world.

Proposition 11.4. *The operation $\oplus$ satisfies* Følner's condition *on $X = [0,1)$, i.e. for each $x_1, \ldots, x_n \in [0,1)$ and each $\varepsilon \in \mathbb{R}_+$ there is a nonempty finite set $A \subseteq [0,1)$ such that*

$$\frac{|A\Delta(A \oplus x_j)|}{|A|} < \varepsilon \qquad (j = 1, \ldots, n) \tag{11.4}$$

where Δ *denotes the* symmetric difference $A\Delta B := (A \setminus B) \cup (B \setminus A)$.

Proof. Without loss of generality, let $x_1 = 0$. We define $kx := x \oplus \cdots \oplus x$ (k times) and $0x := 0$ and prove by induction on n that the set A may even be chosen such that each $a \in A$ may be written in the form $a = k_1 x_1 \oplus \cdots \oplus k_n x_n$.

Since $x_1 = 0$, the induction start is trivial. Assume the claim has already been proved for $n - 1$. Let $\varepsilon \in \mathbb{R}_+$ and $x_1, \ldots, x_n$ be given. We distinguish two cases: First assume that there exist numbers $k_1, \ldots, k_n \in \mathbb{N} \cup \{0\}$, $k_n \neq 0$, with $k_n x_n = k_1 x_1 \oplus \cdots \oplus k_{n-1} x_{n-1}$. By induction hypothesis, we find a nonempty finite set A satisfying (for any $y \in [0, 1)$):

$$|(y \oplus A)\Delta(y \oplus A \oplus x_j)| < |A|\, \varepsilon_0 \qquad (j = 1, \ldots, n-1) \tag{11.5}$$

where $\varepsilon_0 := \varepsilon/(k_1 + \cdots + k_{n-1})$. Put $A_0 := A \cup (A \oplus x_n) \cup \cdots \cup (A \oplus (k_n - 1)x_n)$. Summing up (11.5) for $y = 0, x_n, \ldots, (k_n - 1)x_n$, we obtain

$$|A_0 \Delta(A_0 \oplus x_j)| < k_n\, |A|\, \varepsilon_0 \leq |A_0|\, \varepsilon \qquad (j = 1, \ldots, n-1).$$

Moreover, a successive application of (11.5) implies

$$\begin{aligned} |A_0 \Delta(A_0 \oplus k_n x_n)| &= |A_0 \Delta(A_0 \oplus k_1 x_1 \oplus \cdots \oplus k_{n-1} x_{n-1})| \\ &< (k_1 + \cdots + k_{n-1})\, |A|\, \varepsilon_0. \end{aligned}$$

The definition of A_0 implies $A_0 \Delta(A_0 \oplus x_n) = A\Delta(A \oplus k_n x_n)$, and so

$$|A_0 \Delta(A_0 \oplus x_n)| < (k_1 + \cdots + k_{n-1})\, |A|\, \varepsilon_0 \leq |A_0|\, \varepsilon.$$

Hence, we are done in this case.

If no numbers $k_1, \ldots, k_n$ as assumed above exist, choose $k > 2/\varepsilon$. By induction assumption, we find a nonempty finite set A such that (11.5) holds with $\varepsilon_0 := \varepsilon$. Put $A_0 := A \cup (A \oplus x_n) \cup \cdots \cup (A \oplus (k-1)x_n)$. By assumption, this is a union of disjoint sets, and so $|A_0| = k\, |A|$. From (11.5), it follows that

$$|A_0 \Delta(A_0 \oplus x_j)| < k\, |A|\, \varepsilon_0 = |A_0|\, \varepsilon \qquad (j = 1, \ldots, n-1).$$

Moreover, since $A_0 \Delta(A_0 \oplus x_n) = A\Delta(A \oplus kx_n)$, we have

$$|A_0 \Delta(A_0 \oplus x_n)| \leq 2\, |A| = 2\, |A_0|\, /k < |A_0|\, \varepsilon. \qquad \square$$

Theorem 11.5 (Measure Extension Theorem for $\mathbb{R}$)**.** *The Lebesgue measure has an extension to a translation invariant additive measure which is defined on all subsets of* $\mathbb{R}$.

Proof. As noted above, it suffices to extend the Lebesgue measure on $X = [0,1)$. By Corollary 11.3, we find *some* extension μ of the Lebesgue measure to all subsets of X. Let $* : \widehat{S} \to \widehat{{}^*S}$ be an X-enlargement. Let $c : \mathcal{P}(X) \to \mathbb{N} \cup \{\infty\}$ be the function which associates to each subset of X its number of elements. Exercise 27 implies ${}^*c(B) = {}^\#B$ for any internal $B \subseteq {}^*X$. Consider the binary relation

$$\varphi := \{(\underline{n}, \underline{x}, \underline{y}) \in (\mathbb{N} \times X) \times \mathcal{P}(X) \mid \text{“}\underline{y} \text{ is finite”} \wedge c(\underline{y}\Delta(\underline{y} \oplus \underline{x}))/c(\underline{y}) \leq 1/\underline{n}\}.$$

Proposition 11.4 implies that φ is concurrent on $\mathbb{N} \times X$. By Theorem 8.10, ${}^*\varphi$ is satisfied on ${}^\sigma(\mathbb{N} \times X)$. The standard definition principle for relations thus implies that we find some *-finite $B \in {}^*X$ such that

$$\frac{{}^\#(B\Delta(B \,{}^*\!\oplus {}^*x))}{{}^\#B} \approx 0 \qquad (x \in X).$$

We claim that

$$\mu_0(A) := \mathrm{st}\left(\frac{1}{{}^\#B} \sum_{x \in B} {}^*\mu({}^*A \,{}^*\!\oplus x)\right) \qquad (A \subseteq X)$$

is the desired measure. Note that μ_0 is defined, because $|{}^*\mu(A_0)| \leq 1$ for all $A_0 \in {}^*\mathcal{P}(X)$ implies

$$\sum_{x \in B} {}^*\mu({}^*A \,{}^*\!\oplus x) \leq {}^\#B.$$

Moreover, if $A_1, A_2 \subseteq X$ are disjoint, then ${}^*A_1 \,{}^*\!\oplus x$ and ${}^*A_2 \,{}^*\!\oplus x$ are disjoint, and the additivity of ${}^*\mu$ and st thus implies that μ is additive. If $A \subseteq X$ is Lebesgue measurable, we have $\mu(A \oplus x) = \mu(A)$ for any $x \in X$, and so

$${}^*(\mu_0(A)) \approx \frac{1}{{}^\#B} \sum_{x \in B} {}^*\mu({}^*A) = {}^*\mu({}^*A) = {}^*(\mu(A))$$

which implies that $\mu_0(A) = \mu(A)$.

To see that μ_0 is translation invariant, let $A \subseteq X$ and $y \in X$ be given. We have

$$\mu_0(A) - \mu_0(A \oplus y) = \mathrm{st}\left(\frac{1}{{}^\#B} \sum_{x \in B} ({}^*\mu({}^*A \,{}^*\!\oplus x) - {}^*\mu({}^*A \,{}^*\!\oplus {}^*y \,{}^*\!\oplus x))\right).$$

Observe now that the sum in the above formula may be written as

$$\sum_{x \in B} {}^*\mu({}^*A \,{}^*\!\oplus x) - \sum_{x \in B \oplus y} {}^*\mu({}^*A \,{}^*\!\oplus x)$$
$$= \sum_{x \in B \setminus (B \oplus y)} {}^*\mu({}^*A \,{}^*\!\oplus x) - \sum_{x \in (B \oplus y) \setminus B} {}^*\mu({}^*A \,{}^*\!\oplus x).$$

Since $0 \leq {}^*\mu({}^*A \,{}^*\!\oplus x) \leq 1$ and the total number of summands in these two sums is ${}^{\#}(B\Delta(B \oplus y))$, we obtain the estimate

$$|\mu_0(A) - \mu_0(A \oplus y)| \leq \mathrm{st}\left(\frac{1}{{}^{\#}B}{}^{\#}(B\Delta(B \,{}^*\!\oplus {}^*y))\right) = 0.$$

Thus, μ_0 is translation invariant. □

The previous proofs are essentially taken from [Wag86, p. 161] and [Hen72b].

The reader will have realized that for Proposition 11.4 we might have replaced X by any *commutative* group. Moreover, the proof of Theorem 11.5 holds even for any group X which satisfies Følner's condition: Any translation invariant finitely additive probability measure on such a group X may be extended to a translation invariant finitely additive probability measure defined on all subsets of X. Groups possessing such a measure are called *amenable*. In particular, our above proofs show that commutative groups and, more generally, groups satisfying Følner's condition, are amenable. Følner has proved that conversely amenable groups must satisfy Følner's condition. Readers more interested in amenable groups are referred to [Wag86]. We only mention that the group of isometries in $\mathbb{R}^n$ is amenable if and only if $n \leq 2$.

Chapter 6

Nonstandard Topology and Functional Analysis

§12 Topologies and Filters

12.1 Topological Spaces

Definition 12.1. Let X be some set, and $\mathscr{O}$ a system of subsets of X. Then $\mathscr{O}$ is called a *topology* (and the pair $(X, \mathscr{O})$ is called a *topological space*), if:

1. $\varnothing, X \in \mathscr{O}$.
2. $A, B \in \mathscr{O}$ implies $A \cap B \in \mathscr{O}$.
3. For any $\mathscr{A} \subseteq \mathscr{O}$ we have $\bigcup \mathscr{A} \in \mathscr{O}$.

The elements of $\mathscr{O}$ are called *open* subsets of X. The complements of open sets are called *closed*.

Often, we do not mention $\mathscr{O}$ explicitly, and simply say (not very precisely) that X is a topological space.

The most important example of topological spaces are metric spaces:

Definition 12.2. Let X be some set, and $d : X \times X \to [0, \infty)$. Then d is called a *pseudometric* on X, if:

1. $d(x, x) = 0$.
2. $d(x, y) = d(y, x)$ (symmetry).
3. $d(x, y) \le d(x, z) + d(z, y)$ (triangle inequality).

Moreover, d is called a *metric* on X if additionally $d(x, y) = 0$ implies $x = y$.

Each (pseudo)metric induces a topology $\mathscr{O}$ on X in a canonical way, namely $\mathscr{O}$ is the system of all sets $O \subseteq X$ with the property that for each $x \in O$ there is some $\varepsilon \in \mathbb{R}_+$ such that the open ball $B(x, \varepsilon) := \{y \in X : d(x, y) < \varepsilon\}$ is contained in O. If we speak of a *(pseudo)metric space*, we always mean that X is equipped with this topology.

Any topology allows us a definition of *neighborhoods* of a point:

Definition 12.3. If X is a topological space and $x \in X$, then $U \subseteq X$ is called a *neighborhood of* x if there is an open set $O \subseteq U$ with $x \in O$. The system $\mathscr{U}(x)$ of all neighborhoods of x is called the *neighborhood filter* of x.

For example, in a (pseudo)metric space X, a set is a neighborhood of x if and only if it contains some ball with positive radius and center x.

Proposition 12.4. *Any neighborhood filter* $\mathscr{U}(x)$ *is a filter.*

Proof. If $U \in \mathscr{U}(x)$, then $x \in O \subseteq U$ for some open set O. Hence $U \neq \varnothing$, and whenever $U \subseteq V \subseteq X$, we have $x \in O \subseteq V$ which implies $V \in \mathscr{U}$. Finally, if $U_1, U_2 \in \mathscr{U}(x)$, say $x \in O_1 \subseteq U_1$ and $x \in O_2 \subseteq U_2$ with open sets O_1, O_2, then $O := O_1 \cap O_2$ is open, and $x \in O \subseteq U_1 \subseteq U_2$, i.e. $U_1 \cap U_2 \in \mathscr{U}(x)$. □

The neighborhoods determine the topology:

Proposition 12.5. *A set* $U \subseteq X$ *is open if and only if it is a neighborhood for each of its elements.*

Proof. Let $\mathscr{A}$ denote the system of all subsets of U which are open. By definition, $\bigcup \mathscr{A} \subseteq U$. If U is a neighborhood for each of its elements, then $U \in \mathscr{A}$, and thus $U \subseteq \bigcup \mathscr{A}$. Hence, $U = \bigcup \mathscr{A}$ is open. The converse implication is trivial. □

12.2 Filters in Nonstandard Analysis

In view of Proposition 12.4, filters play an important role in the study of topologies. We first describe filters in more detail by nonstandard methods.

For the rest of this section, let X be an entity of the standard world $\widehat{S}$, and $* : \widehat{S} \to \widehat{{}^*S}$ be a $\mathcal{P}(X)$-enlargement. If $\mathscr{F}$ is a filter over X, we define its *monad* as the set

$$\operatorname{mon}(\mathscr{F}) := \bigcap{}^{\sigma} \mathscr{F} = \bigcap \{ {}^*F : F \in \mathscr{F} \}.$$

Theorem 12.6. *The monads of filters satisfy:*

1. $\operatorname{mon}(\mathscr{F}) \neq \varnothing$.
2. *If* $A \subseteq \operatorname{mon}(\mathscr{F})$ *is internal, then there is some* $B \in {}^*\mathscr{F}$ *with* $A \subseteq B \subseteq \operatorname{mon}(\mathscr{F})$.
3. *If* $A \subseteq {}^*X$ *is internal with* $\operatorname{mon}(\mathscr{F}) \subseteq A$*, then* $A \in {}^*\mathscr{F}$.
4. *If* $A \subseteq X$ *with* $\operatorname{mon}(\mathscr{F}) \subseteq {}^*A$*, then* $A \in \mathscr{F}$.
5. *Finer filters have smaller monads:* $\mathscr{F}_1 \supseteq \mathscr{F}_2$ *if and only if* $\operatorname{mon}(\mathscr{F}_1) \subseteq \operatorname{mon}(\mathscr{F}_2)$.

6. *The monad characterizes the filter:* $\mathscr{F}_1 = \mathscr{F}_2$ *if and only if* $\operatorname{mon}(\mathscr{F}_1) = \operatorname{mon}(\mathscr{F}_2)$.

Proof. 1. Since $\mathscr{F}$ has the finite intersection property, $\operatorname{mon}(\mathscr{F}) \neq \varnothing$ follows from the definition of enlargements.
2. By Theorem 8.10, there is a *-finite set $\mathscr{F}_0$ with ${}^\sigma\mathscr{F} \subseteq \mathscr{F}_0 \subseteq {}^*\mathscr{F}$. Then $\mathscr{F}_1 := \{F \in \mathscr{F}_0 : A \subseteq F\}$ is an internal subset of $\mathscr{F}_0$ by the internal definition principle. Theorem 6.13 thus implies that $\mathscr{F}_1$ is *-finite. Since $A \subseteq \operatorname{mon}(\mathscr{F})$, the definition of $\operatorname{mon}(\mathscr{F})$ implies ${}^\sigma\mathscr{F} \subseteq \mathscr{F}_1$ and thus also $B := \bigcap \mathscr{F}_1 \subseteq \operatorname{mon}(\mathscr{F})$. Since $\mathscr{F}$ is a filter, the sentence

$$\forall \underline{x} \in \mathcal{P}(\mathscr{F}) : (\text{“}\underline{x} \text{ is finite”} \implies \bigcap \underline{x} \in \mathscr{F})$$

is true. The transfer principle implies that for any *-finite internal subset x of ${}^*\mathscr{F}$, we have $\bigcap x \in \mathscr{F}$. In particular, $B = \bigcap \mathscr{F}_1$ is an element of ${}^*\mathscr{F}$.
3. By 2., we find some $B \in {}^*\mathscr{F}$ with $B \subseteq \operatorname{mon}(\mathscr{F})$; in particular $B \subseteq A$. The transfer of the sentence

$$\forall \underline{x} \in \mathcal{P}(X) : ((\exists \underline{y} \in \mathscr{F} : \underline{y} \subseteq \underline{x}) \implies \underline{x} \in \mathscr{F})$$

implies that ${}^*\mathscr{F}$ contains all internal subsets of ${}^*X$ which contain some element of ${}^*\mathscr{F}$ as a subset. In particular, $A \in {}^*\mathscr{F}$.
4. Since ${}^*A \subseteq {}^*X$ is internal with $\operatorname{mon}(\mathscr{F}) \subseteq {}^*A$, we find by 3. that ${}^*A \in {}^*\mathscr{F}$ which by the inverse form of the transfer principle implies $A \in \mathscr{F}$.
5. If $\mathscr{F}_1 \supseteq \mathscr{F}_2$, then trivially $\operatorname{mon}(\mathscr{F}_1) \subseteq \operatorname{mon}(\mathscr{F}_2)$. Conversely, suppose that $\operatorname{mon}(\mathscr{F}_1) \subseteq \operatorname{mon}(\mathscr{F}_2)$. For any $F \in \mathscr{F}_2$, the definition of the monad implies ${}^*F \supseteq \operatorname{mon}(\mathscr{F}_2) \supseteq \operatorname{mon}(\mathscr{F}_1)$, and so $F \in \mathscr{F}_1$ by 4.
6. Swapping the roles of $\mathscr{F}_1$ and $\mathscr{F}_2$ in 5., the statement follows. □

Exercise 60. Prove that a filter $\mathscr{U}$ on X is an ultrafilter if and only if for any filter $\mathscr{F}$ on X we have either $\operatorname{mon}(\mathscr{U}) \subseteq \operatorname{mon}(\mathscr{F})$ or $\operatorname{mon}(\mathscr{U}) \cap \operatorname{mon}(\mathscr{F}) = \varnothing$. Prove also that in this case, any different ultrafilter $\mathscr{U}'$ on X satisfies $\operatorname{mon}(\mathscr{U}) \cap \operatorname{mon}(\mathscr{U}') = \varnothing$.

Hint: Consider the system $\mathscr{F}_0 := \{F \cap U : F \in \mathscr{F},\ U \in \mathscr{U}\}$.

Theorem 12.7. *If* $*$ *is even a* compact $\mathcal{P}(X)$*-enlargement, then the following holds:*

If $\mathscr{F}$ *is a filter on* X*, then any internal set* $\mathscr{B} \supseteq \{F \in {}^*\mathscr{F} : F \subseteq \operatorname{mon}(\mathscr{F})\}$ *contains some standard element* $F \in {}^\sigma\mathscr{F}$.

Proof. Assume that $\mathscr{B}$ contains no standard element. Consider the internal binary relation

$$\varphi := \{(\underline{x}, \underline{y}) \in {}^*\mathscr{F} \times {}^*\mathscr{F} \mid \underline{y} \subseteq \underline{x} \wedge \underline{y} \notin \mathscr{B}\}.$$

Then φ is concurrent on ${}^\sigma\mathscr{F}$: Indeed, if $x_1 = {}^*F_1, \dots, x_n = {}^*F_n$ with $F_k \in \mathscr{F}$ are given, put $y := {}^*(F_1 \cap \dots \cap F_n) = {}^*F_1 \cap \dots \cap {}^*F_n$. Since y is a standard element our assumption implies $y \notin \mathscr{B}$. Since $y \in {}^*\mathscr{F}$ satisfies $y \subseteq x_k$, we thus have $(x_1, y), \dots, (x_n, y) \in \varphi$, as desired. Since $*$ is a compact $\mathcal{P}(X)$-enlargement, φ is satisfied on ${}^\sigma\mathscr{F}$, i.e. there is some $F_0 \in {}^*\mathscr{F}$ such that $({}^*F, F_0) \in \varphi$ for each $F \in \mathscr{F}$. But then $F_0 \subseteq \mathrm{mon}(\mathscr{F})$ and $F_0 \notin \mathscr{B}$, a contradiction to the definition of $\mathscr{B}$. □

Definition 12.8. A filter $\mathscr{F}$ over X is called *principal* if there is some $A \subseteq X$ such that

$$\mathscr{F} = \{F \subseteq X : A \subseteq F\};$$

otherwise $\mathscr{F}$ is called *nonprincipal*.

Example 12.9. A filter $\mathscr{F}$ is principal if and only if $\bigcap \mathscr{F} \in \mathscr{F}$ (consider $A := \bigcap \mathscr{F}$ to see this). An ultrafilter $\mathscr{U}$ is nonprincipal if and only if it is free (Exercise 10).

Theorem 12.10 (Luxemburg)**.** *If the filter $\mathscr{F}$ is principal, then* $\mathrm{mon}(\mathscr{F}) = {}^*(\bigcap \mathscr{F})$ *is a standard entity. Otherwise,* $\mathrm{mon}(\mathscr{F})$ *is external.*

Proof. (For the case that $*$ is a *compact* $\mathcal{P}(X)$-enlargement):

The first statement is almost trivial: Let $\mathscr{F}$ be principal, i.e. $\mathscr{F} = \{F \subseteq X : A \subseteq F\}$ where $A := \bigcap \mathscr{F}$. The transfer principle implies ${}^*A \subseteq {}^*F$ for any $F \in \mathscr{F}$, and so ${}^*A \subseteq \mathrm{mon}(\mathscr{F})$. Since $A \in \mathscr{F}$, we have $\mathrm{mon}(\mathscr{F}) \subseteq {}^*A$, and so $\mathrm{mon}(\mathscr{F}) = {}^*A$ is a standard entity.

Conversely, suppose that $\mathrm{mon}(\mathscr{F})$ is internal. Then the set $\mathscr{B} := \{A \in {}^*\mathscr{F} : A \subseteq \mathrm{mon}(\mathscr{F})\}$ is internal. Since we assume that $*$ is a compact enlargement, Theorem 12.7 implies that $\mathscr{B}$ contains some standard element, i.e. there is some $A \in \mathscr{F}$ with ${}^*A \subseteq \mathrm{mon}(\mathscr{F})$, i.e. ${}^*A \subseteq {}^*F$ for each $F \in \mathscr{F}$. This means $A \subseteq F$ for each $F \in \mathscr{F}$ (Lemma 3.5), and so $\bigcap \mathscr{F} = A \in \mathscr{F}$ which means that $\mathscr{F}$ is principal. □

We intend to give a proof for Theorem 12.10 also for the case that $*$ is only an enlargement (not necessarily a compact enlargement). The proof is rather deep and is taken from [Lux69a]. We need some preparation:

A system $\mathscr{B}$ of subsets of X is called a *subbase* of the filter $\mathscr{F}$, if $\mathscr{F}$ is the filter generated by $\mathscr{B}$. The monad of a filter is determined by any subbase:

Lemma 12.11. *If the filter $\mathscr{F}$ is generated by $\mathscr{B}$, then*

$$\mathrm{mon}(\mathscr{F}) = \bigcap{}^\sigma\mathscr{B}.$$

Proof. Since ${}^\sigma\mathscr{B} \subseteq {}^\sigma\mathscr{F}$, we have $\bigcap{}^\sigma\mathscr{B} \supseteq \bigcap{}^\sigma\mathscr{F} = \mathrm{mon}(\mathscr{F})$. For the converse inclusion, let $A \in \bigcap{}^\sigma\mathscr{B}$, i.e. $A \subseteq {}^*B$ for any $B \in \mathscr{B}$. If $F \in \mathscr{B}$ is given, we have $F \supseteq B_1 \cap \dots \cap B_n$ for $B_1, \dots, B_n \in \mathscr{B}$, and so ${}^*F \supseteq {}^*B_1 \cap \dots \cap {}^*B_n \supseteq A$. Thus, $A \subseteq {}^*F$ for any $F \in \mathscr{F}$, i.e. $A \in \mathrm{mon}(\mathscr{F})$. □

The smallest cardinality of all subbases of $\mathscr{F}$ is called the *dimension* of $\mathscr{F}$.

Lemma 12.12. *If the filter $\mathscr{F}$ is principal, its dimension is 1. Otherwise its dimension is infinite.*

Proof. If $\mathscr{F}$ is principal, then the single set $\bigcap \mathscr{F}$ constitutes a subbase for $\mathscr{F}$. If $\mathscr{F}$ has finite dimension, then there is a finite subbase $\mathscr{B} = \{B_1, \dots, B_n\}$ for $\mathscr{F}$. Putting $A := B_1 \cap \cdots \cap B_n$, we have $F \in \mathscr{F}$ if and only if $A \subseteq F \subseteq X$, i.e. $\mathscr{F}$ is principal. □

One might hope to obtain a "minimal subbase" for $\mathscr{F}$ by considering a subbase of $\mathscr{F}$ with smallest possible cardinality and then to generate from this subbase a "minimal subbase" by successively choosing only those elements which are really "needed" to generate $\mathscr{F}$. Of course, if the subbase is uncountable, one has to use a transfinite induction for this procedure. This idea leads to a minimal subbase in the following sense:

Lemma 12.13. *Let $\mathscr{F}$ be a filter of infinite dimension κ. Then there exists a subbase $\mathscr{B}$ for $\mathscr{F}$ with the following property:*

There is an injection $i : \mathscr{B} \to \kappa$ such that whenever $\mathscr{E} \subseteq \mathscr{B}$ is finite and $\bigcap \mathscr{E} \subseteq E \in \mathscr{B}$, then $i(E) \leq \max\{i(F) : F \in \mathscr{E}\}$ (with respect to the order in the ordinal number κ).

Proof. Let $\mathscr{B}_0$ be a subbase with cardinality κ, and let F_α $(0 \leq \alpha < \kappa)$ be a corresponding enumeration of the elements of $\mathscr{B}_0$. Now we put $\mathscr{C}_\alpha := \{F_\alpha : 0 \leq \beta < \alpha\}$, and for $0 < \alpha < \kappa$, let $\mathscr{F}_\alpha$ denote the filter generated by $\mathscr{C}_\alpha$. Now we define $\mathscr{B}$ as follows:

$$\mathscr{B} := \{F_\alpha \mid \alpha = 0 \text{ or } F_\alpha \notin \mathscr{F}_\alpha\}.$$

The mapping i is defined by $i(F_\alpha) := \alpha$.

We prove by transfinite induction on α that the filter generated by $\mathscr{A}_\alpha := \{F_\beta : 0 \leq \beta \leq \alpha\} = \mathscr{C}_\alpha \cup \{F_\alpha\}$ is a subset of the filter $\mathscr{B}_{\mathscr{F}}$ generated by $\mathscr{B}$: Indeed, this is true for $\alpha = 0$. Suppose as induction assumption that this is true for all $\alpha < \alpha_0$, i.e. $\mathscr{C}_{\alpha_0} \subseteq \mathscr{B}_{\mathscr{F}}$. By definition of $\mathscr{B}$, we have either $F_{\alpha_0} \in \mathscr{B}$, or $F_{\alpha_0} \in \mathscr{F}_{\alpha_0}$. Since the induction assumption implies $\mathscr{F}_{\alpha_0} \subseteq \mathscr{B}_{\mathscr{F}}$, we have in both cases $F_{\alpha_0} \in \mathscr{B}_{\mathscr{F}}$, and so $\mathscr{A}_{\alpha_0} = \mathscr{C}_{\alpha_0} \cup \{F_{\alpha_0}\} \subseteq \mathscr{B}_{\mathscr{F}}$, as required.

We thus have proved, in particular, that each F_α is contained in the filter $\mathscr{B}_{\mathscr{F}}$ generated by $\mathscr{B}$. Since the system of all F_α generates $\mathscr{F}$, we have $\mathscr{F} \subseteq \mathscr{B}_{\mathscr{F}}$. But in view of $\mathscr{B} \subseteq \mathscr{B}_0$, we have $\mathscr{B}_{\mathscr{F}} \subseteq \mathscr{F}$, and so $\mathscr{B}_{\mathscr{F}} = \mathscr{F}$, i.e. $\mathscr{B}$ actually generates the filter $\mathscr{F}$.

The other property follows from our construction: Let $\mathscr{E} \subseteq \mathscr{B}$ be finite, say $\mathscr{E} = \{F_{\alpha_1}, \dots, F_{\alpha_n}\}$, and $\bigcap \mathscr{E} \subseteq E \in \mathscr{B}$. Assume by contradiction that $\alpha := i(E) > \{i(F) : F \in \mathscr{E}\} = \max\{\alpha_1, \dots, \alpha_n\}$. Then $E = F_\alpha$ belongs to the filter $\mathscr{F}_\alpha$ generated by $\mathscr{C}_\alpha \supseteq \mathscr{E}$ which contradicts our definition of $\mathscr{B}$. □

Proof of Theorem 12.10. The first statement has already been proved (the compactness of $*$ was not needed in our previous proof concerning the first statement).

For the second statement, let $\mathscr{F}$ be nonprincipal. Then $\mathscr{F}$ has infinite dimension κ by Lemma 12.12. Without loss of generality, we assume that κ is an entity of $\widehat{S}$. Choose $\mathscr{B}$ and i as in Lemma 12.13, and put

$$\mathscr{P} := \{\underline{x} \in {}^*\mathcal{P}(\mathscr{B}) \mid \text{“}\underline{x} \text{ is } {}^*\text{-finite”} \wedge \bigcap \underline{x} \subseteq \operatorname{mon}(\mathscr{F})\}.$$

If we assume by contradiction that $\operatorname{mon}(\mathscr{F})$ is internal, then $\mathscr{P}$ is internal by the internal definition principle. By Theorem 8.10, there is some *-finite set $\mathscr{B}_0 \subseteq {}^*\mathscr{B}$ with $\mathscr{B}_0 \supseteq {}^\sigma\mathscr{B}$. In particular, $\bigcap \mathscr{B}_0 \subseteq \bigcap {}^\sigma\mathscr{B} = \operatorname{mon}(\mathscr{F})$ (Lemma 12.11). Hence, $\mathscr{B}_0 \in \mathscr{P}$, and so $\mathscr{P} \neq \varnothing$.

Note that $\mathscr{P}$ consists of (internal) *-finite entities $\mathscr{A} \subseteq {}^*\mathscr{B}$. Hence, we may define an internal function $j : \mathscr{P} \to {}^*\kappa$ by $j(\mathscr{A}) := \max\{{}^*i(F) : F \in \mathscr{A}\}$ (Exercise 26). In particular, $\operatorname{rng}(j)$ is an internal nonempty subset of $^*\kappa$ (Theorem 3.19). The transfer of the sentence “κ is well-ordered” implies that $\operatorname{rng}(j)$ has a smallest element $\alpha_0 \in {}^*\kappa$.

We show first that we have for any standard element $A \in {}^\sigma\mathscr{B}$ the relation ${}^*i(A) \leq \alpha_0$: Indeed, by $\mathscr{B}_0 \supseteq {}^\sigma\mathscr{B}$, we have $A \in \mathscr{B}_0 \in \mathscr{P}$, and so the definition of j and α_0 imply ${}^*i(A) \leq j(\mathscr{B}_0) \leq \alpha_0$.

It follows that α_0 is not a standard number: Indeed, assume to the contrary that $\alpha_0 = {}^*\beta_0$ for some $\beta_0 \in \kappa$. Then we would have for any $A \in \mathscr{B}$ that ${}^*(i(A)) = {}^*i({}^*A) \leq \alpha_0 = {}^*\beta_0$, i.e. $i(A) \leq \beta_0$. Thus, i is an injection from $\mathscr{B}$ into the set $\{\beta \in \kappa : \beta \leq \beta_0\}$ which has a strictly smaller cardinality than κ, a contradiction to the fact that $\mathscr{B}$ has the cardinality of κ (because κ is the dimension of $\mathscr{F}$). Hence α_0 is not a standard number.

We thus have for any standard $A \in {}^\sigma\mathscr{B}$ that ${}^*i(A) \neq \alpha_0$; since we already proved ${}^*i(A) \leq \alpha_0$, we even have ${}^*i(A) < \alpha_0$. Consider now the set $\mathscr{B}_1 := \{\underline{x} \in \mathscr{B}_0 : {}^*i(\underline{x}) < \alpha_0\}$. By the internal definition principle, the set $\mathscr{B}_1$ is an internal subset of the *-finite set $\mathscr{B}_0$. Hence, $\mathscr{B}_1$ is *-finite (Theorem 6.13). Moreover, since $\mathscr{B}_0 \supseteq {}^\sigma\mathscr{B}$ and since any $A \in {}^\sigma\mathscr{B}$ satisfies ${}^*i(A) < \alpha_0$ (as we have proved above), we have ${}^\sigma\mathscr{B} \subseteq \mathscr{B}_1$. Consequently, $\bigcap \mathscr{B}_1 \subseteq \bigcap {}^\sigma\mathscr{B} = \operatorname{mon}(\mathscr{F})$ (Lemma 12.11). Thus, $\mathscr{B}_1 \in \mathscr{P}$. Since $j(\mathscr{B}_1) < \alpha_0$ (by definition of j and $\mathscr{B}_1$ and the transfer principle), this contradicts the definition of α_0. □

12.3 Topologies in Nonstandard Analysis

Now we return to the study of topological spaces. As before, let X be a topological space where $X \in \widehat{S}$ is an entity, and $* : \widehat{S} \to \widehat{{}^*S}$ is a $\mathcal{P}(X)$-enlargement.

For $x \in X$, let $\mathscr{U}(x)$ denote its neighborhood filter. In practice, one usually does not calculate with $\mathscr{U}(x)$ but only with a neighborhood base:

Definition 12.14. A system $\mathscr{B} \subseteq \mathscr{U}(x)$ is called a *neighborhood base* for x, if for each neighborhood U of x there is some $B \in \mathscr{B}$ with $B \subseteq U$.

Example 12.15. The system of all open neighborhoods of x is a neighborhood base for x. Indeed, if $U \in \mathscr{U}(x)$, there is some open $O \subseteq X$ with $x \in O \subseteq U$, i.e. O is an open neighborhood of x with $O \subseteq U$.

Example 12.16. If X is a metric space with metric d, then a neighborhood base for x is given by the system of open balls $B(x,r) := \{y \in X : d(x,y) < r\}$ $(r \in \mathbb{R}_+)$ or also by the system of all closed balls $\overline{B}(x,r) := \{y \in X : d(x,y) \leq r\}$ $(r \in \mathbb{R}_+)$. Another neighborhood base is given by the system $B(x,1/n)$ $(n \in \mathbb{N})$ or by $\overline{B}(x,1/n)$ $(n \in \mathbb{N})$.

Example 12.17. Let $U \in \mathscr{U}(x)$. Then the system $\{V \in \mathscr{U}(x) : V \subseteq U\}$ is a neighborhood base for x (since for $W \in \mathscr{U}(x)$ the set $V := W \cap U$ is a neighborhood for x). Moreover, the system of all open neighborhoods O of x with $O \subseteq U$ is also a neighborhood base for x.

Definition 12.18. The *monad of* x, $\operatorname{mon}(x)$, is the monad of its neighborhood filter $\mathscr{U}(x)$.

Proposition 12.19. *If x is a point in a topological space X, we have*

$$\operatorname{mon}(x) = \bigcap{}^{\sigma}\mathscr{B} = \bigcap\{{}^*U : U \in \mathscr{B}\}$$

for any neighborhood base $\mathscr{B}$ of x. In particular, $y \in \operatorname{mon}(x)$ if and only if $y \in {}^*O$ for each open $O \subseteq X$ with $x \in O$. Moreover, there is some set $B \in {}^*\mathscr{B}$ with

$$B \subseteq \operatorname{mon}(x).$$

Proof. It follows immediately from the definition that the neighborhood filter $\mathscr{U}(x)$ of x is generated by any neighborhood base. Hence, the first statement follows from Lemma 12.11. For the last statement, observe that Theorem 12.6 2. implies that there is some some $F \in {}^*\mathscr{F}$ with $F \subseteq \operatorname{mon}(x)$. The transfer of the statement

$$\forall \underline{x} \in \mathscr{F} : \exists \underline{y} \in \mathscr{B} : \underline{y} \subseteq \underline{x}$$

implies that we find some $B \in {}^*\mathscr{B}$ with $B \subseteq F \subseteq \operatorname{mon}(x)$. □

Corollary 12.20. *If X is a pseudometric space, we have*

$$\operatorname{mon}(x) = \bigcap_{r \in \mathbb{R}_+} {}^*B(x,r) = \bigcap_{n \in \mathbb{N}} {}^*B(x,1/n),$$

where $B(x,r) := \{y \in X : d(x,y) < r\}$. In particular, for $X = \mathbb{R}$ with the natural topology, the definition of $\operatorname{mon}(x)$ *coincides with our previous Definition* 5.22.

We also want to define monads for points in the nonstandard world; we define even monads of sets:

Definition 12.21. If $A \subseteq {}^*X$ is nonempty, we denote the filter generated by the system of all open sets $O \subseteq X$ with $A \subseteq {}^*O$ the *standard filter* of A. Its monad is called the *monad of* A and denoted by $\mu(A)$.

Note that the system $\mathscr{B}$ of all open sets $O \subseteq X$ with $A \subseteq {}^*O$ indeed generates a filter, since it has the finite intersection property: If $O_1, \ldots, O_n \in \mathscr{B}$, then $O = O_1 \cap \cdots \cap O_n$ is open, and ${}^*O = {}^*O_1 \cap \cdots \cap {}^*O_n$ contains A and thus is nonempty. Hence, $O \neq \varnothing$. By Lemma 12.11, we have $\mu(A) = \bigcap^\sigma \mathscr{B}$. In other words:

Proposition 12.22. *We have $y \in \mu(A)$ if and only if $y \in {}^*O$ for each open set $O \subseteq X$ with $A \subseteq {}^*O$.* □

Corollary 12.23. *If $x \in X$, then $\mu(\{{}^*x\}) = \mathrm{mon}(x)$.*

Proof. By the transfer principle, $\{{}^*x\} \subseteq {}^*O$ if and only if $x \in O$. Now apply Propositions 12.22 and 12.19. □

Definition 12.24. We call two nonstandard points $x, y \in {}^*X$ of a standard topological space X *infinitely close* to each other if for each open set $O \subseteq X$ with $x \in {}^*O$ we also have $y \in {}^*O$, i.e. if $y \in \mu(\{x\})$. In this case, we write $y \approx_{\mathscr{O}} x$.

In the case $X = \mathbb{R}$, we have $\mathrm{mon}(x) = \{y \in X : y \approx {}^*x\} = \{y \in X : y \approx_{\mathscr{O}} {}^*x\}$. Indeed, Corollary 12.23 implies for any topological space X:

Corollary 12.25. *We have for any $x \in X$ that*

$$\mathrm{mon}(x) = \{y \in {}^*X : y \approx_{\mathscr{O}} {}^*x\}.$$ □

We emphasize that $\approx_{\mathscr{O}}$ is in literature sometimes only defined when x is a standard point (e.g. in [LR94]).

The reader should be warned that $\approx_{\mathscr{O}}$ is in general *not* an equivalence relation and not even symmetric, i.e. $x \approx_{\mathscr{O}} y$ does not imply that $y \approx_{\mathscr{O}} x$, even if x and y are both standard points:

Example 12.26. Let $X = \{a, b\}$ $(a \neq b)$ where only the three sets $\varnothing, \{b\}, \{a, b\}$ are open (this is a topology!). Since X is finite, we have ${}^\sigma X = {}^*X$. Similarly, ${}^*O = {}^\sigma O$ for each open set $O \subseteq X$. The only set O with ${}^*a \in {}^*O$ is thus $O = \{a, b\}$, and so ${}^*b \approx_{\mathscr{O}} {}^*a$. However, for $O = \{b\}$, we have ${}^*b \in {}^*O$ and ${}^*a \notin {}^*O$ which implies ${}^*a \not\approx_{\mathscr{O}} {}^*b$.

There is another danger when dealing with $\approx_{\mathscr{O}}$ for nonstandard points:

For $X = \mathbb{N}$ with the natural (metric) topology, one might suspect that $n \approx_{\mathscr{O}} m$ only for $n = m$. This is indeed true if either n or m is finite, but fails for nonstandard points:

Theorem 12.27. *Let $X = \mathbb{N}$ and $h \in \mathbb{N}_\infty$. Then there are infinitely many $n \in {}^*\mathbb{N}$ with $n \approx_{\mathscr{O}} h$. Moreover, the relation $n \approx_{\mathscr{O}} h$ implies that either $n = h$ or that $|n - h|$ is infinite.*

Proof. Recall that $\mu(\{h\})$ is the filter monad of the standard filter of $\{h\}$. This filter is nonprincipal, since for any $n \in \mathbb{N}$ the set $\mathbb{N} \setminus \{n\}$ belongs to the filter (because $h \in {}^*\mathbb{N} \setminus \{{}^*n\} = {}^*(\mathbb{N} \setminus \{n\})$). Consequently, $\mu(\{h\}) \subseteq {}^*\mathbb{N}$ is external by Theorem 12.10. Hence, $\mu(\{h\})$ is infinite by Exercise 6.

Assume that $k := |n - h| > 0$ is finite, i.e. $k \in {}^\sigma\mathbb{N}$. Put $F_j := \{2ki + j : i \in \mathbb{N}\}$. Then $F_0 \cup \cdots \cup F_{2k-1} = \mathbb{N}$, and so ${}^*\mathbb{N} = {}^*(F_0) \cup \cdots \cup {}^*(F_{2k-1})$, i.e. we find some $j \in \mathbb{N}$ with $h \in {}^*F_j$. Then F_j belongs to the standard filter of $\{h\}$, and so $n \in {}^*F_j$. By the standard definition principle, we have ${}^*F_j = \{2ki + j : i \in {}^*\mathbb{N}\}$. In view of $h, n \in {}^*F_j$, we thus find that $n - h = (n + j) - (h + j)$ is a multiple of $2k$, a contradiction to $k = |n - h| > 0$. □

Corollary 12.28. *On $X = \mathbb{R}$ (with the natural topology) the relation $y \approx x$ is not equivalent to $y \approx_{\mathscr{O}} x$.* □

However, we will see that

$$y \approx {}^*x \iff y \approx_{\mathscr{O}} {}^*x.$$

For the above reasons, $\approx_{\mathscr{O}}$ may not appear a "natural" notion. For so-called uniform spaces (like $X = \mathbb{R}$) we will later learn another relation which is more natural and which for standard points coincides with $\approx_{\mathscr{O}}$; for $X = \mathbb{R}$ this new relation becomes the same as $\approx$.

One of the most useful concepts in real nonstandard analysis was the mapping st. It appears natural to call $x \in X$ the standard part of $y \in {}^*X$ if $y \approx_{\mathscr{O}} {}^*x$, i.e. $\operatorname{mon}(x)$ consists precisely of all those points $y \in {}^*X$ whose standard part is x. Recall that in case $X = \mathbb{R}$, the standard part mapping was not defined on ${}^*R$ but only on $\operatorname{fin}({}^*\mathbb{R})$. Hence, we cannot expect to define st on all of X. Even worse, in general it may happen that $\operatorname{st}(y)$ is not uniquely determined. Nevertheless, we can define:

Definition 12.29. The *standard part relation* st is a relation on ${}^*X \times X$, defined by

$$(y, x) \in \operatorname{st} \iff y \approx_{\mathscr{O}} {}^*x \qquad (\iff y \in \operatorname{mon}(x)).$$

Points $y \in \operatorname{dom}(\operatorname{st})$ are called *nearstandard*. The set of all nearstandard points of X is denoted by $\operatorname{ns}(X)$.

Of course, in case $X = \mathbb{R}$, the relation st is a function, and we end up with the old Definition 5.20. Recall that a topological space is called a *Hausdorff space* if each two points $x \neq y$ have disjoint neighborhoods. For example, a pseudometric space is a Hausdorff space if and only if it is a metric space (indeed, if two points $x \neq y$ satisfy $d(x, y) = 0$, they have the same neighborhoods).

Proposition 12.30. *For a topological space X the following three statements are equivalent:*

1. *X is a Hausdorff space.*
2. *The relation* st *is a function.*
3. *Monads to different points are disjoint, i.e. $x \neq y$ implies* $\mathrm{mon}(x) \cap \mathrm{mon}(y) = \varnothing$.

Proof. The equivalence of the last two statements follows immediately from the definition. If $x \neq y$ have disjoint monads, choose $U \in {}^*\mathscr{U}(x)$, $V \in {}^*\mathscr{U}(y)$ with $U \subseteq \mathrm{mon}(x)$ and $V \subseteq \mathrm{mon}(y)$ (Proposition 12.19). Then the sentence

$$\exists \underline{u} \in {}^*\mathscr{U}(x), \underline{v} \in {}^*\mathscr{U}(y) : \underline{u} \cap \underline{v} = \varnothing$$

is true, and the inverse form of the transfer principle implies that x and y have disjoint neighborhoods. Conversely, if $x \neq y$ have disjoint neighborhoods U and V, respectively, then $U \cap V = \varnothing$ implies ${}^*U \cap {}^*V = \varnothing$, and since $\mathrm{mon}(x) \subseteq {}^*U$, $\mathrm{mon}(y) \subseteq {}^*V$, it follows that $\mathrm{mon}(x) \cap \mathrm{mon}(y) = \varnothing$. □

For compact enlargements we have the following generalization of the Cauchy principle:

Theorem 12.31 (Permanence principle for ${}^*X$ (Cauchy principle))**.** *Let $*$ be a compact $\mathcal{P}(X)$-enlargement. Let $\alpha(\underline{y})$ be an internal predicate with $\underline{y}$ as its only free variable. If $\alpha(y)$ holds for all $y \in \mathrm{mon}(x)$, then there is some standard neighborhood U of x such that $\alpha(y)$ holds for all $y \in {}^*U$.*

Proof. The set

$$\mathscr{B} := \{\underline{u} \in {}^*\mathscr{U}(x) \mid \forall \underline{y} \in \underline{u} : \alpha(\underline{y})\}$$

is internal by the internal definition principle. Moreover, any $F \in {}^*\mathscr{U}(x)$ with $F \subseteq \mathrm{mon}(x)$ belongs to $\mathscr{B}$. Hence, Theorem 12.7 implies that $\mathscr{B}$ contains some standard element $V = {}^*U$ with $U \in \mathscr{U}(x)$. □

As one might expect, monads can be used to characterize topological sets:

Definition 12.32. Let $A \subseteq X$. A point $x \in A$ is called an *interior point* of A, if A is a neighborhood of x. The *closure* $\overline{A}$ is the set of all points $x \in X$ with the property that any neighborhood of x intersects A.

We first recall some facts of the standard world:

Proposition 12.33. *We have:*

1. *The set $\overline{A}$ is the smallest closed set which contains A. In particular, A is closed if and only if $A = \overline{A}$.*

2. *The set of all interior points of A is the smallest open set which is contained in A. In particular, A is open if and only if each $x \in A$ is interior, i.e. if and only if A is a neighborhood for each of its elements.*

Proof. 1. We claim that $C := X \setminus \overline{A}$ is the union of all open sets which are contained in $X \setminus A$: Then C is open, and thus $\overline{A}$ is closed, and moreover, if $B \supseteq A$ is closed, then $X \setminus B \subseteq C$, i.e. $\overline{A} \subseteq B$.

If $O \subseteq X \setminus A$ is open and $x \in O$, then O is a neighborhood of x which does not intersect A, and so $x \in C$ by definition of $\overline{A}$. Conversely, if $x \in C$, then x has a neighborhood U which does not intersect A, and so there is some open $O \subseteq U$ with $x \in O$. We have $O \subseteq C$ by definition of $\overline{A}$.
2. We prove that the set I of interior points of A is the union of all open sets contained in A: If $O \subseteq A$ is open and $x \in O$, then $x \in I$. Conversely, if $x \in I$, then there is some open $O \subseteq A$ with $x \in O$; hence, x is contained in an open set $O \subseteq A$. □

Now we turn to the nonstandard characterizations for topological properties of sets:

Theorem 12.34. *Let $A \subseteq X$. A point $x \in A$ is an interior point of A if and only if* $\operatorname{mon}(x) \subseteq {}^*A$. *The set A is open if and only if*

$$\bigcup_{x\in A} \operatorname{mon}(x) \subseteq {}^*A.$$

Proof. The second statement follows from the first by Proposition 12.33. Let $x \in A$ be interior, i.e. $A \in \mathscr{U}(x)$. The system $\mathscr{B} := \{V \in \mathscr{U}(x) : V \subseteq A\}$ is a neighborhood base for x such that any $V \in \mathscr{B}$ satisfies $V \subseteq A$, i.e. ${}^*V \subseteq {}^*A$. Hence, Proposition 12.19 implies

$$\operatorname{mon}(x) = \bigcap{}^{\sigma}\mathscr{B} = \{{}^*V : V \in \mathscr{B}\} \subseteq {}^*A.$$

Conversely, suppose that $\operatorname{mon}(x) \subseteq {}^*A$. By Proposition 12.19, we find some $U \in {}^*\mathscr{U}(x)$ with ${}^*U \subseteq \operatorname{mon}(x)$. Hence,

$$\exists \underline{u} \in {}^*\mathscr{U}(x) : \underline{u} \subseteq {}^*A.$$

The inverse form of the transfer principle implies that $\mathscr{U}(x)$ contains a subset of A, i.e. x is an interior point of A. □

For the second statement, we could also have applied the Cauchy principle, but this would require that $*$ be a *compact* enlargement.

We think of st as a multivalued function, and thus use for $y \in {}^*X$ the notation

$$\operatorname{st}(y) := \{x \in X : (y, x) \in \operatorname{st}\} = \{x \in X : y \approx_{\mathscr{O}} {}^*x\}$$

which is slightly ambiguous if st is single-valued, because in this case st is a function, i.e. $\mathrm{st}(y)$ was already defined as the unique point $x \in X$ with $y \approx_{\mathscr{O}} {}^*x$ (and not as the *set* $\{x\}$). However, we hope that no confusion will arise. We emphasize that by the above definition, we have $\mathrm{st}(y) = \varnothing$ if (and only if) $y \notin \mathrm{dom}(\mathrm{st}) = \mathrm{ns}({}^*X)$.

We also use for $A \subseteq {}^*X$ the corresponding notation

$$\mathrm{st}(A) := \bigcup_{y \in A} \mathrm{st}(y) = \{x \in X : \text{There is some } y \in A \text{ with } y \approx_{\mathscr{O}} {}^*x\}.$$

Moreover, for the inverse relation $\mathrm{st}^{-1} := \{(x, y) : (y, x) \in \mathrm{st}\}$ we write correspondingly for $x \in X$,

$$\mathrm{st}^{-1}(x) := \{y : (x, y) \in \mathrm{st}^{-1}\} = \{y \in {}^*X : y \approx_{\mathscr{O}} {}^*x\}$$

and for $A \subseteq X$,

$$\mathrm{st}^{-1}(A) := \bigcup_{x \in A} \mathrm{st}(x) = \{y \in {}^*X : \text{There is some } x \in A \text{ with } y \approx_{\mathscr{O}} {}^*x\}.$$

Theorem 12.35. *Let $A \subseteq X$. Then*

$$\mathrm{st}({}^*A) = \overline{A},$$

i.e. a point $x \in X$ belongs to $\overline{A}$ if and only if there is some $y \in {}^*A$ with $y \approx_{\mathscr{O}} {}^*x$. In particular, A is closed if and only if

$$\mathrm{st}({}^*A) \subseteq A$$

(and then equality holds, because the converse inclusion is always true).

Proof. Let $x \in \overline{A}$, i.e. for any $U \in \mathscr{U}(x)$, we have $U \cap A \neq \varnothing$. Then the system $\mathscr{A} := \{U \cap A : U \in \mathscr{U}(x)\}$ has the finite intersection property. Since $*$ is a $\mathcal{P}(X)$-enlargement, this implies $\bigcap{}^\sigma \mathscr{A} \neq \varnothing$, i.e. there is some y with $y \in {}^*(U \cap A) = {}^*U \cap {}^*A$ for all $U \in \mathscr{U}(x)$. Hence, $y \in \mathrm{mon}(x) \cap {}^*A$, i.e. $y \in {}^*A$ satisfies $y \approx_{\mathscr{O}} {}^*x$, and so $x \in \mathrm{st}({}^*A)$.

Conversely, let $x \in \mathrm{st}({}^*A)$, i.e. there is some $y \in {}^*A$ with $y \approx_{\mathscr{O}} {}^*x$, i.e. $y \in \mathrm{mon}(x)$. For any $U \in \mathscr{U}(x)$, we thus have $y \in {}^*U$, and so ${}^*U \cap {}^*A \neq \varnothing$ which implies $U \cap A \neq \varnothing$ by the transfer principle. Consequently, $x \in \overline{A}$. $\square$

Corollary 12.36. *If X is a Hausdorff space, then $A \subseteq X$ is closed if and only if each $y \in {}^*A \cap \mathrm{ns}({}^*X)$ is infinitely close to some (standard) point of ${}^\sigma A$, i.e. if and only if*

$${}^*A \cap \mathrm{ns}({}^*X) \subseteq \mathrm{st}^{-1}(A).$$

Proof. Since X is Hausdorff, any $y \in {}^*A$ is infinitely close to at most one point ${}^*x$ with $x \in X$ (Proposition 12.30). Theorem 12.35 implies that A is closed if and only if the points x arising in this way all belong to A. The latter means that for any $y \in {}^*A$ with the additional property that y is infinitely close to some standard point x (i.e. $y \in \mathrm{ns}({}^*X)$), we have $x \in {}^\sigma A$. □

For compact enlargements, we have a generalization of the previous result for internal (not necessarily standard) sets:

Theorem 12.37. *Let* $*$ *be a compact* $\mathcal{P}(X)$*-enlargement. If* $A \subseteq {}^*X$ *is internal, then* $\mathrm{st}(A)$ *is closed.*

Proof. Let $x \in \overline{\mathrm{st}(A)}$, and let $\mathscr{B}$ be the system of all open neighborhoods of x. Consider the internal relation

$$\varphi := \{(\underline{x}, \underline{y}) \in {}^*\mathscr{B} \times A \mid \underline{y} \in \underline{x}\}.$$

We claim that φ is concurrent on ${}^\sigma\mathscr{B}$. Indeed, if $O_1, \dots, O_n \in \mathscr{B}$, then $O := O_1 \cap \dots \cap O_n$ also belongs to $\mathscr{B}$, and so $O \cap \mathrm{st}(A)$ contains some point x_0. In view of $x_0 \in \mathrm{st}(A)$, there is some $y \in A$ with $y \in \mathrm{mon}(x_0)$. Since O is a neighborhood of x_0, this implies $y \in {}^*O$. Hence, we found some $y \in A$ with $y \in {}^*O = {}^*O_1 \cap \dots \cap {}^*O_n$, and so φ is concurrent on ${}^\sigma\mathscr{B}$, as claimed.

Since ${}^\sigma\mathscr{B}$ consists only of standard sets (and has not a larger cardinality than $\mathcal{P}(X)$), φ is satisfied on ${}^\sigma\mathscr{B}$, i.e. there is some $y \in A$ such that $y \in {}^*O$ for any $O \in \mathscr{B}$. This means $y \approx_{\mathscr{O}} {}^*x$. Hence $x \in \mathrm{st}(A)$, and so $\overline{\mathrm{st}(A)} = \mathrm{st}(A)$ is closed. □

Definition 12.38. A set $A \subseteq X$ is *compact* if each open cover of A has a finite subcover. This means that whenever $\mathscr{C}$ is a system of open sets with $A \subseteq \bigcup \mathscr{C}$, there exist finitely many $O_1, \dots, O_n \in \mathscr{C}$ with $A \subseteq O_1 \cup \dots \cup O_n$.

One of the most important aspects of nonstandard topology is that compact sets have a very natural characterization:

Theorem 12.39. *A set* $A \subseteq X$ *is compact if and only if each point of* ${}^*A$ *is infinitely close to some (standard) point of* ${}^\sigma A$*, i.e. if and only if*

$${}^*A \subseteq \mathrm{st}^{-1}(A).$$

Proof. Let A be compact, and $y \in {}^*A$. Assume by contradiction that we find no $x \in A$ with $y \approx_{\mathscr{O}} {}^*x$. Then we find for each $x \in A$ some open set $O \subseteq X$ with $x \in O$ and $y \notin {}^*O$. Hence, the set $\mathscr{C}$ of all open sets $O \subseteq X$ with $y \notin {}^*O$ is an open cover of X. Since A is compact, we have $A \subseteq O_1 \cup \dots \cup O_n$ with $O_k \in \mathscr{C}$. But since $*$ is a superstructure monomorphism, this implies $y \in {}^*A \subseteq {}^*O_1 \cup \dots \cup {}^*O_n$, and so $y \in {}^*O_k$ for some $O_k \in \mathscr{C}$, a contradiction to the definition of $\mathscr{C}$.

Conversely, let each $y \in {}^*A$ be infinitely close to some standard point of ${}^\sigma A$, and $\mathscr{C}$ be an open cover of A. For any $y \in {}^*A$, we find some $x \in A$ with $y \approx_{\mathscr{O}} {}^*x$ and some $O \in \mathscr{C}$ with $x \in O$; hence, $y \in {}^*O$. This proves ${}^*A \subseteq \bigcup {}^\sigma\mathscr{C}$. By Exercise 44, we thus find a finite $\mathscr{C}_0 \subseteq \mathscr{C}$ with $A \subseteq \bigcup \mathscr{C}_0$, i.e. A is compact. □

For $X = \mathbb{R}$, we defined compact sets as closed and bounded. However, since the nonstandard characterization of Theorem 12.39 is the same as the characterization of Corollary 7.10, we may conclude that the definitions actually are equivalent. This nonstandard argument implies in particular the classical Heine-Borel theorem:

Corollary 12.40 (Heine-Borel)**.** *A subset $A \subseteq \mathbb{R}$ is compact in the sense of Definition* 12.38 *if and only if it is closed and bounded.*

In literature, the term "compact" is sometimes reserved for compact Hausdorff spaces, because these spaces have particularly convenient properties. This appears natural from the nonstandard characterization:

Corollary 12.41. *The topological space X is a compact Hausdorff space if and only if*

$$\mathrm{st} : {}^*X \to X,$$

i.e. if and only if for each $y \in {}^*X$ there is precisely one $x \in X$ with $y \approx_{\mathscr{O}} {}^*x$.

Proof. Combine Proposition 12.30 and Theorem 12.39. □

Exercise 61. Prove the following statements by nonstandard methods:

1. A closed set which is contained in a compact set is compact.
2. Compact sets in Hausdorff spaces are closed.

Exercise 62. Let $*$ be a compact $\mathcal{P}(X)$-enlargement. Prove that $A \subseteq X$ is compact if and only if $\mathrm{st}^{-1}(A) \cap {}^*A$ is internal.

Hint: Apply Theorem 8.16.

Definition 12.42. A topological space is called a T_3 *space* if points may be divided from closed sets by open sets, i.e. whenever A is closed and $x \notin A$, there exist disjoint open sets O_0, O_1 with $x \in O_0$ and $A \subseteq O_1$.

Lemma 12.43. *Let X be a T_3 space. Then a set $M \subseteq X$ is compact if and only if for each open cover $\mathscr{C}$ of M there is a finite $\mathscr{C}_0 \subseteq \mathscr{C}$ with $M \subseteq \overline{\bigcup \mathscr{C}_0}$.*

Proof. Necessity is trivial. To prove sufficiency, consider the family $\mathscr{C}_1$ of all open sets O with the property that there is some $O' \in \mathscr{C}$ with $\overline{O} \subseteq O'$. Then $\mathscr{C}_1$ is an open cover of M. Indeed, for any $x \in M$ there is some $O \in \mathscr{C}$ with $x \in O$. Putting $A := X \setminus O$, we find disjoint open sets O_0, O_1 with $x \in O_0$ and $A \subseteq O_1$. Then $A_0 := X \setminus O_1$ is closed and contains O_0 which means $\overline{O}_0 \subseteq A_0 \subseteq O$. Hence, $O_0 \in \mathscr{C}_1$, and so $x \in O_0 \subseteq \bigcup \mathscr{C}_1$.

By assumption, we find $O_1, \dots, O_n \in \mathscr{C}_1$ with $M \subseteq \overline{O_1 \cup \dots \cup O_n} = \overline{O}_1 \cup \dots \cup \overline{O}_n$. The definition of $\mathscr{C}_1$ implies that there are $O_i' \in \mathscr{C}$ with $O_i' \supseteq \overline{O}_i$, and so $M \subseteq O_1' \cup \dots \cup O_n'$. $\square$

Theorem 12.44. *Let X be a T_3 space, and $A \subseteq X$. Then $\overline{A}$ is compact if and only if ${}^*A \subseteq \mathrm{ns}({}^*X)$.*

Proof. If $\overline{A}$ is compact, then Theorem 12.39 implies ${}^*A \subseteq {}^*(\overline{A}) \subseteq \mathrm{st}^{-1}(\overline{A}) \subseteq \mathrm{ns}({}^*X)$.

Conversely, let ${}^*A \subseteq \mathrm{ns}({}^*X)$, and let $\mathscr{C}$ be an open cover of $\overline{A}$. For any $y \in {}^*A$ we find some $x \in X$ with $y \approx_{\mathscr{O}} {}^*x$. By Theorem 12.35, we have $x \in \overline{A}$, and so there is some $O \in \mathscr{C}$ with $x \in O$; hence, $y \in {}^*O$. This proves ${}^*A \subseteq \bigcup {}^\sigma\mathscr{C}$. By Exercise 44, we thus find a finite $\mathscr{C}_0 \subseteq \mathscr{C}$ with $A \subseteq \bigcup \mathscr{C}_0$. Then $\overline{A} \subseteq \overline{\bigcup \mathscr{C}_0}$, and so $\overline{A}$ is compact by Lemma 12.43. $\square$

For compact enlargements, we have a generalization to internal sets (similar to Theorem 12.37):

Exercise 63. Let $*$ be a compact $\mathcal{P}(X)$-enlargement. If X is T_3 and $A \subseteq \mathrm{ns}({}^*X)$ is internal, prove that $\mathrm{st}(A)$ is compact.

Hint: Apply Lemma 12.43 and Theorem 8.16.

We now discuss convergence in a topological space:

Definition 12.45. A sequence $x_n \in X$ converges to $x \in X$, if for each $U \in \mathscr{U}(x)$, we have $x_n \in U$ for all except finitely many n. A filter $\mathscr{F}$ over X converges to x, if $\mathscr{U}(x) \subseteq \mathscr{F}$. We write $x_n \to x$ resp. $\mathscr{F} \to x$.

These definitions are related:

Proposition 12.46. *A sequence x_n converges to x if and only if the filter $\mathscr{F}$ generated by the sets $F_n = \{x_n, x_{n+1}, x_{n+2}, \dots\}$ $(n = 1, 2, \dots)$ converges to x.*

Proof. We have $\mathscr{U}(x) \subseteq \mathscr{F}$ if and only if for each $U \in \mathscr{U}(x)$ there are $n_1, \dots, n_k$ with $F_{n_1} \cap \dots \cap F_{n_k} \subseteq U$. Since $F_{n_1} \cap \dots \cap F_{n_k} = F_{\max\{n_1,\dots,n_k\}}$, this is the case if and only if for each $U \in \mathscr{U}(x)$ there is some n with $F_n \subseteq U$, i.e. if and only if all except finitely many x_n belong to U. $\square$

Theorem 12.6 implies:

Proposition 12.47. *We have $\mathscr{F} \to x$ if and only if $\mathrm{mon}(\mathscr{F}) \subseteq \mathrm{mon}(x)$.*

Proof. By Theorem 12.6, $\mathscr{U}(x) \subseteq \mathscr{F}$ if and only if $\mathrm{mon}(\mathscr{F}) \subseteq \mathrm{mon}(\mathscr{U}(x)) = \mathrm{mon}(x)$. $\square$

Corollary 12.48. *If X is a Hausdorff space, then $\mathscr{F}$ converges to at most one point.*

Proof. The relations $\mathrm{mon}(\mathscr{F}) \subseteq \mathrm{mon}(x)$ and $\mathrm{mon}(\mathscr{F}) \subseteq \mathrm{mon}(y)$ imply $\mathrm{mon}(x) \cap \mathrm{mon}(y) \neq \varnothing$ (because $\mathrm{mon}(\mathscr{F}) \neq \varnothing$), and so $x = y$ by Proposition 12.30. □

Exercise 64. Prove by nonstandard methods that in a compact space each ultrafilter converges. (Actually, this characterizes compact spaces, but the converse implication is proved more easily by standard methods.)

We prove now that standard open sets can be characterized in terms of monads:

Theorem 12.49. *Let $*$ be a compact $\mathcal{P}(X)$-enlargement, and $A \subseteq {}^*X$ be internal.*

1. *If $\mu(\{a\}) \subseteq A$ for some $a \in A$, then A contains some standard open set ${}^*O$ (i.e. $O \subseteq X$ is open) with $a \in {}^*O$.*
2. *If $\mu(\{a\}) \subseteq A$ for any $a \in A$, then A is a standard open set.*

Proof. 1. Let $\mathscr{F}$ be the standard filter generated by a, and let $\mathscr{B}$ denote the system of all $F \in {}^*\mathscr{F}$ with $a \in F \subseteq A$. By $\mathrm{mon}(\mathscr{F}) = \mu(\{a\}) \subseteq A$, we have $\mathscr{B} \supseteq \{F \in {}^*\mathscr{F} : F \subseteq \mathrm{mon}(\mathscr{F})\}$, and so Theorem 12.7 implies that ${}^*F \in \mathscr{B}$ for some $F \in \mathscr{F}$, i.e. ${}^*F \subseteq A$. By definition of $\mathscr{F}$, we have $F \supseteq O_1 \cap \cdots \cap O_n = O$ for open $O_k \subseteq X$ with $a \in {}^*O_k$. Then O is open with ${}^*O = {}^*O_1 \cap \cdots \cap {}^*O_n$, and so $a \in {}^*O \subseteq {}^*F \subseteq A$.
2. Let $\mathscr{A}$ denote the system of all open sets $O \subseteq X$ with ${}^*O \subseteq A$. In view of 1., we find for any $a \in A$ some open $O \subseteq X$ with $a \in {}^*O \subseteq A$, and so $A \subseteq \bigcup^\sigma \mathscr{A}$. Theorem 8.16 implies that there is a finite $\mathscr{A}_0 \subseteq \mathscr{A}$ with $A \subseteq \bigcup^\sigma \mathscr{A}_0$. We have $\mathscr{A}_0 = \{O_1, \ldots, O_n\}$ where $O_k \subseteq X$ are open, and ${}^*O_k \subseteq A$. Then $O := O_1 \cup \cdots \cup O_n$ is open, and ${}^*O = A$: Indeed, $A \subseteq \bigcup^\sigma \mathscr{A}_0 = {}^*O_1 \cup \cdots \cup {}^*O_n \subseteq A$, and so $A = {}^*O_1 \cup \cdots \cup {}^*O_n = {}^*O$.

□

By choosing an appropriate topology, we can characterize standard elements in terms of monads: The *discrete topology* on X is $\mathscr{O} = \mathcal{P}(X)$, i.e. any subset of X is open with respect to the discrete topology.

Corollary 12.50. *Let X be equipped with the discrete topology. Let $*$ be a compact $\mathcal{P}(X)$-enlargement, and $A \subseteq {}^*X$ be internal.*

1. *A contains a standard element if and only if $\mu(\{a\}) \subseteq A$ for some $a \in A$.*
2. *A is standard if and only if $\mu(\{a\}) \subseteq A$ for any $a \in A$.*

Proof. Sufficiency follows immediately from Theorem 12.49. For the converse implications observe that if ${}^*x \in A$ for some $x \in X$, then we have for $a = {}^*x$ that $O := \{x\}$ is open and $a \in {}^*O$, and so $\mu(\{a\}) \subseteq {}^*O = \{{}^*x\} = \{a\} \subseteq A$. Similarly, if $A = {}^*O$ for some $O \subseteq X$, then O is open, and so we have for any $a \in A$ in view of $a \in {}^*O$ that $\mu(\{a\}) \subseteq {}^*O = A$. □

As a nice application of nonstandard methods, let us give a simple proof of the famous Tychonoff compactness theorem.

Definition 12.51. Let $X = \prod X_i$ with topological spaces X_i $(i \in I)$. Then the following topology on X is called the *product topology*: Let $\mathscr{B}$ denote the system of all sets of the form

$$\prod_{i \in I} O_i$$

where $O_i = X_i$ for all except finitely many $i \in I$, and $O_i \subseteq X_i$ is open for all $i \in I$. Then O is open in the product topology if and only if it is a union of sets from $\mathscr{B}$.

The condition that $O_i = X_i$ for all except finitely many $i \in I$ appears rather artificial. However, there are two main reasons why this should be included in the above definition of the product topology (besides the fact that the above definition has many applications):

It turns out that the above definition is the smallest topology such that each of the projections $p_i : X \to X_i$ is continuous (continuity will be defined in Section 12.4). In nonstandard terms, this fact reads as follows:

We will assume henceforth that we have given a family X_i $(i \in I)$ of topological spaces where X_i, I, and $\{X_i : i \in I\}$ are all entities of $\widehat{S}$. Then also $U := \bigcup X_i$, U^I, and thus also $X := \prod X_i$ are entities of $\widehat{S}$. As before, we assume that $*$ is an $\mathcal{P}(X)$-enlargement.

We use the notation of Corollary A.6, i.e. for $i \in {}^*I$, we define ${}^*X_i := {}^*f(i)$ where f denotes the function $i \mapsto X_i$. Recall (Exercise 80), that the elements of ${}^*X$ are precisely the internal elements of $\prod_{i \in {}^*I} {}^*X_i$. In particular (Corollary A.6), each $x \in {}^*X$ is an internal function

$$x : {}^*I \to \bigcup_{i \in {}^*I} {}^*X_i.$$

Theorem 12.52. *In the product topology, we have* $y \approx_{\mathscr{O}} {}^*x$ *if and only if* $y({}^*i) \approx_{\mathscr{O}} {}^*x({}^*i) = {}^*(x(i))$ *for each* $i \in I$.

Proof. Suppose that $y({}^*i) \approx_{\mathscr{O}} {}^*x({}^*i)$ for each $i \in I$. Let $O \subseteq X$ be open with ${}^*x \in {}^*O$. We have to prove that $y \in {}^*O$. But ${}^*x \in {}^*O$ implies $x \in O$, and so we have $x \in B \subseteq O$ for some $B \in \mathscr{B}$ where $\mathscr{B}$ is as in Definition 12.51. By definition of $\mathscr{B}$, we find finitely many $i_1, \ldots, i_n \in I$ and open sets $O_{i_k} \subseteq X_{i_k}$ such that $B = \prod_{i \in I} O_i$ for each $i \in I$ where we put $O_i := X_i$ for $i \neq i_1, \ldots, i_n$. Exercise 80 implies that ${}^*B$ consists of all internal elements of $\prod_{i \in {}^*I} {}^*O_i$. Note that the transfer principle implies

$$\forall \underline{x} \in {}^*I : ((\underline{x} \neq {}^*i_1 \wedge \cdots \wedge \underline{x} \neq {}^*i_n) \implies {}^*O_{\underline{x}} = {}^*X_{\underline{x}}).$$

Since $y({}^*i_k) \approx_{\mathscr{O}} {}^*x(i_k) \in {}^*(O_{i_k})$, we have $y({}^*i_k) \in {}^*(O_{i_k}) = {}^*O_{{}^*i_k}$ (recall the remarks preceding Corollary A.6). Thus, y is an internal function which satisfies $y(i) \in {}^*O_i$ for all $i \in {}^*I$ (for $i \neq {}^*i_1, \ldots, {}^*i_k$ we have ${}^*O_i = {}^*X_i$, as we have shown above). This proves $y \in {}^*B \subseteq {}^*O$, as desired.

Conversely, let $y \approx_{\mathscr{O}} {}^*x$. If $i_0 \in I$ and some open $O_{i_0} \subseteq X_{i_0}$ with $x(i_0) \in O_{i_0}$ are given, put $O_i := X_i$ $(i \neq i_0)$ and $O := \prod_{i\in I} O_i$. Then O is open with $x \in O$. Hence, ${}^*x \in {}^*O$, and so our assumption implies $y \in {}^*O$. Since ${}^*O$ consists of all internal elements of $\prod_{i\in {}^*I} {}^*O_i$ and since ${}^*O_{{}^*i_0} = {}^*(O_{i_0})$, we must have $y({}^*i_0) \in {}^*(O_{i_0})$. Hence, $y({}^*i_0) \approx_{\mathscr{O}} {}^*(x(i_0))$. □

Corollary 12.53. $X := \prod_{i\in I} X_i$ *is a Hausdorff space if and only if each* X_i *is a Hausdorff space.*

Proof. If X_{i_0} is not a Hausdorff space, there are points $a, b \in X_{i_0}$, $a \neq b$ such that $\operatorname{mon}(a) \cap \operatorname{mon}(b)$ contains some point c (Proposition 12.30). Choose some $x \in X$ with $x(i_0) = a$, and put $y(i) = x(i)$ $(i \in I \setminus \{i_0\})$ and $y(i_0) = b$. Consider the function $z(i) := {}^*y(i)$ $(i \in I \setminus \{i_0\})$ and $z(i_0) := c$. Then z is an internal function (Exercise 8), and in view of Theorem 12.52, we have $z \in \operatorname{mon}(x) \cap \operatorname{mon}(y)$ although $x \neq y$. Hence, Proposition 12.30 implies that X is not a Hausdorff space.

Conversely, if X is not a Hausdorff space, we find elements $x \neq y$ in X such that $\operatorname{mon}(x) \cap \operatorname{mon}(z)$ contains some element z. Choose some i_0 with $x(i_0) \neq y(i_0)$. Then X_{i_0} is not a Hausdorff space, because Theorem 12.52 implies $z({}^*i_0) \in \operatorname{mon}(x(i_0)) \cap \operatorname{mon}(y(i_0))$. □

The other reason for the definition of the product topology is that the following important theorem of Tychonoff holds, which has many applications. We note that all known proofs of the Tychonoff theorem are rather technical so that the following nonstandard proof is an essential simplification:

Corollary 12.54 (Tychonoff)**.** $X := \prod_{i\in I} X_i$ *is compact if and only if each* X_i *is compact.*

Proof. Let X be compact. A standard argument immediately implies that X_{i_0} is compact (because each projection $p_i : X \to X_i$ is continuous, as mentioned above). However, for completeness we provide a nonstandard proof: By Theorem 12.39, we have to prove that for each $b \in {}^*X_{i_0}$ there is some $a \in X_{i_0}$ with $b \approx_{\mathscr{O}} {}^*a$. There is some $y \in {}^*X$ with $y({}^*i_0) = b$. Since X is compact, we find some $x \in X$ with $y \approx_{\mathscr{O}} {}^*x$. Then $y({}^*i_0) \approx_{\mathscr{O}} {}^*(x(i_0))$ by Theorem 12.52, and so $a = x(i_0) \in X_{i_0}$ satisfies $b \approx_{\mathscr{O}} {}^*a$.

The converse direction is the one which is hard to prove by standard methods: Suppose that all X_i are compact. Let $y \in {}^*X$. For each $i \in I$, we have $y({}^*i) \in {}^*X_{{}^*i} = {}^*(X_i)$. Since X_i is compact, we find some $x(i) \in \operatorname{st}(y({}^*i))$. Then $x \in X$ (axiom of choice!), and Theorem 12.52 implies $y \approx_{\mathscr{O}} {}^*x$. □

Some notes are in order: It lies in the nature of things that we had to use the axiom of choice in its full generality to prove the Tychonoff theorem. In fact, the Tychonoff theorem is actually equivalent to the axiom of choice [Kel50] (this paper contains a minor mistake which however can be corrected, see [LRN51]). However, to prove Tychonoff's theorem for Hausdorff spaces, one does not need the full power of the axiom of choice: In fact, the Tychonoff theorem for Hausdorff spaces is actually equivalent to the so-called maximal ideal theorem [LRN54] which in turn is equivalent to Theorem 4.9 [Sik64]. The most difficult of these implications follows from our above proof of the Tychonoff theorem: In fact, if all X_i are compact Hausdorff spaces, then st is a *function* by Proposition 12.30, and so no axiom of choice is required to define the function x in the above proof. So in this case, we only made use of the axiom of choice in the construction of the ultrapower model. A careful analysis shows that a map $*$ sufficient for our proof may be defined by only applying Theorem 4.9 and no other form of the axiom of choice.

12.4 Functions in Nonstandard Topology

Let X and Y be topological spaces. Assume that $X, Y \in \widehat{S}$ are entities and that $* : \widehat{S} \to \widehat{{}^*S}$ is an enlargement (actually, it suffices for the following considerations that $*$ is a $\mathcal{P}(X) \cup \mathcal{P}(Y)$-enlargement).

Let $f : X \to Y$. Recall that if $\mathscr{F}$ is a filter over X, then $f(\mathscr{F})$ denotes the filter generated by $\{f(F) : F \in \mathscr{F}\}$.

Theorem 12.55. *We always have* $\operatorname{mon}(f(\mathscr{F})) \supseteq {}^*f(\operatorname{mon}(\mathscr{F}))$. *Moreover, equality holds if* $*$ *is also a* compact $\mathcal{P}(X)$*-enlargement.*

Proof. Since $\mathscr{B} := \{f(F) : F \in \mathscr{F}\}$ generates $f(\mathscr{F})$, Lemma 12.11 implies in view of ${}^*(f(F)) = {}^*f({}^*F)$ (Theorem 3.13) that

$$\begin{aligned} \operatorname{mon}(f(\mathscr{F})) &= \bigcap{}^\sigma\mathscr{B} = \bigcap\{{}^*(f(F)) : F \in \mathscr{F}\} \\ &= \bigcap\{{}^*f(F) : F \in {}^\sigma\mathscr{F}\} \supseteq {}^*f(\bigcap{}^\sigma\mathscr{F}) = {}^*f(\operatorname{mon}(\mathscr{F})). \end{aligned}$$

To see that we have equality if $*$ is a compact $\mathcal{P}(X)$-enlargement, assume that there is some $y \in \bigcap\{{}^*f(F) : F \in {}^\sigma\mathscr{F}\}$ which is not contained in ${}^*f(\operatorname{mon}(\mathscr{F}))$. Put $\mathscr{B}_y := \{F \in {}^*\mathscr{F} : y \notin {}^*f(F)\}$. Then $\mathscr{B}_y$ is internal by the internal definition principle, and $\mathscr{B}_y \supseteq \{F \in {}^*\mathscr{F} : F \subseteq \operatorname{mon}(\mathscr{F})\}$, because $y \notin {}^*f(\operatorname{mon}(\mathscr{F}))$. Theorem 12.7 now implies that $\mathscr{B}_y$ contains some standard $F \in {}^\sigma\mathscr{F}$, i.e. $y \notin {}^*f(F)$, a contradiction to our choice of y. □

Definition 12.56. A function $f : X \to Y$ is called *continuous* at $x \in X$, if for each neighborhood $V \in \mathscr{U}(f(x))$ there is a neighborhood $U \in \mathscr{U}(x)$ with $f(U) \subseteq V$.

The function f is called *continuous*, if it is continuous at each $x \in X$.

Theorem 12.57. *For $f : X \to Y$ the following statements are equivalent:*

1. *f is continuous at x.*
2. *$\mathscr{F} \to x$ implies $f(\mathscr{F}) \to x$ for any filter $\mathscr{F}$ over X.*
3. *$\mathscr{U}(f(x)) \subseteq f(\mathscr{U}(x))$.*
4. *$\operatorname{mon}(f(\mathscr{U}(x))) \subseteq \operatorname{mon}(\mathscr{U}(f(x)))$.*
5. *${}^*f(\operatorname{mon}(x)) \subseteq \operatorname{mon}(f(x))$.*
6. *$y \approx_{\mathscr{O}} {}^*x$ implies ${}^*f(y) \approx_{\mathscr{O}} {}^*(f(x)) = {}^*f({}^*x)$.*

Proof. We first prove the equivalence of the first three statements: If f is continuous at x, and $\mathscr{F} \to x$ (i.e. $\mathscr{U}(x) \subseteq \mathscr{F}$), then we find for each $V \in \mathscr{U}(f(x))$ some $U \in \mathscr{U}(x) \subseteq \mathscr{F}$ with $f(U) \subseteq V$, and so V belongs to the filter generated by $\{f(U) : U \in \mathscr{F}\}$. Thus 2. holds. If 2. holds, then we find for the particular choice $\mathscr{F} = \mathscr{U}(x)$ that $f(\mathscr{U}(x)) \to f(x)$ which means that 3. is satisfied. Finally, assume that 3. holds. Recall that by Lemma 5.27, $f(\mathscr{U}(x))$ consists precisely of those sets $V \subseteq Y$ for which $U := \{x : f(x) \in V\} \in \mathscr{U}(x)$. Hence, if $\mathscr{U}(f(x)) \subseteq f(\mathscr{U}(x))$ and $V \in \mathscr{U}(f(x))$, we find some $U \in \mathscr{U}(x)$ with $f(U) \subseteq V$, and so f is continuous at x.

Theorem 12.6 5. implies that the inclusions 3. and 4. are equivalent.

Noting that $\operatorname{mon}(\mathscr{U}(f(x))) = \operatorname{mon}(f(x))$, and that ${}^*f(\operatorname{mon}(x)) = {}^*f(\mathscr{U}(f(x))) \subseteq \operatorname{mon}(f(\mathscr{U}(x)))$ in view of Theorem 12.55, we see that 4. implies 5.; moreover, equivalence follows if $*$ is a compact enlargement. To see the converse inclusion without this additional requirement, assume that 5. holds. Choose some internal $U \in {}^*\mathscr{U}(x)$ with $U \subseteq \operatorname{mon}(x)$ (Proposition 12.19). Given $V \in \mathscr{U}(f(x))$, we have by assumption ${}^*f(U) \subseteq \operatorname{mon}(f(x)) \subseteq {}^*V$, i.e. we have proved

$$\exists \underline{u} \in {}^*\mathscr{U}(x) : {}^*f(\underline{u}) \subseteq {}^*V.$$

The inverse form of the transfer principle implies that there is some $U \in \mathscr{U}(x)$ with $f(U) \subseteq V$. Hence, f is continuous at x. The equivalence of 6. with 5. is trivial. □

Corollary 12.58. *If f is continuous at x, then $x_n \to x$ implies $f(x_n) \to f(x)$.*

Proof. Let $x_n \to x$, and $\mathscr{F}$ be the filter generated by the sets $\{x_n, x_{n+1}, \ldots\}$ $(n = 1, 2, \ldots)$. Then $\mathscr{F} \to x$ (Proposition 12.46), and so $f(\mathscr{F}) \to f(x)$ by Theorem 12.57. By definition, $f(\mathscr{F})$ is the filter generated by the sets $\{f(x_n), f(x_{n+1}), \ldots\}$ $(n = 1, 2, \ldots)$. Thus, Proposition 12.46 implies $f(x_n) \to f(x)$. □

We point out that the converse to Corollary 12.58 does not hold in general. For counterexamples (and assumptions which imply that the converse implication holds), we refer the reader to books on topology.

Exercise 65. Prove by nonstandard methods that a continuous function maps compact sets into compact sets.

Using Proposition 12.5, one can prove that a function f is continuous if and only if preimages of open sets are open. For one direction of this statement, one can give an easier nonstandard proof:

Exercise 66. Prove by nonstandard methods that for any continuous function $f : X \to Y$ preimages of open sets are open.

§13 Uniform Structures

13.1 Uniform Spaces

There are some concepts which appear topological but which actually cannot be described in topological spaces: Uniform convergence, uniform continuity, or Cauchy sequences. In (pseudo)metric spaces, one can define these concepts: For example, a sequence f_n of functions with values in a pseudometric space *converges uniformly* to a function f, if for each $\varepsilon \in \mathbb{R}_+$ one finds some index N such that $d(f_n(x), f(x)) < \varepsilon$ $(n \geq N)$ simultaneously for all x. This definition is possible, since ε determines not only a neighborhood, but actually a system of neighborhoods for each point in the space (in particular for each of the points $f(x)$). Thus, if one intends to introduce such a concept in more general topological spaces, one should consider families of neighborhoods. This is the motivation for the definition of a so-called *uniform structure*.

Recall that sets $U \subseteq X \times X$ are relations. Hence, the notation

$$U^{-1} := \{(y, x) : (x, y) \in U\}$$

is natural, and if $V \subseteq X \times X$ and $x \in X$ also

$$U \circ V := \{(x, z) \mid \exists y \in X : ((x, y) \in V \wedge (y, z) \in U)\}$$

and

$$U(x) := \{y : (x, y) \in U\}.$$

We write also U^2 for $U \circ U$. We already made use of these conventions for the particular relation st (see the remarks in front of Theorem 12.35).

Similar arguments as in Theorem 3.13 show that $({}^*U)^{-1} = {}^*(U^{-1})$, ${}^*U \circ {}^*V = {}^*(U \circ V)$, and ${}^*U({}^*x) = {}^*(U(x))$.

Definition 13.1. A *uniform structure* over a space X is a filter $\mathscr{U}$ over $X \times X$ such that each $U \in \mathscr{U}$ satisfies:

1. $\Delta := \{(x, x) : x \in X\} \subseteq U$.
2. $U^{-1} \in \mathscr{U}$.
3. There is some $V \in \mathscr{U}$ with $V^2 \subseteq U$.

Each uniform structure generates a topology in a canonical way: Given $x \in X$, put $\mathscr{U}(x) := \{U(x) : U \in \mathscr{U}\}$.

Proposition 13.2. *Let $\mathscr{O}$ be the system of all sets O with the property that $O \in \mathscr{U}(x)$ for any $x \in O$. Then $\mathscr{O}$ is a topology on X, and $\mathscr{U}(x)$ is the corresponding neighborhood filter of x.*

The topology is Hausdorff if and only if the relation $(x, y) \in U$ for any $U \in \mathscr{U}$ implies $x = y$.

Proof. We prove first that $\mathscr{O}$ is a topology: Clearly, $\varnothing, X \in \mathscr{O}$. Let $O_1, O_2 \in \mathscr{O}$ and $O := O_1 \cap O_2$. For any $x \in O$, we find sets $U_1, U_2 \in \mathscr{U}$ with $O_i = U_i(x)$. Since $\mathscr{U}$ is a filter, we have $U := U_1 \cap U_2 \in \mathscr{U}$. Hence $O = U(x) \in \mathscr{U}(x)$, and so $O \in \mathscr{O}$. Similarly, if $\mathscr{O}_0 \subseteq \mathscr{O}$ the set $O := \bigcup \mathscr{O}_0$ belongs to $\mathscr{O}$: For any $x \in O$, there is some $U \in \mathscr{U}$ with $U(x) \subseteq O$. Putting $V := U \cup \{(x,y) : y \in O\}$, we have $V \in \mathscr{U}$ (because $V \supseteq U \in \mathscr{U}$) and $O = V(x) \in \mathscr{U}(x)$. Hence, $O \in \mathscr{O}$, and so $\mathscr{O}$ is a topology.

If W is a neighborhood of x, then there is some $U \in \mathscr{U}$ with $W \supseteq U(x)$. Then $V := U \cup \{(x,y) : y \in W\}$ belongs to $\mathscr{U}$, and so $W = V(x) \in \mathscr{U}(x)$.

The only nontrivial part of the proof is the converse implication, i.e. that any $W \in \mathscr{U}(x)$ actually is a neighborhood of x: Let $W \in \mathscr{U}(x)$, i.e. $W = U(x)$ for some $U \in \mathscr{U}$. Let O be the set of all $y \in X$ with the property that there is some $V \in \mathscr{U}$ with $V(y) \subseteq W$. Then $O \in \mathscr{O}$: Given $y \in O$, choose some $V \in \mathscr{U}$ with $V(y) \subseteq W$. There is some $V_0 \in \mathscr{U}$ with $V_0^2 \subseteq V$. For any $z \in V_0(y)$, we have $V_0(z) \subseteq V_0^2(y) \subseteq V(y) \subseteq W$, and so $z \in O$. Hence, $V_0(y) \subseteq O$, which implies that $O = V_1(y)$ for $V_1 := V_0 \cup \{(y,z) : z \in O\} \in \mathscr{U}$. Thus, $O \in \mathscr{U}(y)$ for any $y \in O$ which means $O \in \mathscr{O}$, as claimed. Since $y \in V(y)$ for $V \in \mathscr{U}$, the definition of O implies $O \subseteq W$. Finally, since $U(x) \subseteq W$, we have $x \in O$. Thus, W is indeed a neighborhood of x.

If the topology is Hausdorff and $x \neq y$, then there are $U, V \in \mathscr{U}$ with $U(x) \cap V(y) = \varnothing$. Since $y \in V(y)$, this implies $y \notin U(x)$, and so $(x,y) \notin U$. Conversely, assume that $(x,y) \in U$ for any $U \in \mathscr{U}$ implies $x = y$. Then the topology is Hausdorff, since for any $x \neq y$ we find some $U \in \mathscr{U}$ with $(x,y) \notin U$. Choose some $V \in \mathscr{U}$ with $V^2 \subseteq U$. Then $V^{-1}(x) \cap V(y) = \{z : (x,z), (z,y) \in V\} = \varnothing$, since otherwise $(x,y) \in V^2 \subseteq U$, a contradiction. Hence, x and y have disjoint neighborhoods $V^{-1}(x)$ and $V(y)$. □

For evident reasons, the above topology is called the *topology generated* by the uniform structure $\mathscr{U}$. We call X equipped with this topology a *uniform space*. (Similarly as we did for topological spaces, we usually do not mention $\mathscr{U}$ explicitly).

Example 13.3. Let X be a (pseudo)metric space. Let $\mathscr{U}$ be the system of all sets $U \subseteq X \times X$ for which there is some $\varepsilon \in \mathbb{R}_+$ with

$$\{(x,y) : d(x,y) < \varepsilon\} \subseteq U.$$

Then $\mathscr{U}$ is a uniform structure. A set O is open in the generated topology if and only if for any $x \in O$ there is some $\varepsilon \in \mathbb{R}_+$ such that $\{y : d(x,y) < \varepsilon\} \subseteq O$. Thus, the topology generated by $\mathscr{U}$ is the topology of X in the usual sense.

From now on, we understand by a (pseudo)metric space always a space with this uniform structure and the corresponding topology.

The previous example is generalized by the following:

Example 13.4. Let D be a family of pseudometrics. Let $\mathscr{U}$ be the filter which is generated by the system of all sets

$$\{(x,y) : d(x,y) < \varepsilon\} \qquad (d \in D,\ \varepsilon \in \mathbb{R}_+).$$

Thus, we have $U \in \mathscr{U}$ if and only if there are finitely many $d_1, \ldots, d_n \in D$ and some $\varepsilon \in \mathbb{R}_+$ such that

$$\{(x,y) : d_1(x,y) < \varepsilon \wedge \cdots \wedge d_n(x,y) < \varepsilon\} \subseteq U.$$

Then $\mathscr{U}$ is a uniform structure. We call $\mathscr{U}$ the uniform structure *induced by the family D*.

It is a rather deep result of elementary topology that each uniform structure is actually induced by a family of pseudometrics. This result is remarkable for two reasons: First, it implies some criteria for the metrizability of uniform spaces. Moreover, it establishes a connection between a general uniform space X and the system $\mathbb{R}$ of real numbers (because each pseudometric takes its values in $\mathbb{R}$). However, we will not make use of this result and so do not provide a proof. The reader who is interested in a proof is referred to books on elementary topology (see e.g. [vQ79, 11.34]).

The importance of uniform structures becomes clear in the nonstandard description. As in the previous sections, we assume in the following that $X \in \widehat{S}$ is an entity and that $* : \widehat{S} \to {}^*\widehat{S}$ is a $\mathcal{P}(X)$-enlargement. Then $*$ is even a $\mathcal{P}(X \times X)$-enlargement (because $X \times X$ has the same cardinality as X if X is infinite).

Given a filter $\mathscr{U}$ over $X \times X$, note that any $U \in \mathscr{U}$ satisfies ${}^*U \subseteq {}^*(X \times X) = {}^*X \times {}^*X$, and so $\operatorname{mon}(\mathscr{U}) \subseteq {}^*X \times {}^*X$. Hence, it makes sense to define

$$x \approx_{\mathscr{U}} y \iff (x,y) \in \operatorname{mon}(\mathscr{U}).$$

If $\mathscr{U}$ is a uniform structure, we call x *infinitely $\mathscr{U}$-close* to y if $x \approx_{\mathscr{U}} y$. It makes sense to consider only uniform structures in this connection:

Proposition 13.5. *A filter $\mathscr{U}$ over $X \times X$ is a uniform structure if and only if $\approx_{\mathscr{U}}$ is an equivalence relation on ${}^*X$.*

Proof. In fact, the three properties of Definition 13.1 correspond to the properties of equivalence relations for $\approx_{\mathscr{U}}$:

The first property is equivalent to the reflexivity of $\approx_{\mathscr{U}}$. Indeed, we have $\Delta := \{(\underline{x}, \underline{y}) \in X \times X \mid \underline{x} = \underline{y}\} \subseteq U$ for any $U \in \mathscr{U}$ if and only if ${}^*\Delta \subseteq {}^*U$ for any $U \in \mathscr{U}$. Since the standard definition principle for relations implies ${}^*\Delta = \{(\underline{x}, \underline{y}) \in {}^*X \times {}^*X \mid \underline{x} = \underline{y}\} = \{(x,x) : x \in {}^*X\}$, this is the case if and only if $x \approx_{\mathscr{U}} x$ for any $x \in {}^*X$, as claimed.

The second property of Definition 13.1 is equivalent to the symmetry of $\approx_{\mathscr{U}}$. In fact, since $*$ is a superstructure monomorphism, we have ${}^*(U^{-1}) = ({}^*U)^{-1}$ for any $U \in \mathscr{U}$. Using our notation for relations, we thus find that $\approx_{\mathscr{U}}$ is symmetric if and only if $\operatorname{mon}(\mathscr{U}) = \operatorname{mon}(\mathscr{U})^{-1}$. Note that $\mathscr{U}_0 := \{U^{-1} : U \in \mathscr{U}\}$ is a filter with $\operatorname{mon}(\mathscr{U}_0) = \operatorname{mon}(\mathscr{U})^{-1}$. Hence, $\approx_{\mathscr{U}}$ is symmetric if and only if $\operatorname{mon}(\mathscr{U}_0) = \operatorname{mon}(\mathscr{U})$ which in view of Theorem 12.6 is equivalent to $\mathscr{U} = \mathscr{U}_0$ and thus to the second property of Definition 13.1.

The last property of Definition 13.1 is equivalent to the transitivity of $\approx_{\mathscr{U}}$. In fact, let this property be satisfied. If $x \approx_{\mathscr{U}} y$ and $y \approx_{\mathscr{U}} z$, we have for any $U \in \mathscr{U}$ that $(x,y),(y,z) \in {}^*U$. Choose some $V \in \mathscr{U}$ with $V^2 \subseteq U$. Then $({}^*V)^2 = {}^*(V^2) \subseteq {}^*U$, and so $(x,y),(y,z) \in {}^*V$ implies $(x,z) \in {}^*U$. Hence, $x \approx_{\mathscr{U}} z$, and so $\approx_{\mathscr{U}}$ is transitive.

Conversely, let $\approx_{\mathscr{U}}$ be transitive. By Theorem 12.6, there is some $V \in {}^*\mathscr{U}$ with $V \subseteq \operatorname{mon}(\mathscr{U})$. Then we have for any $(x,y) \in V$ that $x \approx_{\mathscr{U}} y$. Since $\approx_{\mathscr{U}}$ is transitive, we have for any $(x,z) \in V^2$ that $x \approx_{\mathscr{U}} z$ and thus $V^2 \in \operatorname{mon}(\mathscr{U})$. Hence, given $U \in \mathscr{U}$, we have $V^2 \in {}^*U$, i.e. the sentence

$$\exists \underline{v} \in {}^*\mathscr{U} : \underline{v}^2 \subseteq {}^*U$$

is true. The inverse form of the transfer principle implies that there is some $V \in \mathscr{U}$ with $V^2 \subseteq U$. □

Proposition 13.6. *If the uniform structure on X is induced by a family D of pseudometrics, then*

$$x \approx_{\mathscr{U}} y \iff {}^*d(x,y) \approx 0 \text{ for any } d \in D.$$

In particular, if X is a pseudometric space, we have

$$x \approx_{\mathscr{U}} y \iff {}^*d(x,y) \approx 0,$$

and for $X = \mathbb{R}$, we have

$$x \approx_{\mathscr{U}} y \iff x \approx y.$$

Proof. By Example 13.4, the sets of the form

$$B_{\varepsilon,d} := \{(\underline{x},\underline{y}) \in X \times X : d(\underline{x},\underline{y}) < \varepsilon\} \qquad (\varepsilon \in \mathbb{R}_+,\ d \in D)$$

generate $\mathscr{U}$. Hence, Lemma 12.11 implies $\operatorname{mon}(\mathscr{U}) = \bigcap\{{}^*(B_{\varepsilon,d}) : \varepsilon \in \mathbb{R}_+, d \in D\}$. Note that, by the standard definition principle for relations,

$${}^*(B_{\varepsilon,d}) = \{(\underline{x},\underline{y}) \in {}^*X \times {}^*X : {}^*d(\underline{x},\underline{y}) < {}^*\varepsilon\}.$$

Hence, $(x,y) \in \operatorname{mon}(\mathscr{U})$ if and only if ${}^*d(x,y) < \varepsilon$ for any $\varepsilon \in \mathbb{R}_+$ and any $d \in D$. □

One might hope that $x \approx_{\mathscr{U}} y$ if and only if $x \approx_{\mathscr{O}} y$ with respect to the topology generated by the uniform structure. Unfortunately, for nonstandard points this may fail even in natural situations:

Example 13.7. Consider $X = \mathbb{N}$ with the canonical (metric) uniform structure. By Proposition 13.6, we have

$$n \approx_{\mathscr{U}} m \iff n = m.$$

On the other hand, for $h \in \mathbb{N}_\infty$, we find some $n \neq h$ with $n \approx_{\mathscr{O}} h$ by Theorem 12.27.

Nevertheless, for standard points the situation is good:

Proposition 13.8. *Let X be a uniform space. Then we have*

$$y \approx_{\mathscr{O}} {}^*x \iff y \approx_{\mathscr{U}} {}^*x \qquad (x \in X,\ y \in {}^*X).$$

Proof. Let $y \approx_{\mathscr{U}} {}^*x$. If $O \subseteq X$ is open with ${}^*x \in {}^*O$, then $x \in O$, and the definition of the topology implies that there is some $U \in \mathscr{U}$ with $O = U(x)$. Since $\approx_{\mathscr{U}}$ is symmetric, we have ${}^*x \approx_{\mathscr{U}} y$ and thus $({}^*x, y) \in {}^*U$, i.e. $y \in {}^*U({}^*x) = {}^*(U(x)) = {}^*O$. Hence, $y \approx_{\mathscr{O}} {}^*x$.

Conversely, if $y \approx_{\mathscr{O}} {}^*x$ and $U \in \mathscr{U}$ is given, then $U(x)$ is a neighborhood of x by Proposition 13.2, and so we find some open $O \subseteq X$ with $x \in O \subseteq U(x)$. Then ${}^*x \in {}^*O$ implies in view of $y \approx_{\mathscr{O}} {}^*x$ that $y \in {}^*O \subseteq {}^*(U(x)) = {}^*U({}^*x)$, i.e. $({}^*x, y) \in {}^*U$. Hence, $({}^*x, y) \in \operatorname{mon}(\mathscr{U})$ which means ${}^*x \approx_{\mathscr{U}} y$ (or, equivalently, $y \approx_{\mathscr{U}} {}^*x$). □

Proposition 13.8 explains why $y \approx_{\mathscr{O}} x$ is in literature sometimes defined only for the case that x is a standard point.

In uniform spaces, it makes sense to speak of *Cauchy sequences*:

Definition 13.9. A sequence $x_n \in X$ is a *Cauchy sequence* if for each $U \in \mathscr{U}$ there is some n_0 such that $(x_n, x_m) \in U$ for $n, m \geq n_0$. A filter $\mathscr{F}$ over X is a *Cauchy filter* if for each $U \in \mathscr{U}$ there is some $F \in \mathscr{F}$ with $F \times F \subseteq U$.

Proposition 13.10. *A sequence $x_n \in X$ is a Cauchy sequence if and only if the filter $\mathscr{F}$ generated by the sets $F_n := \{x_n, x_{n+1}, x_{n+2}, \ldots\}$ $(n = 1, 2, \ldots)$ is a Cauchy filter.*

Proof. If x_n is a Cauchy sequence and $U \in \mathscr{U}$, then there is some n_0 with $F_{n_0} \in U$. Conversely, if $\mathscr{F}$ is a Cauchy filter and $U \in \mathscr{U}$, then there are $n_1, \ldots, n_k$ such that $F := F_{n_1} \cap \cdots \cap F_{n_k}$ satisfies $F \times F \subseteq U$. The latter means $(x_n, x_m) \in U$ for $n, m \geq \max\{n_1, \ldots, n_k\}$. □

Theorem 13.11. *A filter $\mathscr{F}$ is a Cauchy filter if and only if $x, y \in \operatorname{mon}(\mathscr{F})$ implies $x \approx_{\mathscr{U}} y$.*

Proof. If $\mathscr{F}$ is a Cauchy filter and $x, y \in \operatorname{mon}(\mathscr{F})$, then we have for any $U \in \mathscr{U}$ that there is some $F \in \mathscr{F}$ with $F \times F \subseteq U$, and so $(x, y) \in {}^*F \times {}^*F \subseteq {}^*U$. Hence, $x \approx_{\mathscr{U}} y$.

Conversely, let $x \approx_{\mathscr{U}} y$ for any $x, y \in \operatorname{mon}(\mathscr{F})$. By Theorem 12.6, there is some $F \in {}^*\mathscr{F}$ with $F \subseteq \operatorname{mon}(\mathscr{F})$. Given $U \in \mathscr{U}$, we have for any $x, y \in F$ that $(x, y) \in {}^*U$, i.e. $F \times F \subseteq {}^*U$. Hence,

$$\exists \underline{x} \in {}^*\mathscr{F} : \underline{x} \times \underline{x} \subseteq {}^*U.$$

The inverse form of the transfer principle implies that there is some $F \in \mathscr{F}$ with $F \times F \subseteq U$, and so $\mathscr{F}$ is a Cauchy filter. □

Definition 13.12. A uniform space is called *complete*, if each Cauchy filter converges.

Propositions 12.46 and 13.10 together imply:

Corollary 13.13. *In a complete space any Cauchy sequence converges.*

Proof. If x_n is a Cauchy sequence, then the filter $\mathscr{F}$ generated by the sets $F_n := \{x_n, x_{n+1}, \ldots\}$ is a Cauchy filter by Proposition 13.10 and thus convergent to some x. Proposition 12.46 implies $x_n \to x$. □

We point out that the converse to Corollary 13.13 does not hold, in general. But the converse is true in (pseudo)metric spaces; although it is rather easy to prove this fact by standard methods, one can also give a nonstandard proof (Exercise 68).

It turns out that completeness of a uniform space X is related to another important notion:

Definition 13.14. A point $y \in {}^*X$ is called a *pre-nearstandard point* if for any $U \in \mathscr{U}$ there is some $x \in X$ with $({}^*x, y) \in {}^*U$. We write $\operatorname{pns}({}^*X)$ for the set of pre-nearstandard points.

Since $U \in \mathscr{U}$ if and only if $U^{-1} \in \mathscr{U}$ and since $({}^*U)^{-1} = {}^*(U^{-1})$, it is equivalent to require that for any $y \in {}^*X$ and any $U \in \mathscr{U}$ there is some $x \in X$ with $(y, {}^*x) \in {}^*U$.

Exercise 67. Prove that in a (pseudo)metric space X a point $y \in {}^*X$ is pre-nearstandard if and only if for each $\varepsilon \in \mathbb{R}_+$ there is some $x \in {}^*X$ with ${}^*d({}^*x, y) < {}^*\varepsilon$.

Lemma 13.15. *If $\mathscr{F}$ is a Cauchy filter, then* $\operatorname{mon}(\mathscr{F}) \subseteq \operatorname{pns}({}^*X)$.

Proof. Let $y \in \operatorname{mon}(\mathscr{F})$. Given $U \in \mathscr{U}$, choose some $F \in \mathscr{F}$ with $F \times F \subseteq U$. Fix some $x \in F$. Then

$$({}^*x, y) \in {}^*F \times {}^*F = {}^*(F \times F) \subseteq {}^*U. \qquad \square$$

Theorem 13.16. $\mathrm{ns}({}^*X) \subseteq \mathrm{pns}({}^*X)$, *i.e. each nearstandard point is a pre-nearstandard point. Moreover, we have equality if and only if X is complete.*

Proof. If $y \in \mathrm{ns}({}^*X)$, then there is some $x \in X$ with $y \approx_{\mathscr{O}} {}^*x$ which by Proposition 13.8 implies $y \approx_{\mathscr{U}} {}^*x$, i.e. $({}^*x, y) \in {}^*U$ for any $U \in \mathscr{U}$. Hence, $y \in \mathrm{pns}({}^*X)$.

Now, let X be complete, and $y \in \mathrm{pns}({}^*X)$. Let $\mathscr{B}$ be the system of all sets of the form $U(x)$ where $U \in \mathscr{U}$ and $({}^*x, y) \in {}^*U$, i.e. $y \in {}^*U({}^*x)$. For $U_1(x_1), \dots, U_n(x_n) \in \mathscr{B}$, the set ${}^*U_1({}^*x_1) \cap \dots \cap {}^*U_n({}^*x_n)$ is not empty (because it contains y), and so $U_1(x_1) \cap \dots \cap U_n(x_n)$ is not empty. Hence, $\mathscr{B}$ has the finite intersection property and thus generates some filter $\mathscr{F}$. We claim that $\mathscr{F}$ is a Cauchy filter: Given $U \in \mathscr{U}$, choose some $V \in \mathscr{U}$ with $V^2 \in \mathscr{U}$. Then $U_0 := V \cap V^{-1} \in \mathscr{U}$. Since $y \in \mathrm{pns}({}^*X)$, we find some $x \in X$ with $({}^*x, y) \in {}^*U_0$. Then

$$U_0(x) \times U_0(x) \subseteq V^{-1}(x) \times V(x) = \{(a, b) : (a, x), (x, b) \in V\} \subseteq V^2 \subseteq U.$$

Since $U_0(x) \in \mathscr{F}$, we may conclude that $\mathscr{F}$ is a Cauchy filter. Since X is complete, we have $\mathscr{F} \to x$ for some $x \in X$, i.e. $\mathrm{mon}(\mathscr{F}) \subseteq \mathrm{mon}(x)$ (Proposition 12.47). Lemma 12.11 implies $\mathrm{mon}(\mathscr{F}) = \bigcap^{\sigma} \mathscr{B}$; since we have for any $B = U(x) \in \mathscr{B}$ that $y \in {}^*U({}^*x) = {}^*B$, it follows that $y \in \mathrm{mon}(\mathscr{F}) \subseteq \mathrm{mon}(x)$, i.e. $y \approx {}^*x$.

Conversely, let $\mathrm{pns}({}^*X) = \mathrm{ns}({}^*X)$, and let $\mathscr{F}$ be a Cauchy filter. Choose some $y \in \mathrm{mon}(\mathscr{F})$. Lemma 13.15 implies $y \in \mathrm{pns}({}^*X) = \mathrm{ns}({}^*X)$. Hence, there is some $x \in X$ with $y \approx_{\mathscr{U}} {}^*x$. We claim that $\mathscr{F} \to x$: If U_0 is a neighborhood of x, Proposition 13.2 implies that there is some $U \in \mathscr{U}$ with $U_0 = U(x)$. For any $z \in \mathrm{mon}(\mathscr{F})$, Theorem 13.11 implies $z \approx_{\mathscr{U}} y \approx_{\mathscr{U}} {}^*x$, and so $z \approx_{\mathscr{U}} {}^*x$. Hence, $({}^*x, z) \in {}^*U$, i.e. $z \in {}^*U_0$. This proves $\mathrm{mon}(\mathscr{F}) \subseteq \mathrm{mon}(\mathscr{U}(x))$ which means $\mathscr{F} \to x$ by Proposition 12.47. Thus, X is complete. □

Exercise 68. Prove by applying the nonstandard Theorem 13.16 that a (pseudo)metric space X is complete if and only if each Cauchy sequence converges.

If $Y \subseteq X$, we equip Y with the uniform structure

$$\mathscr{U}_Y := \{U \cap (Y \times Y) : U \in \mathscr{U}\}$$

which we call the *inherited uniform structure*.

Exercise 69. Prove that $\mathscr{U}_Y$ is indeed a uniform structure on Y.

The definition immediately implies:

Proposition 13.17. *Let $Y \subseteq X$. Then the neighborhoods of $y \in Y$ with respect to the induced uniform structure $\mathscr{U}_Y$ are precisely the sets of the form $U \cap Y$ where $U \subseteq X$ is a neighborhood of y with respect to $\mathscr{U}$.* □

If X is a topological space and $Y \subseteq X$, one can also define an *inherited topology* on Y: The open sets in Y with respect to this topology are by definition precisely the sets of the form $O \cap Y$ where $O \subseteq X$ is open. Fortunately, Proposition 13.17 implies that there cannot arise any confusion if X is a uniform space:

Corollary 13.18. *Let $Y \subseteq X$. Then the inherited topology on Y is induced by the inherited uniform structure $\mathscr{U}_Y$.* □

Proposition 13.19. *Let $Y \subseteq X$. For $x, y \in {}^*Y$, we have $x \approx_{\mathscr{U}} y$ if and only if $x \approx_{\mathscr{U}_Y} y$.*

Proof. We have $x \approx_{\mathscr{U}} y$ if and only if $(x, y) \in {}^*U$ for any $U \in \mathscr{U}$, i.e. if and only if $(x, y) \in {}^*U \cap ({}^*Y \times {}^*Y) = {}^*(U \cap (Y \times Y))$ for any $U \in \mathscr{U}$. This means $(x, y) \in {}^*V$ for any $V \in \mathscr{U}_Y$, i.e. $x \approx_{\mathscr{U}_Y} y$. □

Exercise 70. Prove by nonstandard methods that a closed subset of a complete uniform space with the inherited uniform structure is complete.

Definition 13.20. A subset A of a uniform space X is called *precompact*, if for each $U \in \mathscr{U}$ there are finitely many $x_1, \ldots, x_n \in X$ with $A \subseteq U(x_1) \cup \cdots \cup U(x_n)$.

Example 13.21. A subset A of a (pseudo)metric space X is precompact if and only if for each $\varepsilon \in \mathbb{R}_+$ there is a finite ε-net in X, i.e. there are finitely many $x_1, \ldots, x_n \in X$ such that the balls with center x_i and radius ε cover X.

Theorem 13.22. *A subset A of uniform space X is precompact if and only if any point of ${}^*A$ is pre-nearstandard, i.e. if and only if ${}^*A \subseteq \operatorname{pns}({}^*X)$.*

Proof. Let A be precompact, and $y \in {}^*A$. For any $U \in \mathscr{U}$, there are finitely many $x_1, \ldots, x_n \in X$ with $y \in {}^*A \subseteq {}^*((U(x_1) \cup \cdots \cup U(x_n)) = {}^*U(x_1) \cup \cdots \cup {}^*U(x_n)$. Hence, there is some x_k with $(y, {}^*x_k) \in {}^*U$, and so $y \in \operatorname{pns}({}^*X)$.

Conversely, if ${}^*A \subseteq \operatorname{pns}({}^*X)$ and $U \in \mathscr{U}$, put $\mathscr{A} := \{U(x) : x \in X\}$. For any $y \in {}^*A \subseteq \operatorname{pns}({}^*X)$ there is some $x \in X$ with $y \in {}^*U({}^*x) = {}^*(U(x))$, i.e. ${}^*A \subseteq \bigcup^\sigma \mathscr{A}$. By Exercise 44, there is a finite $\mathscr{A}_0 \subseteq \mathscr{A}$ with $A \subseteq \bigcup \mathscr{A}_0$. But this means that A is precompact. □

As an application, we give an easy nonstandard proof of an important standard result:

Corollary 13.23. *A uniform space X is compact if and only if it is precompact and complete.*

Proof. Theorem 13.16 and 13.22 imply that X is complete and precompact if and only if $\operatorname{ns}({}^*X) = \operatorname{pns}({}^*X) = {}^*X$. Since $\operatorname{ns}({}^*X) \subseteq \operatorname{pns}({}^*X) \subseteq {}^*X$ (Theorem 13.16), this is the case if and only if ${}^*X = \operatorname{ns}({}^*X)$. By Theorem 12.39, this means that ${}^*X$ is compact. □

Exercise 71. Let $*$ be a compact $\mathcal{P}(X)$-enlargement. Prove that X is precompact if and only if $\operatorname{pns}({}^*X)$ is internal.

Hint: ${}^\sigma X \subseteq \operatorname{pns}({}^*X)$.

13.2 Nonstandard Hulls

In connection with nonstandard analysis, uniform structures have an important advantage over general topologies: They allow us to define a so-called *nonstandard hull*. To define this, we first have to consider a uniform structure on ${}^*X$:

Proposition 13.24. *The system ${}^\sigma\mathscr{U}$ generates a filter $\mathscr{U}_*$ which is a uniform structure over ${}^*X$. We have $V \in \mathscr{U}_*$ if and only if there is some $U \in \mathscr{U}$ with ${}^*U \subseteq V \subseteq {}^*X \times {}^*X$.*

Proof. Since $\mathscr{U}$ has the finite intersection property, also ${}^\sigma\mathscr{U}$ has the finite intersection property, and so the set $\mathscr{U}_*$ is a filter over ${}^*(X \times X) = {}^*X \times {}^*X$. By definition of a generated filter, we have $V \in \mathscr{U}_*$ if and only if there are finitely many $U_1, \dots, U_n \in \mathscr{U}_*$ with $U_1 \cap \dots \cap U_n \subseteq V \subseteq {}^*X \times {}^*X$. Since $\mathscr{U}$ is a filter, we have $U_1 \cap \dots \cap U_n \in \mathscr{U}$, and so $\mathscr{U}_*$ can be described as in the claim.

Since $\Delta := \{(\underline{x}, \underline{y}) \in X \times X : \underline{x} = \underline{y}\} \in \mathscr{U}$, we have ${}^*\Delta \in \mathscr{U}_*$. The standard definition principle for relations implies

$$ {}^*\Delta = \{(\underline{x}, \underline{y}) \in {}^*X \times {}^*X : \underline{x} = \underline{y}\}, $$

and so the first property of Definition 13.1 is verified for $\mathscr{U}_*$. If $V \in \mathscr{U}_*$, choose some $U \in \mathscr{U}$ with ${}^*U \subseteq V$. Since $U^{-1} \in \mathscr{U}$ and ${}^*(U^{-1}) = ({}^*U)^{-1} \subseteq V^{-1}$, we have $V^{-1} \in \mathscr{U}_*$. Moreover, there is some $W \in \mathscr{U}$ with $W^2 \subseteq U$. Then ${}^*W \in \mathscr{U}_*$ satisfies $({}^*W)^2 = {}^*(W^2) \subseteq {}^*U \subseteq V$, and so $\mathscr{U}_*$ is a uniform structure. □

Definition 13.25. If X is a uniform space, we equip ${}^*X$ with the uniform structure $\mathscr{U}_*$ of Proposition 13.24 and the corresponding topology.

We point out that the uniform structure on ${}^*X$ is *not* internal (except for trivial cases). In particular, the uniform structure $\mathscr{U}_*$ usually differs from ${}^*\mathscr{U}$.

Theorem 13.26. *Let $*$ be $\mathcal{P}(X)$-saturated. Then ${}^*X$ is complete.*

Proof. Let $\mathscr{F}$ be some Cauchy filter over ${}^*X$. For each $U \in \mathscr{U}$, there is some $F_U \in \mathscr{F}$ with $F_U \times F_U \subseteq {}^*U$ and some $x_U \in F_U$. Consider the system $\mathscr{A} := \{{}^*U(x_U) : U \in \mathscr{U}\}$ (axiom of choice!). We claim that $\mathscr{A}$ has the finite intersection property: If $U_1, \dots, U_n \in \mathscr{U}$, then $F_{U_1} \cap \dots \cap F_{U_n}$ contains some element x. For each $k = 1, \dots, n$, we have $(x_{U_k}, x) \in F_{U_k} \times F_{U_k} \subseteq {}^*U_k$, and so $x \in {}^*U_k(x_{U_k})$. Moreover, $\mathscr{A}$ has at most the cardinality of $\mathcal{P}(X)$ (if X is infinite). Since $*$ is $\mathcal{P}(X)$-saturated, we thus find some $x \in \bigcap \mathscr{A}$, i.e. $x \in {}^*U(x_U)$ for any $U \in \mathscr{U}$.

We claim that $\mathscr{F} \to x$: Indeed, by Proposition 13.2, any neighborhood of x can be written in the form $W(x)$ where $W \in \mathscr{U}_*$. We have to prove that $W(x) \in \mathscr{F}$. There is some $U \in \mathscr{U}$ with ${}^*U \subseteq W$ and some $V \in \mathscr{U}$ with $V^2 \subseteq U$. We may assume that $V = V^{-1}$ (otherwise replace V by $V \cap V^{-1}$). Since $x \in {}^*V(x_V)$, we have $x_V \in ({}^*V)^{-1}(x) = {}^*(V^{-1})(x) = {}^*V(x)$. In particular, ${}^*V(x_V) \subseteq ({}^*V)^2(x)$. In view of $F_V \times F_V \subseteq {}^*V$ and $x_V \in F_V$, we obtain $F_V \subseteq {}^*V(x_V) \subseteq ({}^*V)^2(x) = {}^*(V^2)(x) \subseteq {}^*U(x) \subseteq W(x)$, and so $W(x) \in \mathscr{F}$, as claimed. □

Exercise 72. Assume that $*$ is comprehensive and a compact $\mathcal{P}(X)$-enlargement. Prove that ${}^*X$ is complete.

If X is a (pseudo)metric space, then a weaker saturation property suffices for the completeness. At this point, we recall that *any* nonstandard ultrapower model is comprehensive and thus $\mathbb{N}$-saturated.

In the definition of a (pseudo)metric space, one may formally allow that the metric d attains the value ∞. If the reader does not like this convention, an alternative is to equivalently replace d by

$$d_0(x, y) = \min\{d(x, y), 1\}$$

which is a (pseudo)metric that does not attain the value ∞ and generates the same uniform structure as d.

Theorem 13.27. *Let the uniform structure on X be induced by a family D of pseudometrics. Then the uniform structure $\mathscr{U}_*$ on ${}^*X$ is induced by the family of pseudometrics*

$$d_*(x, y) := \begin{cases} \operatorname{st}({}^*d(x, y)) & \text{if } {}^*d(x, y) \text{ is finite,} \\ \infty & \text{if } {}^*d(x, y) \text{ is infinite.} \end{cases} \qquad (d \in D).$$

Moreover, if $*$ is $(\mathcal{P}(D) \times \mathbb{N})$-saturated, then ${}^*X$ is complete.

Proof. Given $d \in D$, the transfer principle implies $\forall \underline{x} \in {}^*X : {}^*d(\underline{x}, \underline{x}) = 0$, and $\forall \underline{x}, \underline{y}, \underline{z} \in {}^*X : {}^*d(x, y) \leq {}^*d(x, z) + {}^*d(z, y)$. Hence, $d_*(x, x) = 0$, and the triangle inequality for d_* follows together with the additivity and monotonicity of st. Thus, d_* is a pseudometric.

By definition, $U \in \mathscr{U}$ if and only if U contains a set of the form

$$U_0 := \{(\underline{x}, \underline{y}) \in X \times X \mid d_1(\underline{x}, \underline{y}) < \varepsilon \wedge \cdots \wedge d_n(\underline{x}, \underline{y}) < \varepsilon\}$$

with $\varepsilon \in \mathbb{R}_+$ and $d_1, \ldots, d_n \in D$. We have $V \in \mathscr{U}_*$ if and only if there is some $U \in \mathscr{U}$ with ${}^*U \subseteq V$, i.e. if and only if there is some U_0 of the above form such that (using the standard definition principle for relations)

$$V \supseteq {}^*U_0 = \{(\underline{x}, \underline{y}) \in {}^*X \times {}^*X \mid d_1(\underline{x}, \underline{y}) < {}^*\varepsilon \wedge \cdots \wedge {}^*d_n(\underline{x}, \underline{y}) < {}^*\varepsilon\}.$$

In view of the monotonicity of st, this means that there are $d_1, \dots, d_n \in D$ and $\varepsilon \in \mathbb{R}_+$ with

$$V \supseteq \{(x,y) : (d_1)_*(x,y) \leq \varepsilon \wedge \cdots \wedge (d_n)_*(x,y) \leq \varepsilon\}.$$

But this means that the uniform structure $\mathscr{U}_*$ is induced by the family of pseudometrics d_* $(d \in D)$.

Let $\mathscr{F}$ be a Cauchy filter over ${}^*X$. Then we find for each $n \in \mathbb{N}$ and each finite $D_0 \subseteq D$ some $F_{D_0,n} \in \mathscr{F}$ such that

$$F_{D_0,n} \times F_{D_0,n} \subseteq \{(\underline{x}, \underline{y}) \in {}^*X \times {}^*X \mid {}^*d(\underline{x}, \underline{y}) < 1/{}^*n \text{ for all } d \in D_0\} =: U_{D_0,n}.$$

Choose some $x_{D_0,n} \in F_{D_0,n}$, and consider the system $\mathscr{A}$ of internal sets $U_{D_0,n}(x_{D_0,n})$ (axiom of choice!). Then $\mathscr{A}$ has the finite intersection property. Indeed, if $n_1, \dots, n_N \in \mathbb{N}$ and $D_1, \dots, D_N$ are finite subset of D, then $F_{D_1,n_1} \cap \cdots \cap F_{D_N,n_N}$ contains some element x. Then $x \in U_{D_1,n_1}(x_{D_1,n_1}) \cap \cdots \cap U_{D_N,n_N}(x_{n_N})$, because for $n = n_k$ and $D_0 = D_k$ $(k = 1, \dots, n)$, we have $(x, x_{D_0,n}) \in F_{D_0,n} \times F_{D_0,n} \subseteq U_{D_0,n}$. Since $*$ is $(\mathcal{P}(D) \times \mathbb{N})$-saturated, we may conclude that $\bigcap \mathscr{A}$ contains some element $x \in {}^*X$.

We claim that $F_n \to x$: We have to prove that any neighborhood V of x belongs to $\mathscr{F}$. We have $V = U(x)$ for some $U \in \mathscr{U}_*$, and $U \supseteq U_{D_0,n}$ for some $n \in \mathbb{N}$ and some finite $D_0 \subseteq D$. Hence, $V \supseteq U_{D_0,n}(x)$. Since $x \in U_{D_0,2n}(x_{D_0,2n})$, we have ${}^*d(x, x_{2n}) < 1/(2{}^*n)$ for each $d \in D_0$, and so by the triangle inequality of ${}^*d$ that $V \supseteq U_{D_0,2n}(x_{2n})$. Hence, $F_{D_0,2n} \times F_{D_0,2n} \subseteq U_{D_0,2n}$ and $x_{D_0,2n} \in F_{D_0,2n}$ imply $F_{D_0,2n} \subseteq U_{D_0,2n}(x_{2n}) \subseteq V$, and so $V \in \mathscr{F}$, as claimed. □

We note that for (pseudo)metric spaces, we needed only a countable version of the axiom of choice in the previous proof.

Even if X is a metric space, the space ${}^*X$ is usually not a Hausdorff space: If $x \approx_{\mathscr{U}} y$, then x and y do not have disjoint neighborhoods. To get a Hausdorff space, one identifies such elements: One may do this, since $\approx_{\mathscr{U}}$ is an equivalence relation. Thus, we put

$$\widetilde{X} := \{[x] : x \in {}^*X\},$$

where $[x]$ denotes the equivalence class of x, i.e. the set of all $y \in {}^*X$ with $y \approx_{\mathscr{U}} x$. Let $\widetilde{\mathscr{U}}$ be the system of all sets of the form

$$\widetilde{U} := \{([x],[y]) : (x,y) \in U\},$$

where U is an element of the uniform structure $\mathscr{U}_*$ of ${}^*X$ from Definition 13.25.

Some care is necessary: The relation $([x],[y]) \in \widetilde{U}$ does not imply that $(x,y) \in U$. However, a weakening is true if $U \in {}^\sigma\mathscr{U}$:

Lemma 13.28. *If $U = {}^*V$ for some $V \in \mathscr{U}$, then the relation $([x],[y]) \in \widetilde{U}$ implies $(x,y) \in U^3$.*

Proof. Since $([x],[y]) \in \widetilde{U}$, there are elements $x_0, y_0 \in {}^*X$ with $x_0 \approx_{\mathscr{U}} x$, $y_0 \approx_{\mathscr{U}} x$ and $(x_0, y_0) \in U = {}^*V$. Since $x_0 \approx_{\mathscr{U}} x$ and $y_0 \approx_{\mathscr{U}} y$, we have $(x, x_0) \in {}^*V$ and $(y_0, y) \in {}^*V$, and so $(x,y) \in ({}^*V)^3$. □

Theorem 13.29. *$\widetilde{\mathscr{U}}$ is a uniform structure on $\widetilde{X}$. For $V \subseteq \widetilde{X} \times \widetilde{X}$ we have $V \in \widetilde{\mathscr{U}}$ if and only if there is some $U \in \mathscr{U}$ with $V \supseteq \widetilde{{}^*U}$.*

The space $\widetilde{\mathscr{U}}$ is always a Hausdorff space. Moreover, if X is Hausdorff, then the embedding $X \hookrightarrow \widetilde{X}$ defined by $x \mapsto [{}^*x]$ is one-to-one. In the sense of this embedding, we have

$$\mathscr{U} = \{\widetilde{U} \cap (X \times X) : \widetilde{U} \in \widetilde{\mathscr{U}}\}, \tag{13.1}$$

i.e. the uniform structure of X is inherited by the uniform structure of $\widetilde{X}$. Moreover, the closure of X in $\widetilde{X}$ is the set

$$\overline{X} := \{[x] : x \in \operatorname{pns}({}^*X)\}.$$

Proof. It follows by definition that $\widetilde{\mathscr{U}}$ consists of subsets of $\widetilde{X} \times \widetilde{X}$ which contain $\Delta := \{([x],[x]) : [x] \in \widetilde{X}\}$. We prove first that $\widetilde{\mathscr{U}}$ is even a filter:

If $\widetilde{U} \in \widetilde{\mathscr{U}}$ and $V \supseteq \widetilde{U}$, let $V_0 := \{(x,y) : ([x],[y]) \in V\}$. Then $V_0 \supseteq U$, hence $V_0 \in \mathscr{U}$, and so $V = \widetilde{V_0} \in \widetilde{\mathscr{U}}$.

Let $\widetilde{U}_1, \widetilde{U}_2 \in \widetilde{\mathscr{U}}$ be given, i.e. $U_1, U_2 \in \mathscr{U}_*$. Then $U := U_1 \cap U_2 \in \mathscr{U}_*$, and so $\widetilde{U} \in \widetilde{\mathscr{U}}$. Since

$$\begin{aligned}\widetilde{U} &= \{([x],[y]) : x,y \in U_1 \cap U_2\} \subseteq \{([x],[y]) : x,y \in U_1\} \cap \{([x],[y]) : x,y \in U_2\} \\ &= \widetilde{U}_1 \cap \widetilde{U}_2,\end{aligned}$$

we have $\widetilde{U}_1 \cap \widetilde{U}_2 \in \widetilde{\mathscr{U}}$, as desired.

Now we prove that $\widetilde{\mathscr{U}}$ is even a uniform structure: If $\widetilde{U} \in \widetilde{\mathscr{U}}$, then $(\widetilde{U})^{-1} = \{([x],[y]) : (y,x) \in U\} = \widetilde{U^{-1}} \in \widetilde{\mathscr{U}}$. Moreover, there is some $V \in \mathscr{U}_*$ with $V^6 \subseteq U$ and some $V_0 \in \mathscr{U}$ with $V \supseteq {}^*V_0$. Then $W := \widetilde{{}^*V_0} \in \widetilde{\mathscr{U}}$ satisfies $W^2 \subseteq \widetilde{U}$. Indeed, if $([x],[y]) \in W^2$, we find some $z \in {}^*X$ with $([x],[z]), ([z],[y]) \in W$. By Lemma 13.28, we have $(x,z), (z,y) \in ({}^*V_0)^3$, and so $(x,y) \in ({}^*V_0)^6 \subseteq V^6 \subseteq U$ which implies $([x],[y]) \in \widetilde{\mathscr{U}}$, as desired.

It is now easily seen that $\widetilde{\mathscr{U}}$ can be described as in the claim: If $V \supseteq \widetilde{{}^*U}$ for some $U \in \mathscr{U}$, then $\widetilde{{}^*U} \in \widetilde{\mathscr{U}}$, and so $V \in \widetilde{\mathscr{U}}$, since $\widetilde{\mathscr{U}}$ is a filter. Conversely, if $V \in \widetilde{\mathscr{U}}$, then $V = \widetilde{W}$ for some $W \in \mathscr{U}_*$. There is some $U \in \mathscr{U}$ with $W \supseteq {}^*U$. Then $V = \widetilde{W} \supseteq \widetilde{{}^*U}$.

To prove that $\widetilde{X}$ is Hausdorff, we apply Proposition 13.2: Let $[x],[y] \in \widetilde{X}$ satisfy $([x],[y]) \in \widetilde{U}$ for any $\widetilde{U} \in \widetilde{\mathscr{U}}$. For any $U \in \mathscr{U}$, we find some $V \in \mathscr{U}$ with

$V^3 \subseteq U$. Since $([x],[y]) \in \widetilde{{}^*V}$, Lemma 13.28 implies $(x,y) \in ({}^*V)^3 = {}^*(V^3) \subseteq {}^*U$. Hence, $(x,y) \in {}^*U$ for any $U \in \mathscr{U}$ which means means $x \approx_{\mathscr{U}} y$, and so $[x]=[y]$, as desired.

Now, let X be Hausdorff. If $x,y \in X$ satisfy $[{}^*x]=[{}^*y]$, we have ${}^*x \approx_{\mathscr{U}} {}^*y$, and so ${}^*x \approx_{\mathscr{O}} {}^*y$ which implies $x = \mathrm{st}({}^*y) = y$, since st is a function by Proposition 12.30. Thus, the map $x \mapsto [{}^*x]$ is one-to-one.

Now we prove (13.1): If $\widetilde{U} \in \widetilde{\mathscr{U}}$, there is some $V \in \mathscr{U}$ with $U \supseteq {}^*V$. Since $U \supseteq {}^\sigma V$, we have in the sense of the embedding

$$V = \{([{}^*x],[{}^*y]) : (x,y) \in V\} \subseteq \widetilde{U} \cap (X \times X),$$

and so $\widetilde{U} \cap (X \times X) \in \mathscr{U}$. Conversely, let $U \in \mathscr{U}$. There is some $V \in \mathscr{U}$ with $V^3 \subseteq U$. Then $\widetilde{{}^*V} \in \widetilde{\mathscr{U}}$, and $([x],[y]) \in \widetilde{{}^*V}$ implies $(x,y) \in ({}^*V)^3 = {}^*(V^3)$ by Lemma 13.28. In particular, $([{}^*x],[{}^*y]) \in \widetilde{{}^*V}$ implies $({}^*x,{}^*y) \in {}^*(V^3)$, and so $(x,y) \in V^3 \subseteq U$. In the sense of the embedding, this means

$$\widetilde{{}^*V} \cap (X \times X) \subseteq U.$$

We thus find some $W \supseteq {}^*V$ (and so $W \in \mathscr{U}_*$) with $\widetilde{W} \cap (X \times X) = U$, as desired.

If $y \in \mathrm{pns}({}^*X)$ and $\widetilde{V} \in \widetilde{\mathscr{U}}$, choose some $U \in \mathscr{U}$ with $V \supseteq {}^*U$. We find some $x \in X$ with $(y,{}^*x) \in {}^*U \subseteq V$. Hence, $([y],[{}^*x]) \in \widetilde{V}$, i.e. $[y] \in \widetilde{V}([{}^*x]) = \widetilde{V}(x)$ (in the sense of our identification). Thus, $[y]$ belongs to the closure of X. Conversely, if $[y]$ belongs to the closure of X and $U \in \mathscr{U}$ is given, choose some $V \in \mathscr{U}$ with $V^3 \subseteq U$. Since $[y]$ belongs to the closure of X, we find some element of X in the neighborhood $\widetilde{{}^*V}([y])$, i.e. there is some $x \in X$ with $([{}^*x],[y]) \in \widetilde{{}^*V}$. Lemma 13.28 implies $({}^*x,y) \in ({}^*V)^3 \subseteq {}^*U$. Hence, $y \in \mathrm{pns}({}^*X)$. □

The set $\widetilde{X}$ with the uniform structure $\widetilde{\mathscr{U}}$ is called the *nonstandard hull* of the uniform space X.

Theorem 13.30. *Let $*$ be $\mathcal{P}(X)$-saturated. Then $\widetilde{X}$ is complete.*

Proof. Let $\mathscr{F}$ be a Cauchy filter over $\widetilde{X}$. Let $\mathscr{F}_0$ be the system of all subsets of ${}^*X$ which contain an element of the form $\{x \in {}^*X : [x] \in F\}$ where $F \in \mathscr{F}$.

Then $\mathscr{F}_0$ is a filter: By definition, $\varnothing \notin \mathscr{F}_0$, and the relations $G_0 \in \mathscr{F}_0$ and $G_0 \subseteq G \subseteq {}^*X$ imply $G \in \mathscr{F}_0$. Moreover, for any $G_1, G_2 \in \mathscr{F}_0$, we find $F_1, F_2 \in \mathscr{F}$ with $G_i \supseteq \{x : [x] \in F_i\}$. Hence,

$$G_1 \cap G_2 \supseteq \{x : [x] \in F_1\} \cap \{x : [x] \in F_2\} \supseteq \{x : [x] \in F_1 \cap F_2\} \in \mathscr{F}_0,$$

and so $G_1 \cap G_2 \in \mathscr{F}_0$.

Moreover, $\mathscr{F}_0$ is a Cauchy filter: Given $U \in \mathscr{U}_*$, choose some $V \in \mathscr{U}_*$ with $V^3 \subseteq U$ and some $W \in \mathscr{U}$ with $V \supseteq {}^*W$. Since $\mathscr{F}$ is a Cauchy filter, we find some $F \in \mathscr{F}$ with $F \times F \subseteq \widetilde{{}^*W}$. Putting $F_0 := \{x : [x] \in F\} \in \mathscr{F}_0$, we have for any $(x, y) \in F_0 \times F_0$ that $([x], [y]) \subseteq \widetilde{{}^*W}$ which by Lemma 13.28 implies $(x, y) \in ({}^*W)^3 \subseteq U$. Hence, $F_0 \times F_0 \subseteq U$.

By Theorem 13.26, ${}^*X$ is complete, and so $\mathscr{F}_0 \to x$ for some $x \in X$. We claim that $\mathscr{F} \to [x]$. Given $\widetilde{U} \in \widetilde{\mathscr{U}}$, we have to prove that $\widetilde{U}([x]) \in \mathscr{F}$. Since $\mathscr{F}_0 \to x$, we have $U(x) \in \mathscr{F}_0$. By definition of $\mathscr{F}_0$, we find some $F \in \mathscr{F}$ with $\{y \in {}^*X : [y] \in F\} \subseteq U(x)$. Then

$$\begin{aligned}\widetilde{U}([x]) &= \{[y] : ([x], [y]) \in \widetilde{U}\} = \{[y] : (x, y) \in U\} \\ &= \{[y] : y \in U(x)\} \supseteq \{[y] : [y] \in F\} = F \in \mathscr{F},\end{aligned}$$

and so $\widetilde{U}([x]) \in \mathscr{F}$, as desired. □

As can be seen from the proof, the saturation property is only needed for the completeness of ${}^*X$. In particular, the same result holds if $*$ is comprehensive and a compact $\mathcal{P}(X)$-enlargement (Exercise 72). Moreover, if the uniform structure is induced by a family D of pseudometrics, it suffices that $*$ is $(\mathcal{P}(D) \times \mathbb{N})$-saturated (Theorem 13.27).

Corollary 13.31. *If X is a Hausdorff space, then X has a completion, i.e. there is a complete uniform Hausdorff space $\overline{X}$ such that $X \subseteq \overline{X}$ carries the inherited uniform structure and such that $\overline{X}$ is the closure of X.*

Proof. By Exercise 70, the closure $\overline{X}$ of X in the complete space $\widetilde{X}$ is complete. □

Proposition 13.32. *Let the uniform structure on X be induced by a family D of pseudometrics. Then the uniform structure on $\widetilde{X}$ is induced by the family of pseudometrics*

$$\widetilde{d}([x], [y]) = \begin{cases} \operatorname{st}({}^*d(x, y)) & \text{if } {}^*d(x, y) \text{ is finite,} \\ \infty & \text{if } {}^*d(x, y) \text{ is infinite.} \end{cases} \qquad (d \in D).$$

If X is a metric space, i.e. $D = \{d\}$, then $\widetilde{d}$ is a metric (which might assume the value ∞). Moreover, $\widetilde{X}$ is complete if $*$ is $(\mathcal{P}(D) \times \mathbb{N})$-saturated.

Proof. The last statement follows by the remarks following Theorem 13.30. To see that $\widetilde{d}$ is well-defined, let $[x] = [x_0]$ and $[y] = [y_0]$. Then $x \approx_{\mathscr{U}} x_0$ and $y \approx_{\mathscr{U}} y_0$, and by Proposition 13.6, this implies ${}^*d(x, x_0) \approx 0 \approx {}^*d(y, y_0)$ for any $d \in D$. By the inverse triangle inequality for ${}^*d$, we thus have $|{}^*d(x, y) - {}^*d(x_0, y_0)| \leq {}^*d(x, x_0) + {}^*d(y, y_0) \approx 0$, and so ${}^*d(x, y)$ and ${}^*d(x_0, y_0)$ are either both infinite or both finite with the same standard part.

If X is a metric space, then $\widetilde{X}$ is a pseudometric space. Since $\widetilde{X}$ is Hausdorff (Theorem 13.29), the pseudometric on $\widetilde{X}$ must even be a metric. □

Corollary 13.33. *The completion of a metric space is a metric space.* □

One should not make the mistake of thinking that the nonstandard hull $\widetilde{X}$ is essentially the completion of X: The nonstandard hull is much larger and more useful (even if X is complete). Consider, for example, $X = \mathbb{N}$. Then $^*X = {}^*\mathbb{N}$, and in the canonical way we have also $\widetilde{\mathbb{N}} = {}^*\mathbb{N}$ (Example 13.7). As another example consider $X = \mathbb{R}$: Then $^*X = {}^*\mathbb{R}$, but the nonstandard hull is not so easy to describe. However, from the previous consideration, we may conclude that $\widetilde{X} \supseteq {}^*\mathbb{N}$ (in a canonical way). In this example, it appears that the subspace of "finite" elements of $\widetilde{X}$ plays an important role, i.e. the subspace of all $[x]$ where $x \in {}^*X$ is "finite" in a certain sense. More generally, this "finiteness" plays a particular role in the context of topological vector spaces which we discuss now.

Unfortunately, there are two natural definitions of "finiteness" which differ in general [Hen72a, HM72]. However, for all "important" spaces, these definitions coincide: In particular, for the large class of so-called locally convex spaces (see the remarks preceding Proposition 14.14), and also for so-called quasinormed spaces (like ℓ_p or L_p with $0 < p \leq \infty$), the two definitions are identical.

Thus, instead of going into technical details, we present just one of these definitions in the context of topological vector spaces in the following. The reader who is interested in more details is referred to the original papers [Hen72a, HM72].

§14 Topological Vector Spaces

Definition 14.1. A linear space (=vector space) X over $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$ with a topology $\mathscr{O}$ is called a *topological vector space*, if the addition and multiplication (by scalars) are continuous operations, i.e. if the mappings $(x, y) \mapsto x + y$ and $(\lambda, x) \mapsto \lambda x$ are continuous as mappings from the topological spaces $X \times X$ resp. $\mathbb{K} \times X$ (with the product topology) into X.

Sometimes, the term *topological vector space* is used for those topological vector spaces which are additionally Hausdorff. If we assume that X is Hausdorff, we will state this explicitly.

Theorem 14.2. *Let X be a vector space endowed with some topology. Then X is a topological vector space if and only if the following three conditions hold:*

1. *The neighborhoods of $x \in X$ are precisely the sets U of the form $U = x + V$ where V is a neighborhood of 0.*

   In particular, the topology is translation invariant (and so the neighborhoods of 0 determine the topology): If O is open, then $x + O$ is a neighborhood for each of its elements and thus open.
2. *For each neighborhood U of 0 and each $\lambda \in \mathbb{K}$ there is some neighborhood V of 0 with $V + V \subseteq U$ and some neighborhood Λ of λ with $\Lambda V \subseteq U$.*

   In particular, if U is a neighborhood of 0, then $\lambda^{-1}U$ is a neighborhood of 0 for all $\lambda \neq 0$, and there is a neighborhood V of 0 with $V - V \subseteq U$.
3. *For each neighborhood U of 0 and each $x \in X$ there is some neighborhood Λ of 0 with $\Lambda x \subseteq U$.*

Proof. Put $A(x, y) := x + y$ and $M(\lambda, x) := \lambda x$. Assume first that A and M are continuous.

To prove 1., we show that for any $x_0, x_1 \in X$ and any neighborhood U of x_0 the set $U_0 := (x_1 - x_0) + U$ is a neighborhood of x_1: For the choice $x_0 := 0$ and $x_1 := x$, we obtain then that for any neighborhood U of 0 the set $U_0 = x + U$ is a neighborhood of x; for the choice $x_0 := x$ and $x_1 := 0$, we obtain conversely that any neighborhood U of x has the form $U = x + U_0$ for some neighborhood U_0 of 0.

Thus, let $x_0, x_1 \in X$, and U be a neighborhood of x_0. Since $A(x_1, x_0 - x_1) = x_0$, we find a neighborhood V of $(x_1, x_0 - x_1)$ with $A(V) \subseteq x_0$. By definition of the product topology, there are neighborhoods V_1 of x_1 and V_2 of $x_0 - x_1$ such that $V_1 \times V_2 \subseteq V$. Then $A(V_1 \times V_2) \subseteq U$; in particular, $V_1 + (x_0 - x_1) \subseteq U$. Hence, $U_0 = U - (x_0 - x_1)$ contains V_1 and thus is a neighborhood of x_1, as claimed.

For the proof of 2. and 3., let a neighborhood U of 0 and points $x \in X$ and $\lambda \in \mathbb{K}$ be given. We find a neighborhood V_0 of (λ, x) such that $M(V_0) \subseteq U$. By definition of the product topology this means that there are neighborhoods Λ of λ

and V_1 of x such that $\Lambda \times V_1 \subseteq V_0$. In particular, $\Lambda V_1 \subseteq M(V_0) \subseteq U$. This already proves 3. for the choice $\lambda := 0$.

To proceed with the proof of 2., consider the choice $x := 0$, and observe that V_1 is then a neighborhood of 0. In view of $A(0,0) = 0$, we find a neighborhood W of $(0,0)$ with $A(W) \subseteq U$. There are neighborhood W_1, W_2 of 0 with $W_1 \times W_2 \subseteq W$. Then the set $V := W_1 \cap W_2 \cap V_1$ is a neighborhood of 0. As observed above, we have $\Lambda V \subseteq \Lambda V_1 \subseteq U$. Moreover, $V + V \subseteq A(W) \subseteq U$. Thus, 2. holds.

Now let the conditions of the theorem be satisfied. To prove that A is continuous at (x, y), let V be a neighborhood of $z := x + y$. Then $U := V - z$ is a neighborhood of 0. Choose a neighborhood V of 0 with $V + V \subseteq U$. Then $U_1 := x + V$ and $U_2 := y + V$ are neighborhoods of x and y, and so $U_1 \times U_2$ is a neighborhood of (x, y) with $A(U_1 \times U_2) \subseteq x + y + U = V$.

To prove that M is continuous at (λ, x), let V be a neighborhood of $y := \lambda x$. Then $V - y$ is a neighborhood of 0, and so we find a neighborhood U of 0 with $U + U \subseteq V - y$. There is a neighborhood Λ of 0 with $\Lambda x \subseteq U$. Moreover, there are neighborhoods Λ_0 of λ and U_0 of 0 with $\Lambda_0 U_0 \subseteq U$. Since $\Lambda_1 := \Lambda_0 \cap (\lambda + \Lambda)$ is a neighborhood of λ, the set $W := \Lambda_1 \times (x + U_0)$ is a neighborhood of (λ, x). We have

$$M(W) \subseteq (\lambda + \Lambda)x + \Lambda_0 U_0 \subseteq y + U + U \subseteq V,$$

and so M is continuous at (λ, x). □

If X is a topological vector space, we let $\mathscr{U}$ denote the system of all sets $U \subseteq X \times X$ which satisfy $U \supseteq \{(x, y) : x - y \in O\}$ for some neighborhood $O \subseteq X$ of 0.

Theorem 14.3. *If X is a topological vector space, then $\mathscr{U}$ is a uniform structure which generates the topology $\mathscr{O}$.*

Proof. Let $\mathscr{O}_0$ denote the system of neighborhoods of 0. Given some $O \in \mathscr{O}_0$, we use the notation $U_O := \{(x, y) : x - y \in O\}$. Then $\mathscr{U}$ is the system of all sets $U \subseteq X \times X$ which contain some U_O.

Observe that $\Delta := \{(x, x) : x \in X\}$ is contained in each U_O and thus in each element of $\mathscr{U}$. Moreover, it follows from the definition that the relation $U \in \mathscr{U}$ and $V \supseteq U$ implies $V \in \mathscr{U}$. To see that $\mathscr{U}$ is a filter, let $U_1, U_2 \in \mathscr{U}$ be given. Then there are $O_1, O_2 \in \mathscr{O}_0$ with $U_i \supseteq U_{O_i}$. Then

$$U_1 \cap U_2 \supseteq U_{O_1} \cap U_{O_2} \supseteq U_{O_1 \cap O_2} \in \mathscr{U},$$

and so $U_1 \cap U_2 \in \mathscr{U}$, as desired.

To see that $\mathscr{U}$ is a uniform structure, let $U \in \mathscr{U}$ be given. Then $U \supseteq U_O$ for some $O \in \mathscr{O}_0$. By Theorem 14.2 2., there is some $O_0 \in \mathscr{O}_0$ with $-O_0 \subseteq O$ and

$O_0 + O_0 \subseteq O$. Then $U^{-1} \supseteq U_O^{-1} = U_{-O} \supseteq U_{O_0} \in \mathscr{U}$ implies $U^{-1} \in \mathscr{U}$. Moreover, for $V := U_{O_0}$, we have

$$V^2 = \{(x,y) \mid \exists z \in X : x - z, z - y \in O_0\}.$$

Since $x - z, z - y \in O_0$ implies $x - y = (x - z) + (z - y) \in O_0 + O_0 \subseteq O$, we find $V^2 \subseteq U_O$.

We prove now that $\mathscr{U}$ generates the topology. In view of Proposition 12.5, it suffices to consider neighborhoods: Given $x \in X$, we have to prove that the neighborhoods of x are precisely the sets of the form $U(x)$ with $U \in \mathscr{U}$ (recall Proposition 13.2). If $U \in \mathscr{U}$, we have $U \supseteq U_O$ for some $O \in \mathscr{O}_0$, and so $U(x) \supseteq U_O(x) = x - O$ is a neighborhood of x by Theorem 14.2. Conversely, if V is a neighborhood of x, we find by Theorem 14.2 some $O \in \mathscr{O}_0$ with $V \supseteq x - O = U_O(x)$. Hence,$U := U_O \cup \{(x,y) : y \in V\} \in \mathscr{U}$ satisfies $U(x) = V$, as desired. □

It can be proved that the uniform structure of Theorem 14.3 is the unique uniform structure such that the addition is uniformly continuous. However, we will not need this fact. If we speak of the *uniform structure* of a topological vector space, we always mean the uniform structure of Theorem 14.3.

Note that if X is a vector space, then also ${}^*X$ becomes equipped with an addition and a scalar multiplication by the standard definition for relations. The transfer principle implies that ${}^*X$ is actually a vector space (with scalar multiplication $\lambda x := ({}^*\lambda)\ {}^*\!\cdot\ x)$. We write $+$ in place of ${}^*+$ and 0 in place of ^{*0} (noting that ^{*0} is also the neutral element of addition in ${}^*X$ by the transfer principle).

Proposition 14.4. *Let X be a topological vector space. Then*

$$x \approx_{\mathscr{U}} y \iff x - y \approx_{\mathscr{O}} 0 \iff x - y \approx_{\mathscr{U}} 0 \qquad (x, y \in {}^*X). \tag{14.1}$$

Moreover, V is a neighborhood of $x \in {}^*X$ if and only if $V \supseteq x + {}^*O$ where $O \subseteq X$ is a neighborhood of 0.

In particular, V is a neighborhood of $x \in {}^*X$ if and only if $V = x + U$ for some neighborhood U of 0.

Proof. We have $x \approx_{\mathscr{U}} y$ if and only if for any $U \in \mathscr{U}$ the relation $(x,y) \in {}^*U$ holds. By Theorem 14.2, this holds if and only if $(x,y) \in {}^*U$ whenever

$$U \supseteq U_O := \{(\underline{x}, \underline{y}) \in X \times X : \underline{x} - \underline{y} \in O\}$$

where O is a neighborhood of 0. By the standard definition principle for relations, this is the case if and only if

$$(x,y) \in {}^*U_O = \{(\underline{x}, \underline{y}) \in {}^*X \times {}^*X : \underline{x} - \underline{y} \in {}^*O\}$$

for any neighborhood $O \subseteq X$ of 0, i.e. if and only if $x-y \in {}^*O$ for any neighborhood O of 0. But this means $x - y \approx_{\mathscr{O}} 0$. The second equivalence of (14.1) follows from Proposition 13.8.

For the second statement, note that V is a neighborhood of x if and only if $V \supseteq {}^*U^{-1}(x)$ for some $U \in \mathscr{U}$. This is the case if and only if $V \supseteq {}^*U_O^{-1}(x)$ for some neighborhood $O \subseteq X$ of 0. But this means $V \supseteq \{\underline{y} \in {}^*X : \underline{y} - x \in {}^*O\} = x + {}^*O$ for some neighborhood O of 0.

For the last statement, observe that by what we just proved, U is a neighborhood of 0 if and only if $U \supseteq 0 + {}^*O = {}^*O$ for some neighborhood $O \subseteq X$ of 0. Hence, if U is a neighborhood of 0, and $V := x + U$, then $V \supseteq x + {}^*O$ for some neighborhood O of 0, and so V is a neighborhood of x. Conversely, if V is a neighborhood of x, then $V \supseteq x + {}^*O$ for some neighborhood O of 0, and so $U := V - x \supseteq {}^*O$ is a neighborhood of 0 with $V = x + U$. □

It is, however, not true that ${}^*X$ is a topological vector space: In fact, by condition 3. of Theorem 14.2, we would otherwise find for any $x \in {}^*X$ and any neighborhood U of 0 some $\lambda \neq 0$ with $\lambda^{-1}x \in U$. Since $U \supseteq {}^*O$ for some neighborhood O of 0, this means that x is finite in the following sense:

Definition 14.5. Let X be a topological vector space. A point $x \in {}^*X$ is called *finite*, if for each neighborhood $U \subseteq X$ of 0 there is some $\lambda \in \mathbb{K}$ with $x \in \lambda {}^*U$. We use the notation

$$\operatorname{fin}({}^*X) := \{x \in {}^*X : x \text{ is finite}\}.$$

As we have mentioned at the end of §13, this is not the only natural definition of the term "finite": One could also use another definition which takes into account only the uniform structure of the space (and not the multiplication operation). Unfortunately, these two definitions may differ in certain cases. However, the above definition appears to be the most natural one in the context of topological vector spaces. For details, we refer the reader to [Hen72a, HM72].

Lemma 14.6. *If X is a topological vector space, then any neighborhood U of 0 contains a* balanced neighborhood O *of 0, i.e. $|\lambda| \leq 1$ implies $\lambda O \subseteq O$.*

Proof. By Theorem 14.2, we have $\Lambda_0 O_0 \subseteq U$ where $\Lambda_0 \subseteq \mathbb{K}$ and $O_0 \subseteq X$ are appropriate neighborhoods of 0. There is some $\varepsilon > 0$ such that $|\lambda| \leq \varepsilon$ implies $\lambda \in \Lambda_0$. Then $O := \{\lambda x : x \in O_0,\ |\lambda| \leq \varepsilon\}$ has the required properties. □

Proposition 14.7. *The following statements are equivalent:*

1. $x \in \operatorname{fin}({}^*X)$.
2. *For each balanced neighborhood U of 0 there is some $\lambda \in \mathbb{K}$ with $x \in \lambda {}^*U$.*
3. *For each neighborhood U of 0 there is some $n \in \mathbb{N}$ with $x \in n {}^*U$.*

Proof. If U is a neighborhood of 0, then there is a balanced neighborhood $U_0 \subseteq U$ of 0. If there is some $\lambda \in \mathbb{K}$ with $x \in \lambda^* U_0$, we also have $x \in \lambda^* U$. Thus, it suffices to consider balanced neighborhoods. But if U is a balanced neighborhood and $x \in \lambda^* U$, then we have $x \in n^* U$ where $n \in \mathbb{N}$ is such that $n \geq |\lambda|$. □

We define the set of *infinitesimals* of X as

$$\inf({}^*X) := \{x \in {}^*X : x \approx_{\mathscr{O}} 0\} = \{x \in {}^*X : x \approx_{\mathscr{U}} 0\} = \mathrm{mon}(0).$$

Exercise 73. Prove that $x \in \mathrm{fin}({}^*X)$ if and only if for any $c \in \inf({}^*\mathbb{R})$, $c > 0$, the relation $cx \in \inf({}^*X)$ holds.

Lemma 14.8. *The set* $\inf({}^*X)$ *is a linear subspace of* ${}^*X$.

Proof. If $x, y \in \inf({}^*X)$, $\lambda \in \mathbb{K}$, and U is a neighborhood of 0, we find by Theorem 14.2 some neighborhood O of 0 with $O + O \subseteq U$ and $\lambda O \subseteq U$. Since $x, y \in {}^*O$, we have $\lambda x \in {}^*(\lambda O) \subseteq {}^*U$ and $x + y \in {}^*O + {}^*O = {}^*(O + O) \subseteq {}^*U$. Hence, $\lambda x, x + y \in \mathrm{mon}(0)$. □

Theorem 14.9. *If* X *is a topological vector space, then* $\mathrm{fin}({}^*X) \subseteq {}^*X$ *is the largest subspace of* ${}^*X$ *with the property that* $\mathrm{fin}({}^*X)$ *is a topological vector space (with the inherited topology).*

The uniform structure on the topological vector space $\mathrm{fin}({}^*X)$ *is precisely the uniform structure inherited from* ${}^*X$.

Moreover, $\mathrm{fin}({}^*X)$ *is closed in* ${}^*X$. *In particular, if* $*$ *is* $\mathcal{P}(X)$*-saturated, then* $\mathrm{fin}({}^*X)$ *is complete.*

Proof. The fact that each topological vector subspace of ${}^*X$ must be contained in $\mathrm{fin}({}^*X)$ follows from our considerations preceding Definition 14.5.

Now we prove that $\mathrm{fin}({}^*X)$ is indeed a linear subspace: If $x, y \in \mathrm{fin}({}^*X)$ and $\lambda \in \mathbb{K}$ are given, then $cx, cy \in \inf({}^*X)$ for any $c \in \inf(\mathbb{R})$, $c > 0$ by Exercise 73. By Lemma 14.8, we have $c(x+y) = cx + cy \in \inf({}^*X)$ and $c\lambda x = \lambda(cx) \in \inf({}^*X)$ for any $c \in \inf(\mathbb{R})$, $c > 0$, and so $x + y, \lambda x \in \mathrm{fin}({}^*X)$ by Exercise 73.

To prove that $\mathrm{fin}({}^*X)$ is a topological vector space, we verify the three conditions of Theorem 14.2. Condition 1. follows from Proposition 14.4: If $x \in \mathrm{fin}({}^*X)$, and $U_0 \subseteq \mathrm{fin}({}^*X)$ is a neighborhood of 0, i.e. $U_0 = U \cap \mathrm{fin}({}^*X)$ for some neighborhood $U \subseteq {}^*X$ of 0, then $x + U$ is a neighborhood of x in ${}^*X$. Since $\mathrm{fin}({}^*X)$ is a vector space, it follows that $x + U_0 = x + (U \cap \mathrm{fin}({}^*X)) = (x + U) \cap \mathrm{fin}({}^*X)$ is a neighborhood of x. Analogously, if $V_0 \subseteq \mathrm{fin}({}^*X)$ is a neighborhood of $x \in \mathrm{fin}({}^*X)$, then we find some neighborhood $V \subseteq {}^*X$ of x with $V_0 = V \cap \mathrm{fin}({}^*X)$. Since $V - x$ is a neighborhood of 0 in ${}^*X$, it follows that $U_0 := V_0 - x = (V - x) \cap \mathrm{fin}({}^*X)$ is a neighborhood of 0 in $\mathrm{fin}({}^*X)$ with $V_0 = x + U_0$.

For conditions 2. and 3., let a neighborhood $U_0 \subseteq \operatorname{fin}({}^*X)$ of 0, $\lambda \in \mathbb{K}$, and $x \in \operatorname{fin}({}^*X)$ be given. We have $U_0 = U \cap \operatorname{fin}({}^*X)$ for some neighborhood $U \subseteq {}^*X$ of 0, and $U \supseteq {}^*V$ for some balanced neighborhood $V \subseteq X$ of 0 (Proposition 14.4 and Lemma 14.6). By Theorem 14.2, we find neighborhoods $O \subseteq X$ of 0 and $\Lambda \subseteq \mathbb{K}$ of λ with $O + O \subseteq V$ and $\Lambda O \subseteq V$. The transfer principle implies ${}^*O + {}^*O \subseteq {}^*V$ and $\Lambda {}^*O \subseteq {}^*V$. Hence, 2. holds with the neighborhoods Λ and $O_0 := {}^*O \cap \operatorname{fin}({}^*X)$. Indeed, ${}^*O$ is a neighborhood of 0 in ${}^*X$ by Proposition 14.4, and so O_0 is a neighborhood of 0 in $\operatorname{fin}({}^*X)$. Moreover, $O_0 + O_0 \subseteq {}^*V \cap \operatorname{fin}({}^*X) \subseteq U_0$, and $\Lambda O_0 \subseteq {}^*V \cap \operatorname{fin}({}^*X) \subseteq U_0$, since $\operatorname{fin}({}^*X)$ is a vector space. For condition 3., observe that we find some $n \in \mathbb{N}$ with $x \in n {}^*V$. For each μ in the neighborhood $\Lambda_0 := \{\mu \in \mathbb{K} : |\mu| < n^{-1}\}$ of 0, we have $\mu x \in {}^*V$, because ${}^*V$ is balanced, and thus $\mu x \in {}^*V \cap \operatorname{fin}({}^*X) \subseteq U_0$. Hence, also condition 3. holds.

Let for a moment $\mathscr{U}_0$ denote the uniform structure corresponding to the topological vector space $\operatorname{fin}({}^*X)$. We have to prove that $\mathscr{U}_0 = \mathscr{U}_{\operatorname{fin}({}^*X)}$ where

$$\mathscr{U}_{\operatorname{fin}({}^*X)} := \{U \cap (\operatorname{fin}({}^*X) \times \operatorname{fin}({}^*X)) \mid U \in \mathscr{U}_*\}$$

with $\mathscr{U}_*$ being the uniform structure of ${}^*X$ introduced in Definition 13.25. By definition, a set $U \subseteq \operatorname{fin}({}^*X) \times \operatorname{fin}({}^*X)$ satisfies $U \in \mathscr{U}_0$ if and only if there is a neighborhood $O \subseteq \operatorname{fin}({}^*X)$ of 0 with

$$U \supseteq \{(\underline{x}, \underline{y}) \in \operatorname{fin}({}^*X) \times \operatorname{fin}({}^*X) \mid \underline{x} - \underline{y} \in O\}.$$

Since the neighborhoods of 0 in $\operatorname{fin}({}^*X)$ are precisely those sets of the form $O \cap \operatorname{fin}({}^*X)$ where $O \subseteq {}^*X$ is a neighborhood of 0, we find that $\mathscr{U}_0$ consists of all sets of the form $U \cap (\operatorname{fin}({}^*X) \times \operatorname{fin}({}^*X))$ where $U \subseteq {}^*X \times {}^*X$ is such that there is some neighborhood $O \subseteq {}^*X$ of 0 with

$$U \supseteq U_O := \{(\underline{x}, \underline{y}) \in {}^*X \times {}^*X \mid \underline{x} - \underline{y} \in O\}.$$

In view of Proposition 14.4, this holds for U if and only if there is a neighborhood $O \subseteq X$ of 0 with $U \supseteq U_{{}^*O}$. The latter means by the inverse form of the standard definition principle for relations that

$$U \supseteq {}^*\{(\underline{x}, \underline{y}) \in X \times X : \underline{x} - \underline{y} \in O\}.$$

By Theorem 14.3, this is satisfied by U if and only if there is some $W \in \mathscr{U}$ with $U \supseteq {}^*W$. In view of Proposition 13.24, this means $U \in \mathscr{U}$.

To see that $\operatorname{fin}({}^*X)$ is closed, let some $x \in {}^*X$ be given with $x \notin \operatorname{fin}({}^*X)$. Then there is a neighborhood U of 0 such that $x \notin n {}^*U$ for any $n \in \mathbb{N}$. By Theorem 14.2 and Lemma 14.6, we find a balanced neighborhood O of 0 with $O - O \subseteq U$ (in particular, $O \subseteq U$). By Proposition 14.4, $V := x + {}^*O$ is a neighborhood of x. Moreover, $V \cap \text{fin}(^*X) = \varnothing$ which implies that $x \notin \overline{\text{fin}(^*X)}$: Indeed, assume that there is some $y \in V \cap \text{fin}(^*X)$. Then we find some n with $y \in n^*O$, i.e. there are $o_1, o_2 \in {}^*O$ with $no_1 = y = x + o_2$. But since $^*O$ is balanced, this implies $x = no_1 - o_2 \in n^*O - {}^*O \subseteq n(^*O - {}^*O) \subseteq n^*U$, a contradiction.

The last statement follows from Theorem 13.26 and Exercise 70. □

Definition 14.10. A subset $A \subseteq X$ of a topological vector space is called *bounded*, if for each neighborhood U of 0 there is some $\lambda \in \mathbb{R}_+$ with $A \subseteq \lambda U$.

Now we can formulate a generalization of Theorem 7.8:

Exercise 74. Prove that $A \subseteq X$ is bounded if and only if $^*A$ contains only finite elements, i.e. if and only if

$$^*A \subseteq \text{fin}(^*X).$$

Exercise 75. Prove the inclusion

$$\text{pns}(^*X) \subseteq \text{fin}(^*X),$$

and conclude that precompact subsets of topological vector spaces are bounded.

Exercise 76. Prove that the vector space $\inf(^*X)$ is a closed subspace of $^*X$ and contained in $\text{fin}(^*X)$. In particular, $\inf(^*X)$ is a closed subspace of $\text{fin}(^*X)$.

Recall that if U is a subspace of some vector space X, one defines the factor space X/U as the set of all equivalence classes with respect to the equivalence relation

$$x \approx y \iff x - y \in U.$$

The space X/U becomes a vector space with the operations $[x] + [y] = [x + y]$ and $[\lambda x] = \lambda[x]$ (which are well-defined). If X is a topological vector space and U is a subspace, then one equips X/U with the following topology: The open sets are the sets of the form $\{[x] : x \in O\}$ where O is open.

Proposition 14.11. *X/U is a topological vector space.*

Proof. We prove first that the sets $\tilde{O} := \{[x] : x \in O\}$ with open sets O form a topology: Clearly, $\varnothing, X/O$ are open. If $\{\tilde{O}_i : i\}$ is a family of open sets, then $\bigcup_i \tilde{O}_i = \{[x] : x \in \bigcup_i O_i\}$ is open. Finally, if $\tilde{O}_1, \tilde{O}_2$ are open, let $\mathscr{A}$ be the family of all open sets which are contained in $A := \tilde{O}_1 \cap \tilde{O}_2$. Then $\bigcup \mathscr{A} \subseteq A$, and if we can prove that even $A = \bigcup \mathscr{A}$, then A is open. Thus, let $[x] \in A$ be given. By choosing a proper representative, we may assume that $x \in O_1$, and we find some $u \in U$ with $x + u \in O_2$. Now we apply Theorem 14.2 several times: $O_3 = O_2 - x - u$ is a neighborhood of 0, and so $x + O_3 := O_2 - u$ is a neighborhood of x. We thus find an open set $O \subseteq O_1$ with $x \in O \subseteq O_2 - u$. Then $\tilde{O}$ is open with $[x] \in \tilde{O} \subseteq A$, and so $[x] \in \bigcup \mathscr{A}$, as desired.

To see that the addition is continuous at $([x],[y])$, let U be a neighborhood of $[x+y]=[x]+[y]$. Without loss of generality, we may assume that U is open, i.e. $U=\tilde{O}$ for some open set O where $x+y+u\in O$ for some $u\in U$. Since the addition is continuous in X, we find open sets O_1,O_2 with $x\in O_1$, $y\in O_2$ and $O_1+O_2\subseteq O-u$. Then $\tilde{O}_1$ and $\tilde{O}_2$ are neighborhoods of $[x]$ resp. $[y]$, and $\tilde{O}_1+\tilde{O}_2\subseteq\tilde{O}$. Hence, the addition is continuous. The continuity of the multiplication is proved analogously. □

Exercise 77. Prove that X/U is Hausdorff if and only if U is closed. In particular, X is Hausdorff if and only if $\{0\}$ is closed.

Definition 14.12. If X is a topological vector space, we define the *nonstandard hull* $\breve{X}$ as the factor space $\operatorname{fin}({}^*X)/\inf({}^*X)=\operatorname{fin}({}^*X)/\operatorname{mon}(0)$.

The reader should be aware that the nonstandard hull of a topological vector space X is in general different from the nonstandard hull of X considered as a uniform space. However, one has an inclusion:

Theorem 14.13. *The space $\breve{X}$ is a Hausdorff topological vector space. It is a subset of $\widetilde{X}$ and has the inherited uniform structure. Moreover, $\breve{X}$ is closed in $\widetilde{X}$. If $*$ is $\mathcal{P}(X)$-saturated, then $\breve{X}$ is complete.*

If X is Hausdorff, then $X\subseteq\breve{X}$ in the sense of the one-to-one embedding $x\mapsto[{}^*x]$.

Proof. The first statement follows from Exercise 77 and Exercise 76 or, alternatively, from the embedding $\breve{X}\subseteq\widetilde{X}$ if one observes that $\widetilde{X}$ is a Hausdorff space (Theorem 13.29) and that subspaces of Hausdorff spaces are Hausdorff.

By (14.1), we have $x-y\in\inf({}^*X)$ if and only if $x\approx_{\mathscr{U}}y$. Moreover, if additionally $x\in\operatorname{fin}({}^*X)$, then $x-y\in\inf({}^*X)\subseteq\operatorname{fin}({}^*X)$ implies that $y\in\operatorname{fin}({}^*X)$. Hence, for $x\in\operatorname{fin}({}^*X)$, the equivalence classes $[x]$ in the spaces $\breve{X}$ and $\widetilde{X}$ consist of the same elements (namely those elements of $\operatorname{fin}({}^*X)$ for which $x\approx_{\mathscr{U}}y$). In particular, $\breve{X}=\{[x]\in\widetilde{X}:x\in\operatorname{fin}({}^*X)\}\subseteq\widetilde{X}$.

To see that $\breve{X}$ carries the inherited uniformity, let first U be an element of the uniformity of $\breve{X}$. By Theorem 14.3, this means that we find some open neighborhood $V\subseteq\breve{X}$ of $[0]$ with

$$U\supseteq U_V:=\{([x],[y]):x,y\in\operatorname{fin}({}^*X),\ [x]-[y]\in V\}.$$

By definition of the topology in $\breve{X}$, we have $V=\{[x]:x\in O\}$ for some open neighborhood $O\subseteq{}^*X$ of 0. By Proposition 14.4, O is a neighborhood of 0 if and only if $O\supseteq{}^*O_0$ for some neighborhood $O_0\subseteq X$ of 0. Hence, whenever $x,y\in\operatorname{fin}({}^*X)$ satisfy $x-y\in{}^*O_0\subseteq O$, we have $[x]-[y]=[x-y]\in V$, and so $([x],[y])\in U_V\subseteq U$. This proves

$$U\supseteq\{([x],[y]):x,y\in\operatorname{fin}({}^*X)\wedge x-y\in{}^*O_0\}.$$

By Theorem 14.3, the set $U_0 := \{(\underline{x}, \underline{y}) \in X \times X : \underline{x} - \underline{y} \in O_0\}$ belongs to the uniform structure $\mathscr{U}$ of X. The standard definition principle for relations thus implies

$$\widetilde{{}^*U_0} = \{([x], [y]) \mid x, y \in {}^*X \wedge x - y \in {}^*O_0\}.$$

It follows that

$$\widetilde{{}^*U_0} \cap (\check{X} \times \check{X}) = \{([x], [y]) : x, y \in \mathrm{fin}({}^*X) \wedge x - y \in {}^*O_0\} \subseteq U,$$

and so U belongs to the uniform structure inherited from $\widetilde{\mathscr{U}}$.

Conversely, let U belong to the uniform structure inherited from $\widetilde{\mathscr{U}}$. Then $U \supseteq \widetilde{{}^*U_0} \cap (\check{X} \times \check{X})$ for some $U_0 \in \mathscr{U}$. By Theorem 14.3, there is some neighborhood $O_0 \subseteq X$ of 0 with

$$U_0 \supseteq \{(\underline{x}, \underline{y}) \in X \times X \mid \underline{x} - \underline{y} \in O_0\}.$$

In view of the standard definition principle, we find

$$\widetilde{{}^*U_0} \supseteq \{([x], [y]) \mid x, y \in {}^*X \wedge x - y \in {}^*O_0\},$$

and so

$$U \supseteq \widetilde{{}^*U_0} \cap (\check{X} \times \check{X}) \supseteq \{([x], [y]) \mid x, y \in \mathrm{fin}({}^*X) \wedge x - y \in {}^*O_0\}. \tag{14.2}$$

Since ${}^*O_0$ is a neighborhood of 0 (Proposition 14.4), we find some open neighborhood $O \subseteq {}^*O_0$ of 0. Then $O_1 := O \cap \mathrm{fin}({}^*X)$ is open in $\mathrm{fin}({}^*X)$, and so the set $V := \{[x] : x \in O_1\}$ is an open neighborhood of $[0]$. Theorem 14.3 implies that

$$V_0 := \{([x], [y]) \in \check{X} \times \check{X} \mid [x] - [y] \in V\}$$

belongs to the uniform structure of $\check{X}$. If we can prove that $V_0 \subseteq U$, then also U must belong to this uniform structure, and we are done. Thus, let $([x], [y]) \in V_0$ be given, i.e. $x, y \in \mathrm{fin}({}^*X)$ with $[x - y] = [x] - [y] \in V$. Then $x - y - u \in O \subseteq {}^*O_0$ for some $u \in \inf({}^*X) \subseteq \mathrm{fin}({}^*X)$. By (14.2), this implies $([x], [y + u]) \in U$. Since $[y] = [y + u]$, we find $([x], [y]) \in U$, as desired.

To see that $\check{X}$ is closed, let $y \in \widetilde{X} \setminus \check{X}$ be given, i.e. $y = [x]$ for some $x \in {}^*X \setminus \mathrm{fin}({}^*X)$. Since $\mathrm{fin}({}^*X)$ is closed, we find some neighborhood $O \subseteq {}^*X$ of x which is disjoint from $\mathrm{fin}({}^*X)$. By definition of the topology of ${}^*X$, we find some $U \in \mathscr{U}$ with $O \supseteq {}^*U(x)$. There is some $V \in \mathscr{U}$ with $V^3 \subseteq U$. Then $\widetilde{{}^*V}([x])$ is a neighborhood of $[x]$ which is disjoint from $\mathrm{fin}({}^*X)$. Indeed, if $[y] \in \widetilde{{}^*V}([x])$ then Lemma 13.28 implies $(x, y) \in ({}^*V)^3 \subseteq {}^*U$, and so $y \in {}^*U(x) \subseteq O$ which implies $y \notin \mathrm{fin}({}^*X)$, as desired.

If X is Hausdorff, Theorem 13.29 implies that the embedding $X \hookrightarrow \widetilde{X}$ defined by $x \mapsto [{}^*x]$ is one-to-one. Since ${}^*x$ is always finite, this is even an embedding into the subspace $\check{X}$. □

A map $\|\cdot\| : X \to [0,\infty)$ is called a *seminorm* if it has all properties of a norm with the exception that $\|x\| = 0$ need not imply $x = 0$. Any seminorm induces a pseudometric by $d(x,y) := \|x-y\|$. Thus, a family N of seminorms induces a family of pseudometrics which in turn induces a uniform structure. It can be verified straightforwardly that X is a topological vector space with this uniform structure. Not all topological vector spaces can be obtained in this way. It turns out that these are precisely those topological vector spaces for which 0 (and equivalently each point) has a neighborhood base of convex sets. For this reason, such spaces are called *locally convex*.

Proposition 14.14. *Let X be a locally convex space generated by the family N of pseudonorms. Then*

$$\operatorname{fin}({}^*X) = \{x \in {}^*X : {}^*\|x\| \textit{ is finite for each } \|\cdot\| \in N\},$$

and

$$\inf({}^*X) = \{x \in {}^*X : {}^*\|x\| \textit{ is infinitesimal for each } \|\cdot\| \in N\}.$$

The uniform structure of $\operatorname{fin}({}^*X)$ *is induced by the family of seminorms*

$$\|x\|_* := \operatorname{st}({}^*\|x\|) \qquad (\|\cdot\| \in N),$$

and the uniform structure of $\check{X}$ is induced by the family of seminorms

$$\|[x]\| \ = \operatorname{st}({}^*\|x\|) \qquad (\|\cdot\| \in N).$$

If X is seminormed, i.e. $N = \{\|\cdot\|\}$, then $\check{X}$ is normed. Moreover, if $*$ is $(\mathcal{P}(N) \times \mathbb{N})$-saturated, then $\operatorname{fin}({}^*X)$ *and $\check{X}$ are complete.*

Proof. We have $x \in \operatorname{fin}({}^*X)$ if and only if for any $\varepsilon \in \mathbb{R}_+$ and each finitely many $\|\cdot\|_1, \dots, \|\cdot\|_k \in N$, we find some $n \in \mathbb{N}$ with

$$x \in n^*\{\underline{y} \in X : \|\underline{y}\|_1 < \varepsilon \wedge \cdots \wedge \|\underline{y}\|_k < \varepsilon\}.$$

By the standard definition principle, this means ${}^*\|x/n\|_1 < {}^*\varepsilon, \dots, {}^*\|x/n\|_k < {}^*\varepsilon$, i.e. ${}^*\|x\|_1 < n^*\varepsilon, \dots, {}^*\|x\|_k < n^*\varepsilon$. Hence $x \in \operatorname{fin}({}^*X)$ if and only if ${}^*\|x\| \in \operatorname{fin}({}^*\mathbb{R})$ for each $\|\cdot\| \in N$.

Similarly, $x \in \inf({}^*X)$ if and only if for any $\varepsilon \in \mathbb{R}_+$ and each finitely many $\|\cdot\|_1, \dots, \|\cdot\|_n \in N$, we have

$$x \in {}^*\{\underline{y} \in X : \|\underline{y}\|_1 < \varepsilon \wedge \cdots \wedge \|\underline{y}\|_n < \varepsilon\}$$

which by the standard definition principle means ${}^*\|x\|_1, \dots {}^*\|x\|_k \le {}^*\varepsilon$. Hence, $x \in \inf({}^*X)$ if and only if ${}^*\|x\| \in \inf({}^*\mathbb{R})$ for each $\|\cdot\| \in N$.

The transfer principle implies for any $\|\cdot\| \in N$ that

$$\forall \underline{x}, \underline{y} \in {}^*X : {}^*\|x+y\| \leq {}^*\|x\| + {}^*\|y\|,$$

and since st is additive and monotone, the triangle inequality for $\|\cdot\|_*$ follows. The equality $\|\lambda x\|_* = |\lambda|\ \|x\|_*$ is proved analogously. Hence, $\|\cdot\|_*$ ($\|\cdot\| \in N$) is a family of seminorms on $\mathrm{fin}({}^*X)$ which generates the same uniform structure as the pseudometrics $d_*(x,y) := \mathrm{st}({}^*d(x,y))$ ($d \in D$) where D is the family of all pseudometrics $d(x,y) := \|x-y\|$ with $\|\cdot\| \in N$. Hence, the statement concerning $\mathrm{fin}({}^*X)$ follows from Theorem 13.27 (recall that $\mathrm{fin}({}^*X)$ is a closed subspace of ${}^*X$, and so the completeness of ${}^*X$ implies the completeness of $\mathrm{fin}({}^*X)$ by Exercise 70). The proof of the statements concerning $\breve{X}$ follows analogously, using Proposition 13.32. □

In general, also the nonstandard hull of a Hausdorff topological vector space is much larger than its completion. Roughly speaking, the elements of the nonstandard hull $\breve{X}$ which are not contained in the completion are some sort of "infinitesimals for the dimensions". In particular, $\breve{X}$ is the completion of a normed space X if and only if X has finite dimension. To prove this, we need the following result:

Theorem 14.15. *Let X be a Hausdorff topological vector space. Then $\breve{X}$ is the closure of X in $\widetilde{X}$ (in the sense of the embedding $x \mapsto [{}^*x]$) if and only if* $\mathrm{fin}({}^*X) = \mathrm{pns}({}^*X)$.

Proof. By Theorem 13.29, the closure of X in $\widetilde{X}$ is equal to

$$\overline{X} := \{[x] : x \in \mathrm{pns}({}^*X)\}.$$

Thus $\breve{X}$ is this closure if and only if

$$\{[x] : x \in \mathrm{fin}({}^*X)\} = \breve{X} = \overline{X} = \{[x] : x \in \mathrm{pns}({}^*X)\}.$$

This is true if $\mathrm{fin}({}^*X) = \mathrm{pns}({}^*X)$. Conversely, if $\mathrm{fin}({}^*X) \neq \mathrm{pns}({}^*X)$, then Exercise 75 implies that we find some $x \in \mathrm{fin}({}^*X) \neq \mathrm{pns}({}^*X)$. Then $[x] \in \breve{X}$, but we have $[x] \notin \overline{X}$, since otherwise $[x] = [y]$ for some $y \in \mathrm{pns}({}^*X)$. But this implies the contradiction $x \in \mathrm{pns}({}^*X)$: Indeed, given $U \in \mathscr{U}$, choose some $V \in \mathscr{U}$ with $V^2 \subseteq U$. Since $y \in \mathrm{pns}({}^*X)$, we find some $z \in X$ with $({}^*z, y) \in {}^*V$. Since $(y,x) \in {}^*V$ (because $x \approx_{\mathscr{U}} y$), we have $({}^*z, x) \in {}^*V^2 \subseteq {}^*U$. Hence, $x \in \mathrm{pns}({}^*X)$, as claimed. □

A topological vector space is called a *Montel space*, if all bounded subsets are precompact. It follows from the well-known lemma of Riesz that the unit ball of an infinite-dimensional normed space is not precompact, and so no infinite-dimensional normed space is Montel. The following result thus implies in particular that infinite-dimensional normed spaces always satisfy $\overline{X} \not\subseteq \widetilde{X}$:

Corollary 14.16. *Let X be a Hausdorff topological vector space. If $\breve{X}$ is the closure of X in $\widetilde{X}$, then X is a Montel space.*

Proof. If $A \subseteq X$ is bounded, then Exercise 74 implies ${}^*A \subseteq \operatorname{fin}({}^*X) = \operatorname{pns}({}^*X)$. Hence, A is precompact by Theorem 13.22. □

It is not known whether a converse of Corollary 14.16 is true. At least, there *are* Montel spaces for which $\breve{X} = \overline{X}$, namely all finite-dimensional spaces (and there are even infinite-dimensional Montel space with this property, see [HM72, Example 4.5]):

Theorem 14.17. *If X is a finite-dimensional Hausdorff topological vector space, then $X \cong \breve{X}$.*

Proof. It is well-known that finite-dimensional Hausdorff topological vector spaces are Montel spaces and that the topology is generated by some norm (see e.g. [Rud90] for a proof). Hence, X is complete, and there is a bounded (and thus precompact) neighborhood U of 0. Given $x \in \operatorname{fin}({}^*X)$, we find some $n \in \mathbb{N}$ with $x \in n{}^*U$. Since U is precompact, also nU is precompact, and so ${}^*(nU) \subseteq \operatorname{pns}({}^*X)$ by Theorem 13.22. It follows that $x \in \operatorname{pns}({}^*X)$, and so we have proved $\operatorname{fin}({}^*X) \subseteq \operatorname{pns}({}^*X)$. Exercise 75 now implies $\operatorname{fin}({}^*X) = \operatorname{pns}({}^*X)$, and by Theorem 14.15, we find that $\breve{X}$ is the closure of X in $\widetilde{X}$ (which is X, since X is complete). □

In particular, we have for $X = \mathbb{R}^n$ that $\operatorname{fin}({}^*\mathbb{R}^n)/\inf({}^*\mathbb{R}^n) = \breve{\mathbb{R}}^n \cong \mathbb{R}^n$. For $n = 1$, this is a new proof of (a part of) Theorem 5.21. The previous results imply:

Corollary 14.18. *If X is normed, then $\breve{X}$ is the closure of X in $\breve{X}$ if and only if X has finite dimension.*

The reader who is interested in deeper results on the nonstandard theory of topological vector spaces and normed spaces is referred to the papers [HM72, HM74, HM83] and also to [Lux69a, Lux92].

Chapter 7

Miscellaneous

§15 Loeb Measures

To define nonstandard measures, we first recall the Carathéodory extension procedure from measure theory:

Let Σ_0 be a set algebra over some set S_0, and $\mu : \Sigma_0 \to [0, \infty]$ a σ-additive measure. It turns out that under this assumption μ may be extended to a σ-additive measure μ^* on a σ-algebra $\Sigma \supseteq \Sigma_0$. This may be done constructively by the Carathéodory extension procedure:

One first defines the so-called *outer measure* for μ on *all* subsets of Σ_0:

$$\mu^*(E) = \inf\{\sum_{n=1}^{\infty} \mu(A_n) \mid E \subseteq \bigcup_{n=1}^{\infty} A_n,\ A_n \in \Sigma_0\}.$$

Clearly, $\mu^*(A) = \mu(A)$ if $A \in \Sigma$. Moreover, μ^* is *monotone* (i.e. $D \subseteq E$ implies $\mu^*(D) \subseteq \mu^*(E)$) and *$\sigma$-subadditive* (i.e. $\mu^*(\bigcup E_n) \leq \sum \mu^*(E_n)$). To see the latter, note that, given $\varepsilon > 0$ and sets E_n, we find for any n a sequence $A_{n,k} \in \Sigma_0$ with $E_n \subseteq \bigcup_k A_{n,k}$ and $\mu^*(E_n) \geq \sum \mu(A_n) - 2^{-n}\varepsilon$. Hence,

$$\mu^*(\bigcup E_n) \leq \sum_{n,k=1}^{\infty} \mu(A_{n,k}) \leq \sum_{n=1}^{\infty}(\mu^*(E_n) + 2^{-n}\varepsilon) = \varepsilon + \sum_{n=1}^{\infty} \mu^*(E_n).$$

In general, μ^* is not σ-additive, but one is only interested in finding a σ-algebra $\Sigma \supseteq \Sigma_0$ such that the restriction of μ^* to Σ is σ-additive. It turns out that such a σ-algebra is given by the system Σ of all sets $E \subseteq S_0$ which satisfy the estimate

$$\mu^*(D) \geq \mu^*(D \cap E) + \mu^*(D \setminus E) \qquad (D \subseteq S_0). \tag{15.1}$$

Theorem 15.1 (Carathéodory)**.** *Σ is a σ-algebra, and the restriction of μ^* to Σ is a measure extending μ.*

Proof. To prove that μ^* extends μ, we only have to prove $\Sigma_0 \subseteq \Sigma$. Thus, let $E \in \Sigma_0$ and $D \subseteq S_0$ be given. Given $\varepsilon > 0$, we find a sequence $A_n \in \Sigma_0$ with $D \subseteq \bigcup A_n$ and $\mu^*(D) \geq \sum \mu(A_n) - \varepsilon$. Then the sets $A_n \cap E$ and $A_n \setminus E$ belong to Σ_0, and since their union covers $D \cap E$ resp. $D \setminus E$, we find

$$\mu^*(D \cap E) + \mu^*(D \setminus E) \leq \sum_{n=1}^{\infty} \mu(A_n \cap E) + \sum_{n=1}^{\infty} \mu(A_n \setminus E) = \sum_{n=1}^{\infty} \mu(A_n) \leq \mu^*(D) + \varepsilon.$$

Hence, (15.1) holds which by definition means $E \in \Sigma$.

For the proof that the restriction of μ^* to Σ is a measure, we only need to show that μ^* is monotone and σ-subadditive. We divide this proof into two parts:

1. We prove first that Σ is an algebra. Since $\mu^*(\varnothing) = 0$, we have $S_0 \in \Sigma$. If $E \in \Sigma$ and $D \subseteq S_0$, then

$$\mu^*(D \cap (S \setminus E)) + \mu^*(D \setminus (S \setminus E)) = \mu^*(D \setminus E) + \mu^*(D \cap E) \leq \mu^*(D),$$

and so $S \setminus E \in \Sigma$. Finally, let $E, F \in \Sigma$. For $D \subseteq S$ we calculate, putting $D_E := D \setminus E$, that

$$\begin{aligned}\mu^*(D \cap (E \cup F)) + \mu^*(D \setminus (E \cup F)) &= \mu^*((D \cap E) \cup (D_E \cap F)) + \mu^*(D_E \setminus F) \\ \leq \mu^*(D \cap E) + \mu^*(D_E \cap F) + \mu^*(D_E \setminus F) &\leq \mu^*(D \cap E) + \mu^*(D_E) \leq \mu^*(D).\end{aligned}$$

2. Now observe that for $E \in \Sigma$ and any $F, D \subseteq S$ with $F \cap E = \varnothing$ we have

$$\begin{aligned}\mu^*(D \cap (E \cup F)) &\geq \mu^*((D \cap (E \cup F)) \cap E) + \mu^*((D \cap (E \cup F)) \setminus E) \\ &= \mu^*(D \cap E) + \mu^*(D \cap F).\end{aligned}$$

An induction by N thus implies that for any sequence $E_n \in \Sigma$ of pairwise disjoint sets and all $D \subseteq S$,

$$\mu^*(D \cap \bigcup_{n=1}^{N} E_n) \geq \sum_{n=1}^{N} \mu^*(D \cap E_n).$$

Since $\bigcup_{n \leq N} E_n \in \Sigma$ by step 1., this implies

$$\mu^*(D) \geq \mu^*(D \cap \bigcup_{n=1}^{N} E_n) + \mu^*(D \setminus \bigcup_{n=1}^{N} E_n) \geq \sum_{n=1}^{N} \mu^*(D \cap E_n) + \mu^*(D \setminus \bigcup_{n=1}^{\infty} E_n).$$

Letting $N \to \infty$, we thus have by the σ-subadditivity

$$\mu^*(D) \geq \sum_{n=1}^{\infty} \mu^*(D \cap E_n) + \mu^*(D \setminus \bigcup_{n=1}^{\infty} E_n) \geq \mu^*(\bigcup_{n=1}^{\infty}(D \cap E_n)) + \mu^*(D \setminus \bigcup_{n=1}^{\infty} E_n)$$
$$= \mu^*(D \cap \bigcup_{n=1}^{\infty} E_n) + \mu^*(D \setminus \bigcup_{n=1}^{\infty} E_n).$$

In particular, $\bigcup E_n \in \Sigma$. Moreover, in view of the subadditivity of μ^*, we even have equality in the above estimate, and for the choice $D = \bigcup E_n$, this implies that μ^* is σ-additive on Σ.

To prove that Σ is a σ-algebra, it now suffices to prove that $E_n \in \Sigma$ implies $\bigcup E_n \in \Sigma$. For pairwise disjoint sets E_n we have just proved this, and the general case reduces to this: Just replace E_n by $E_n \setminus \bigcup_{k<n} E_k$. □

Sets of finite measure may be approximated by elements of Σ up to an arbitrary small error $\varepsilon \in \mathbb{R}_+$:

Theorem 15.2. *Let $E \in \Sigma$ satisfy $\mu^*(E) < \infty$. Then for any $\varepsilon \in \mathbb{R}_+$ there is some $D \in \Sigma_0$ such that $\mu^*(D \Delta E) < \varepsilon$. In particular, $\mu^*(D)$ and $\mu^*(E)$ differ at most by ε.*

Proof. By definition of μ^*, we find a sequence $A_n \in \Sigma_0 \subseteq \Sigma$ with $E \subseteq A := \bigcup A_n$ and

$$\mu^*(E) \geq \sum_{n=1}^{\infty} \mu(A_n) - \varepsilon/2 = \sum_{n=1}^{\infty} \mu^*(A_n) - \varepsilon/2 \geq \mu^*(A) - \varepsilon/2.$$

Put $D_n := A_1 \cup \cdots \cup A_n$ and $D_0 := \varnothing$. Then the sets $B_n := D_n \setminus D_{n-1}$ are pairwise disjoint with $\bigcup B_n = \bigcup D_n = A$, and so

$$\mu^*(E) - \mu^*(E \setminus D_n) = \mu^*(E \cap D_n) = \mu^*(\bigcup_{k=1}^{n}(E \cap B_k)) = \sum_{k=1}^{n} \mu^*(E \cap B_k)$$
$$\to \sum_{k=1}^{\infty} \mu^*(E \cap B_k) = \mu^*(\bigcup_{k=1}^{\infty}(E \cap B_k)) = \mu^*(E \cap A) = \mu^*(E).$$

Hence, we find some n such that

$$\mu^*(E \setminus D_n) \leq \varepsilon/2.$$

For this n, we have

$$\mu^*(E \Delta D_n) \leq \mu^*(E \setminus D_n) + \mu^*(D_n \setminus E)$$
$$\leq \varepsilon/2 + \mu^*(A \setminus E) = \varepsilon/2 + (\mu^*(A) - \mu^*(E)) \leq \varepsilon.$$

□

If $\mu^*(E) = \infty$, we still can approximate E in a weaker sense by elements of D, if E is a so-called *σ-finite set*. The latter means that E is the union of countably many sets of finite measure.

Corollary 15.3. *Let $E \in \Sigma$ be σ-finite. Then for any $\varepsilon > 0$ there is a countable union D of sets from Σ_0 such that $\mu^*(E\Delta D) \leq \varepsilon$.*

Proof. Since E is σ-finite, it is the union of countably many sets $E_n \in \Sigma$ with $\mu^*(E_n) < \infty$. By Theorem 15.2, we find for each n some $D_n \in \Sigma_0$ with $\mu^*(E_n \Delta D_n) < 2^{-n}\varepsilon$. Since $E = \bigcup E_n$, we have for $D := \bigcup D_n$ that

$$\mu^*(E\Delta D) \leq \mu^*(\bigcup_{n=1}^{\infty}(E_n\Delta D_n)) \leq \sum_{n=1}^{\infty}\mu^*(E_n\Delta D_n) < \varepsilon.$$

□

We point out that the previous results hold true if Σ is not necessarily an algebra: It suffices that Σ is a so-called *semi-ring*. For generalizations in this direction, we refer the reader to [Zaa67].

So far, we have only recalled some results of the standard world. The interesting point in connection with nonstandard analysis is that any additive measure μ is *automatically* σ-additive, if it is internal:

Through the rest of this section, we assume that $S_0 \in \widehat{{}^*S}$ is a nonstandard entity, and that $[0,\infty] \in \widehat{S}$ is an entity. Let Σ be an algebra over S_0. Then an internal function $\mu : \Sigma \to {}^*[0,\infty]$ is called an *internal additive measure*, if $\mu(A \cup B) = \mu(A) + \mu(B)$ for all $A, B \in \Sigma$ with $A \cap B = \varnothing$. Note that Σ is internal by Theorem 3.19, and so any $A \in \Sigma$ is internal, too (Proposition 3.16).

Example 15.4. If μ is an additive measure in the standard world, then ${}^*\mu$ is an internal measure. This may be straightforwardly verified by the transfer principle.

Exercise 78. (Cumbersome). With Σ and μ as above, prove that for any *-finite sequence $A_1, \ldots, A_h \in \Sigma$ ($h \in {}^*\mathbb{N}$) the relation $A_1 \cup \cdots \cup A_h \in \Sigma$ holds. Moreover, if the sets A_k are pairwise disjoint, prove also

$$\mu(A_1 \cup \cdots \cup A_h) = \mu(A_1) + \cdots + \mu(A_h).$$

Each internal additive measure μ gives rise to an additive measure with standard values, defined by

$$\mu_0(A) := \begin{cases} \mathrm{st}(\mu(A)) & \text{if } \mu(A) \text{ is finite,} \\ \infty & \text{if } \mu(A) \text{ is infinite.} \end{cases} \qquad (A \in \Sigma).$$

The fact that this is indeed an additive measure follows immediately from the additivity of st (Theorem 5.21). As mentioned above, even more holds true automatically:

Proposition 15.5. *If $*$ is $\mathbb{N}$-saturated then, for any sequence $A_n \subseteq S_0$ of pairwise disjoint nonempty internal sets, the union $\bigcup A_n$ is not internal. In particular, with the above notation:*

1. *Σ is either finite or fails to be a σ-algebra.*
2. *μ and μ_0 are σ-additive.*

Proof. If $A := \bigcup A_n$ were internal, then each of the sets $B_n := A \setminus (A_1 \cup \cdots \cup A_n)$ were internal by Theorem 3.19. Then $\mathscr{B} := \{B_n : n \in \mathbb{N}\}$ is a countable family of nonempty internal sets with the finite intersection property, and so $\bigcap \mathscr{B} \neq \varnothing$. This contradicts the definition of A.

If Σ is infinite, we find a sequence $B_n \in \Sigma$ of nonempty internal sets. Putting $A_n := B_n \setminus \bigcup_{k<n} B_k$, we have $A_n \in \Sigma$ but not $\bigcup A_n \in \Sigma$. Hence, Σ is not a σ-algebra.

Finally, if $A_n \in \Sigma$ are pairwise disjoint with $\bigcup A_n \in \Sigma$, then all except finitely many A_n must be empty. Hence, the σ-additivity of μ and μ_0 follows from their additivity. □

In this connection, we recall that any ultrapower model is automatically comprehensive and thus $\mathbb{N}$-saturated.

Note that although μ_0 takes values in the standard world, its domain of definition is in the nonstandard world. Of course, by identifying $[0, \infty]$ with ${}^\sigma[0, \infty]$, one might consider μ_0 as a mapping of the nonstandard world. However, this mapping must be external if it attains infinitely many values: This follows from Theorem 3.19, because the range of this mapping is ${}^\sigma\mathrm{rng}(\mu_0)$ which is external by Theorem 3.22.

The essential point of our above considerations is that μ_0 is *automatically* a σ-additive function defined on the algebra Σ. In particular, we may apply the Carathéodory extension procedure as explained in the beginning:

Let μ_0^* denote the corresponding outer measure, and Σ_L be the system of all sets satisfying (15.1) (for μ_0^*). Then Σ_L is a σ-algebra, and the restriction of μ_0^* to Σ_L is a measure. This measure μ_L is called the *Loeb measure* for the given additive internal measure μ.

For the rest of this section, we keep this notation.

In contrast to the general Carathéodory extension procedure, the Loeb measure has particularly nice properties. In particular, Theorem 15.2 and Corollary 15.3 hold even with $\varepsilon = 0$.

The proof of this fact is not as easy as one might suspect at first glance, because to apply the saturation property, one has to consider *internal* sets. In particular, one has to consider μ in place of the Loeb measure μ_L.

Theorem 15.6. *Let $*$ be $\mathbb{N}$-saturated. Let E be measurable with respect to the Loeb measure μ_L and such that $\mu_L(E) < \infty$. Then there is some $D \in \Sigma$ with $\mu_L(D\Delta E) = 0$.*

Proof. By Theorem 15.2, we find for any n some set $D_n \in \Sigma$ with $\mu_L(D_n\Delta E) < n^{-1}$. We may conclude that

$$\mu_L(D_n\Delta D_k) \le \mu_L((D_n\Delta E) \cup (E\Delta D_k)) \le n^{-1} + k^{-1}.$$

It follows that the binary relation

$$\varphi := \{(\underline{x}, \underline{n}, \underline{y}) \in (\Sigma \times {}^{\sigma}\mathbb{N}) \times \Sigma \mid \mu(\underline{x}\Delta\underline{y}) \le 3/\underline{n}\}$$

is concurrent on the set $M := \{(D_n, {}^*n) : n \in \mathbb{N}\}$: Indeed, given $n_1, \ldots, n_m \in \mathbb{N}$, put $k := \max\{n_1, \ldots, n_m\}$. Then

$$\mu(D_{n_i}\Delta D_k) \approx {}^*(\mu_L(D_n\Delta D_k)) \le n_i^{-1} + k^{-1} \le 2/n_i^{-1} \qquad (i = 1, \ldots, m).$$

By Theorem 8.12, it follows that φ is satisfied on M, i.e. there is some $D \in \Sigma$ such that $\mu(D_n\Delta D) \le 3/{}^*n$ for any $n \in \mathbb{N}$. Since st is monotone, we find that $\mu_L(D_n\Delta D) \le 3/n$ for any $n \in \mathbb{N}$. Hence,

$$\mu_L(D\Delta E) \le \mu_L((D\Delta D_n) \cup (D_n\Delta E)) \le 3/n + 1/n = 4/n.$$

Since $\mu_L(D\Delta E)$ is standard, this implies $\mu_L(D\Delta E) = 0$. □

Corollary 15.7. *Let $*$ be $\mathbb{N}$-saturated, and S_0 be σ-finite with respect to the Loeb measure. Then a set $E \subseteq S_0$ is Loeb measurable if and only if there is a countable union D of sets from Σ with $\mu_0^*(D\Delta E) = 0$ (and in this case $\mu_L(D\Delta E) = 0$).*

Proof. By assumption, we find Loeb measurable sets F_n with $S_0 = \bigcup F_n$ and $\mu_L(F_n) < \infty$. Hence, if E is Loeb measurable, the sets $E_n := E \cap F_n$ satisfy $\mu(E_n) < \infty$. By Theorem 15.6, we find for each n some D_n with $\mu_L(D_n\Delta E_n) = 0$. Since $E = \bigcup E_n$, we have for $D = \bigcup D_n$ that

$$\mu_0^*(D\Delta E) \le \mu_0^*(\bigcup(D_n\Delta E_n)) \le \sum \mu_0^*(D_n\Delta E_n) = 0.$$

For the converse, recall that the Loeb measurable sets constitute a σ-algebra. Hence, if D is the countable union of sets from Σ, then D is Loeb measurable. Now, if $E \subseteq S_0$ satisfies $\mu_0^*(E\Delta D) = 0$, put $F := E\Delta D$. For any $C \subseteq S_0$, we have

$$\mu_0^*(C) \ge \mu_0^*(C \setminus F) = \mu_0^*(C \cap F) + \mu_0^*(C \setminus F),$$

i.e. (15.1) holds for μ_0 which by definition means that F is Loeb measurable. Hence, also $E = D\Delta F$ is Loeb measurable. □

Also the following property does not hold for general Carathéodory extensions. Recall that any nonstandard ultrapower model is $\mathbb{N}$-saturated and comprehensive.

Theorem 15.8. *Let $*$ be $\mathbb{N}$-saturated and comprehensive. Let $E \subseteq S_0$ be contained in a set of finite Loeb measure. Then E is Loeb measurable if and only if for each $\varepsilon > 0$ there are sets $C, D \in \Sigma$ such that $C \subseteq E \subseteq D$ and $\mu_L(D \setminus C) < \varepsilon$.*

Proof. First, assume that E is Loeb measurable. We find by definition of μ_0^* a sequence $D_n \in \Sigma$ with $E \subseteq \bigcup D_n$ and

$$\mu_L(E) = \mu_0^*(E) \geq \sum_{n=1}^{\infty} \mu_0(D_n) - \varepsilon.$$

It is no loss of generality to assume that the sets D_n are pairwise disjoint: Otherwise replace D_n by $D_n \setminus \bigcup_{k<n} D_k$.

Define $f : {}^\sigma\mathbb{N} \to \Sigma$ by $f({}^*n) = D_n$. Since $*$ is comprehensive, we find an extension of f to an internal mapping $f : {}^*\mathbb{N} \to \Sigma$. Let

$$M := \{\underline{n} \in {}^*\mathbb{N} \mid (\exists \underline{k} \in {}^*\mathbb{N} : (\underline{k} \leq \underline{n} \wedge F(\underline{n}+1) \cap F(\underline{k}) \neq \varnothing)) \\ \vee \sum_{\underline{k}=1}^{\underline{n}+1} \mu(F(\underline{k})) > {}^*(\mu_L(E) + \varepsilon)\}.$$

By the internal definition principle, M is internal. Hence, either M has some smallest element h, or $M = \varnothing$ in which case we choose some arbitrary $h \in \mathbb{N}_\infty$: In both cases the sets $F(1), \ldots, F(h)$ are pairwise disjoint, and we have

$$\sum_{\underline{k}=1}^{h} \mu(F(\underline{k})) \leq {}^*(\mu_L(E) + \varepsilon).$$

Moreover, by construction, h cannot be finite. The set $D := F(1) \cup \cdots \cup F(h)$ belongs to Σ (Exercise 78). Since $h \in \mathbb{N}_\infty$, we have $D \supseteq \bigcup D_n \supseteq E$. Moreover, since the sets $F(1), \ldots, F(h)$ are pairwise disjoint, we have

$${}^*(\mu_L(D)) \approx \mu(D) = \sum_{\underline{k}=1}^{h} \mu(F(\underline{k})) \leq {}^*(\mu_L(E) + \varepsilon),$$

and so $\mu_L(D) \leq \mu_L(E) + \varepsilon$.

Replacing E by $E_0 := D \setminus E$ in the above argument, we find a set $D_0 \in \Sigma$ with $D_0 \supseteq E_0$ and $\mu_L(D_0) \leq \mu_L(D_0) \leq \mu_L(E_0) + \varepsilon$. Putting $C := D \setminus D_0$, we have then $C \subseteq E \subseteq D$ and

$$\mu_L(D \setminus C) \leq \mu_L(D_0) \leq \mu_L(E_0) + \varepsilon = (\mu_L(D) - \mu_L(E)) + \varepsilon \leq 2\varepsilon.$$

Conversely, assume that there are sequences $C_n, D_n \in \Sigma$ with $C_n \subseteq E \subseteq D_n$ and $\mu_L(D_n \setminus C_n) \to 0$. For $C := \bigcup C_n$ and $D := \bigcap D_n$, we have then $C \subseteq E \subseteq D$ and $\mu_L(D \setminus C) \leq \inf_n \mu_L(D_n \setminus C_n) = 0$. Hence, $E \Delta D$ (and also $E \Delta C$) are null sets, and so E is Loeb measurable by Corollary 15.7. □

Theorem 15.8 is essentially the main result from [Loe75].

§16 Distributions

In this section, we show how nonstandard analysis can be used to describe distributions. We only sketch some ideas and leave the details to the reader.

Throughout this section, we assume that $\mathbb{R} \in \widehat{S}$ is an entity.

The idea of distributions goes back to *Dirac's δ-function*. The physical idea was that this is a function δ with the property that $\delta(x) = 0$ for all $x \neq 0$ but satisfying

$$\int_{\mathbb{R}} \delta(x)\, dx = 1.$$

Of course, no function δ with this property does exist. Nevertheless, a formal calculation with the δ-function in physics was successful, in particular due to the essential property

$$\int_{\mathbb{R}} \delta(x) f(x)\, ds = f(0).$$

An appropriate mathematical framework for the treatment of the δ-function is the following:

The closure of the set $\{x \in \mathbb{R} : f(x) \neq 0\}$ is called the *support of* f. One identifies a locally integrable function φ with the linear functional

$$F_\varphi(f) = \int_{\mathbb{R}} \varphi(x) f(x)\, ds$$

defined on e.g. the system of all smooth functions f with compact support. Then the "δ-function" can be identified with the linear functional

$$F_\delta(f) = f(0).$$

If one now considers integration as an application of the corresponding linear functional, we may indeed say that F_δ "is" the δ-function. Note that there is no function φ such that $F_\delta = F_\varphi$.

In nonstandard analysis, we can define the "δ-function" as an actual (internal) function δ, and we really may consider (internal) integrals in place of abstract linear functionals. The above mentioned fact that the "δ-function" is not a function in the usual sense is explained by the fact that δ is not a *standard* function.

Definition 16.1. Let L_1 denote the system of all integrable functions $f : \mathbb{R} \to \mathbb{R}$. Write $I : L_1 \to \mathbb{R}$ for the mapping which associates to each $f \in L_1$ its integral. Then we call ${}^*L_1$ the system of *internal integrable functions* and define the *internal integral*

$$\int f(x)\, dx := {}^*I(f) \qquad (f \in {}^*L_1).$$

It follows from the transfer principle that the internal integral is linear in the strong sense that

$$\int (\lambda f(x) + \mu g(x))\, ds = \lambda \int f(x)\, ds + \mu \int g(x)\, dx$$

holds even for not necessarily standard numbers $\lambda, \mu \in {}^*\mathbb{R}$.

Let L_1^{loc} denote the system of locally integrable functions $f : \mathbb{R} \to \mathbb{R}$, and C_0 denote the system of all continuous functions with compact support. Functions from ${}^*L_1^{\mathrm{loc}}$ are called *locally integrable internal functions*. It follows from the transfer principle that for all $\varphi \in {}^*L_1^{\mathrm{loc}}$ and $f \in {}^*C_0$ the function $g := \varphi f$ (defined by $g(x) := \varphi(x) f(x)$) belongs to ${}^*L_1$. However, the integral is a value from ${}^*\mathbb{R}$ and need not necessarily be finite. Let $L \subseteq {}^*L_1^{\mathrm{loc}}$ be the system of all locally integrable internal functions φ with the property that the integral $\int \varphi(x) f(x)\, dx$ is finite for any $f \in {}^\sigma C_0$. Any $\varphi \in L$ defines a linear functional F_φ on C_0 by means of the formula

$$F_\varphi(f) := \mathrm{st}\left(\int \varphi(x)^* f(x)\, dx\right) \qquad (f \in C_0). \tag{16.1}$$

We shall prove now that every linear functional on C_0 or on a linear subspace $U \subseteq C_0$ has this form, even with a particular φ:

Let C_0^∞ denote the system of all functions $f : \mathbb{R} \to \mathbb{R}$ which are differentiable arbitrarily many times and which have a compact support. It is well-known that C_0^∞ is not trivial: It contains e.g. the "bump" function

$$\Phi(x) := \begin{cases} e^{-1/(1-x^2)} & \text{if } |x| < 1, \\ 0 & \text{if } |x| \geq 1. \end{cases}$$

Theorem 16.2. *If $*$ is a $\mathcal{P}(\mathbb{N})$-enlargement, then any linear functional F on a linear subspace $U \subseteq C_0$ can be written in the form*

$$^*(F(f)) = \int \varphi(x)^* f(x)\, dx \qquad (f \in U) \tag{16.2}$$

with some $\varphi \in L \cap {}^*C_0^\infty$. In particular, (16.1) *holds.*

Proof. Using standard linear algebra (and the axiom of choice) one can extend F to a linear (not necessarily bounded) functional on C_0. Hence, without loss of generality, we can assume $U = C_0$.

The essential step is to prove that the binary relation

$$\psi := \{(\underline{x}, \underline{y}) \in C_0 \times C_0^\infty \mid \int \underline{y}(t)\underline{x}(t)\, dt = F(\underline{x})\}$$

is concurrent: To see this, let $f_1, \dots, f_n \in C_0$ be given. We have to prove that there is some $\varphi \in C_0^\infty$ such that

$$\int \varphi(t) f_k(t)\, dt = F(f_k) \qquad (k = 1, \dots, n). \tag{16.3}$$

It is no loss of generality to assume that the functions f_k are linearly independent (since all expressions in (16.3) are linear, we may successively eliminate those f_i which are linear combinations of the remaining ones).

We prove by induction on n that for any linearly independent $f_1, \dots, f_n \in C_0$ there are $\varphi_1, \dots, \varphi_n \in C_0^\infty$ satisfying

$$\int \varphi_k(t) f_j(t)\, dt = \begin{cases} 1 & \text{if } k = j, \\ 0 & \text{if } k \neq j. \end{cases} \tag{16.4}$$

Then $\varphi := F(f_1)\varphi_1 + \cdots + F(f_n)\varphi_n$ has the required properties. Assume that the claim has already been proved for $n-1$. Under this assumption, we will construct for any given linearly independent $f_1, \dots, f_n \in C_0$ and any given $k \in \{1, \dots, n\}$ a function $\varphi_k \in C_0^\infty$ which satisfies (16.4) for $j = 1, \dots, k$. Then the induction step is complete. Renumbering the functions f_j if necessary, it suffices to describe this construction for the case $k = n$.

By induction hypothesis, we find functions $\psi_1, \dots, \psi_{n-1} \in C_0^\infty$ with

$$\int \psi_k(t) f_j(t)\, dt = \begin{cases} 1 & \text{if } k = j \text{ and } j = 1, \dots, n-1, \\ 0 & \text{if } k \neq j \text{ and } k, j = 1, \dots, n-1. \end{cases}$$

Since $f_1, \dots, f_n$ are linearly independent, the function

$$f := f_n - \sum_{j=1}^{n-1} \left(\int \psi_j(t) f_n(t)\, dt \right) f_j$$

does not vanish. In particular, there is a nonempty interval on which f is either strictly positive or strictly negative. Choose a nontrivial nonnegative function of C_0^∞ which has its support in this interval (such a function can be obtained by rescaling the "bump" function described above). Multiplying this function with a constant, one finds some $\Phi \in C_0^\infty$ such that

$$\int \Phi(t) f(t)\, dt = 1.$$

Putting $\lambda_k := \int \Phi(t) f_k(t)\, dt$ $(k = 1, \dots, n-1)$, we claim that the function

$$\varphi_n := \Phi - \sum_{k=1}^{n-1} \lambda_k \psi_k$$

has the required properties. Indeed, for $j = 1, \ldots, n-1$, we have

$$\int \varphi_n(t) f_j(t)\, dt = \int \Phi(t) f_j(t)\, dt - \lambda_j = 0.$$

Moreover, the definitions of φ_n, λ_j, and f imply

$$\begin{aligned}
\int \varphi_n(t) f_n(t)\, dt &= \int \left(\Phi(t) - \sum_{j=1}^{n-1} \lambda_j \psi_j(t) \right) f_n(t)\, dt \\
&= \int \Phi(s) f_n(s)\, ds - \sum_{j=1}^{n-1} \left(\int \Phi(s) f_j(s)\, ds \right) \left(\int \psi_j(t) f_n(t)\, dt \right) \\
&= \int \Phi(s) \left(f_n(s) - \sum_{j=1}^{n-1} \left(\int \psi_j(t) f_n(t)\, dt \right) f_j(s) \right) ds \\
&= \int \Phi(s) f(s)\, ds = 1,
\end{aligned}$$

as desired.

Note that to determine a function from C_0, it suffices to determine it for rational arguments. Hence, C_0 has the cardinality of $\mathbb{R}^{\mathbb{Q}}$ which in turn has the cardinality $|\mathbb{R}^{\mathbb{N}}| = |(2^{\mathbb{N}})^{\mathbb{N}}| = |2^{\mathbb{N}\times\mathbb{N}}| = |2^{\mathbb{N}}| = |\mathcal{P}(\mathbb{N})|$. Since $*$ is a $\mathcal{P}(\mathbb{N})$-enlargement, we have that ${}^*\psi$ is satisfied on ${}^\sigma C_0$, i.e. we find some $\varphi \in {}^*C_0^\infty$ with $({}^*f, \varphi) \in {}^*\psi$ for any $f \in C_0$. By the standard definition principle, this means

$$\int \varphi(t)\, {}^*f(t)\, dt = {}^*F({}^*f) = {}^*(F(f)) \qquad (f \in C_0).$$

Hence (16.2) holds, and since ${}^*(F(f))$ is always finite, this implies also $\varphi \in L$. □

Considering the functional $F(f) := f(0)$, we thus find indeed a function $\delta \in L \cap {}^*C_0^\infty$ satisfying

$$\int \delta(x)\, {}^*f(x)\, dx = {}^*f(0) \qquad (f \in C_0).$$

If we replace here $=$ by $\approx$, it may even be arranged that $\delta(x) = 0$ for $x \not\approx 0$. Moreover, we do not even need that $*$ is an enlargement to find such a function:

In fact, any locally integrable internal function δ with the following properties will do:

1. $\delta \geq 0$.
2. There is some infinitesimal $c > 0$ with $\delta(x) = 0$ for $|x| > c$.
3. $\displaystyle\int \delta(t)\, dt = 1$.

If $*$ is a nonstandard map, we find such a function: Choose some $f \in C_0^\infty$ with $f \geq 0$, $\int f(t)\,dt = 1$ and $f(x) = 0$ for $|x| > 1$. Then $\delta(x) := c^{-1*}f(c^{-1}x)$ has the required properties for any infinitesimal $c > 0$: The identity $\int \delta(x)\,dx = 1$ follows by the substitution rule.

Exercise 79. Prove that any function δ with the above properties satisfies

$$\int \delta(x)^* f(x)\,dx \approx {}^*f(0) \qquad (f \in C_0),$$

i.e. for $\varphi = \delta$ the functional (16.1) satisfies $F_\varphi(f) = f(0)$ $(f \in C_0)$.

Appendix A

Some Important $*$-Values

Theorem A.1. *Let* $* : \widehat{S} \to \widehat{{}^*S}$ *be elementary. Let* S_n *and* T_n *denote the level sets of the superstructure* $\widehat{S}$ *and* $\widehat{{}^*S}$*, respectively, i.e.* S_n *is as in* Section 2.1. *Then*

$$ {}^*S_n = \{x \in T_n : x \textit{ is internal}\} \qquad (n = 0, 1, 2, \ldots). \tag{A.1} $$

Proof. The proof is by induction on n: For $n = 0$, we have ${}^*S_0 = {}^*S = T_0$. If (A.1) is already proved for some n, recall that $S_{n+1} = S_0 \cup \mathcal{P}(S_n)$, and so ${}^*S_{n+1} = {}^*S_0 \cup {}^*\mathcal{P}(S_n)$. Thus, Theorem 3.21 implies

$$ {}^*S_{n+1} = {}^*S_0 \cup \{A \subseteq {}^*S_n : A \text{ is internal}\}. $$

Since by induction assumption ${}^*S_n \subseteq T_n$, this already implies

$$ {}^*S_{n+1} \subseteq T_0 \cup \mathcal{P}(T_n) = T_{n+1}, $$

and since all elements of ${}^*S_{n+1}$ are internal, we thus have even

$$ {}^*S_{n+1} \subseteq \{x \in T_{n+1} : x \text{ is internal}\}. $$

For the converse inclusion, observe that if $A \in T_{n+1} \backslash T_0$ is internal, then each $x \in A$ is internal, because $\mathscr{I}$ is transitive; moreover, since $A \in T_{n+1} = T_0 \cup \mathcal{P}(T_n)$, we must have $x \in T_n$ which by induction assumption implies $x \in {}^*S_n$. Consequently,

$$ \{A \in T_{n+1} \setminus T_0 : A \text{ is internal}\} \subseteq \{A \subseteq {}^*S_n : A \text{ is internal}\}. $$

Since $T_0 = {}^*S_0$, we thus have

$$ \{A \in T_{n+1} : A \text{ is internal}\} \subseteq {}^*S_0 \cup \{A \subseteq {}^*S_n : A \text{ is internal}\} = {}^*S_{n+1}, $$

which proves the desired converse inclusion. $\square$

We intend to prove now, roughly speaking, that

$$\bigcup {}^*\mathscr{A} = {}^*\Big(\bigcup \mathscr{A}\Big).$$

However, since we work with a set theory with atoms, some care is needed:

Proposition A.2. *Let S_n and T_n be as in Theorem* A.1.

Given an entity $A \in \widehat{S}$, let A_n be the collection of all elements of $x \in A$ which are of type n, i.e. $x \in S_n \setminus S_{n-1}$. Then

$${}^*A_n = \{\underline{x} \in {}^*A : \underline{x} \in T_n \setminus T_{n-1}\},$$

i.e. ${}^*A_n$ contains all elements from ${}^*A$ which are of type n. Here, we put $S_{-1} = T_{-1} = \varnothing$.

Proof. We have

$$A_n = \{\underline{x} \in A : \underline{x} \in S_n \setminus S_{n-1}\}.$$

The standard definition principle implies

$${}^*A_n = \{\underline{x} \in {}^*A : \underline{x} \in {}^*S_n \setminus {}^*S_{n-1}\}.$$

Since Theorem A.1 implies that ${}^*S_n \setminus {}^*S_{n-1}$ contains all internal elements of $T_n \setminus T_{n-1}$, the statement follows. □

Corollary A.3. *Let $x \in \widehat{{}^*S}$ be internal, and of type n. Then there is some entity $A \in \widehat{S}$ with $x \in {}^*A$ such that all elements of A are of type n.*

In particular, any internal entity is contained in a set ${}^*A$ where A is a set consisting of entities.

Proof. Since x is internal, there is an entity $B \in \widehat{S}$ with $x \in {}^*B$. Let A be the collection of all elements $y \in B$ of type n. By Proposition A.2, we have $y \in {}^*A$. □

Theorem A.4. *Let $\mathscr{A} \in \widehat{S}$ be an entity. Let $\mathscr{A}_0$ be the collection of all elements of $\mathscr{A}$ which are entities. Then*

$${}^*\Big(\bigcup \mathscr{A}_0\Big) = \bigcup \{A : A \in {}^*\mathscr{A} \textit{ is an entity}\}$$

and

$${}^*\Big(\bigcap \mathscr{A}_0\Big) = \bigcap \{A : A \in {}^*\mathscr{A} \textit{ is an entity}\}.$$

Proof. Since entities are the elements of at least type 1, Proposition A.2 implies that ${}^*\mathscr{A}_0 = \{A \in {}^*\mathscr{A} : A \text{ entity}\}$. Hence, it is no loss of generality to assume that $\mathscr{A} = \mathscr{A}_0$.

Put $U := \bigcup \mathscr{A}$. The transitively bounded sentence

$$\forall \underline{x} \in U : \exists \underline{y} \in \mathscr{A} : \underline{x} \in \underline{y}$$

is true, and so the transfer principle implies

$$\forall \underline{x} \in {}^*U : \exists \underline{y} \in {}^*\mathscr{A} : \underline{x} \in \underline{y},$$

i.e. ${}^*U \subseteq \bigcup {}^*\mathscr{A}$. Conversely, the transfer of

$$\forall \underline{x} \in \mathscr{A} : \forall \underline{y} \in \underline{x} : \underline{y} \in U$$

implies analogously $\bigcup {}^*\mathscr{A} \subseteq {}^*U$. This proves the first equality. For the second equality, put $D := \bigcap \mathscr{A}$, and observe that

$$D = \{\underline{x} \in U \mid \forall \underline{y} \in \mathscr{A} : \underline{x} \in \underline{y}\}.$$

The standard definition principle shows

$${}^*D = \{\underline{x} \in {}^*U \mid \forall \underline{y} \in {}^*\mathscr{A} : \underline{x} \in \underline{y}\}.$$

Since, ${}^*U = \bigcup {}^*\mathscr{A}$ (as we just have proved), we find ${}^*D = \bigcap {}^*\mathscr{A}$. □

Corollary A.5. *Let $\mathscr{A}, \mathscr{B} \in \widehat{S}$ be sets of entities. Then*

$$\begin{aligned} {}^*\{A \cup B : A \in \mathscr{A},\ B \in \mathscr{B}\} &= \{A \cup B : A \in {}^*\mathscr{A},\ B \in {}^*\mathscr{B}\}, \\ {}^*\{A \cap B : A \in \mathscr{A},\ B \in \mathscr{B}\} &= \{A \cap B : A \in {}^*\mathscr{A},\ B \in {}^*\mathscr{B}\}, \end{aligned}$$

and

$${}^*\{A \setminus B : A \in \mathscr{A},\ B \in \mathscr{B}\} = \{A \setminus B : A \in {}^*\mathscr{A},\ B \in {}^*\mathscr{B}\}.$$

Proof. Put $\mathscr{C} := \{A \cup B : A \in \mathscr{A}, B \in \mathscr{B}\}$, and $U := \bigcup(\mathscr{A} \cup \mathscr{B})$. Then

$$\mathscr{C} = \{\underline{z} \in U \mid \exists \underline{x} \in \mathscr{A}, \underline{y} \in \mathscr{B} : \underline{z} = \underline{x} \cup \underline{y}\}.$$

The standard definition principle thus implies that

$${}^*\mathscr{C} = \{\underline{z} \in {}^*U \mid \exists \underline{x} \in {}^*\mathscr{A}, \underline{y} \in {}^*\mathscr{B} : \underline{z} = \underline{x} \cup \underline{y}\}.$$

Since Theorem A.4 implies ${}^*U = \bigcup({}^*\mathscr{A} \cup {}^*\mathscr{B}) = \bigcup {}^*(\mathscr{A} \cup \mathscr{B})$, the first statement follows. The proof of the other formulas is analogous (choose the same set U as above). □

Although Theorem A.4 is rather satisfactory from a theoretical point of view, one often considers unions of indexed sets:

If X_i $(i \in I)$ is a family of sets, one would like to describe the $*$-value of $\bigcup_{i \in I} X_i$. A natural conjecture is that this value is $\bigcup_{i \in {}^*I} {}^*X_i$. To make this more precise, we have to define what we mean by ${}^*X_i$ when $i \in {}^*I$:

Let X_i, I and $\mathscr{X} = \{X_i : i \in I\}$ all be entities of $\widehat{S}$. Then we may define a bijection $f : I \to \mathscr{X}$ by $f(i) = X_i$. Then ${}^*f : {}^*I \to {}^*\mathscr{X}$. In a slight misuse of notation, we define ${}^*X_i := {}^*f(i)$ $(i \in {}^*I)$. The reader should not confuse ${}^*X_i$ with ${}^*(X_i)$. However, there is no real danger of such confusion, since for $i \in I$ we have ${}^*X_{{}^*i} = {}^*f({}^*i) = {}^*(f(i)) = {}^*(X_i)$.

Corollary A.6. *With the above notation, we have*

$${}^*\left(\bigcup_{i\in I} X_i\right) = \bigcup_{i\in {}^*I} {}^*X_i.$$

Proof. The functions $f : I \to \mathscr{X}$ and ${}^*f : {}^*I \to {}^*\mathscr{X}$ are both onto (Theorem 3.13). Hence, $\bigcup_{i\in {}^*I} X_i = \bigcup \mathscr{X}$ and $\bigcup_{i\in {}^*I} {}^*X_i = \bigcup {}^*\mathscr{X}$. Thus, the statement follows from Theorem A.4 □

Definition A.7. With the above notation, one defines the *Cartesian product*

$$\prod_{i\in I} X_i := \{f : I \to \bigcup_{i\in I} X_i \mid f(i) \in X_i \text{ for each } i \in I\}.$$

Exercise 80. Prove that

$${}^*\left(\prod_{i\in I} X_i\right) = \{x \in \prod_{i\in {}^*I} {}^*X_i \mid x \text{ internal}\}.$$

Exercise 81. Show that for each internal entity X the system of all internal subsets of X is an internal entity.

Exercise 82. Show that for each pair of internal entities A, B the system F of all internal functions $f : A \to B$ is an internal entity. Is also the system $\mathscr{F}$ consisting of all internal functions defined on subsets of A with values in B an internal entity?

We obtain similarly also results for sets of higher type:

Theorem A.8. *Let $\mathscr{A} \in \widehat{S}$ be a system of entities. Then*

$${}^*\{\mathcal{P}(A) : A \in \mathscr{A}\} = \{P_A : A \in {}^*\mathscr{A}\},$$

where P_A denotes the system of all internal subsets of A.

Proof. Let S_n, T_n be as in Theorem A.1, and put $P := \{\mathcal{P}(A) : A \in \mathscr{A}\}$. We find some n with $\mathscr{A}, P \in S_n$. Then $P \in S_{n+2}$, and since each S_n is transitive, we have

$$P = \{\underline{x} \in S_{n+2} \mid \exists \underline{y} \in \mathscr{A} : \forall \underline{z} \in S_{n+1} : (\underline{z} \in \underline{x} \iff \underline{z} \subseteq \underline{y})\}.$$

By the standard definition principle, we find

$${}^*P = \{\underline{x} \in {}^*S_{n+2} \mid \exists \underline{y} \in {}^*\mathscr{A} : \forall \underline{z} \in {}^*S_{n+1} : (\underline{z} \in \underline{x} \iff \underline{z} \subseteq \underline{y})\}.$$

For fixed $y \in {}^*\mathscr{A}$, the following holds: Since ${}^*\mathscr{A} \in {}^*S_n \subseteq T_n$ and T_n is transitive, any internal $z \subseteq y$ satisfies $z \subseteq T_n$, and so $z \in T_{n+1}$. Since z is internal, we find $z \in {}^*S_{n+1}$ by Theorem A.1. In other words: If $z \subseteq y$ is internal, then $z \in {}^*S_{n+1}$. Conversely, each $z \in {}^*S_{n+1}$ is internal. We thus have

$$ {}^*P = \{P_A \in {}^*S_{n+2} \mid A \in {}^*\mathscr{A}\}. $$

It remains to prove that $P_A \in {}^*S_{n+1}$ for any $A \in {}^*\mathscr{A}$. Since ${}^*\mathscr{A} \in {}^*S_n \subseteq T_n$ and since T_n is transitive, we have for any $A \in {}^*\mathscr{A}$ that $A \subseteq T_n$, and so $P_A \subseteq \mathcal{P}(A) \in T_{n+1}$. Since T_{n+1} is transitive, this proves $P_A \subseteq T_{n+1}$, and so $P_A \in T_{n+2}$. Since P_A is internal by Exercise 81, we thus have $P_A \in {}^*S_{n+2}$ by Theorem A.1, as claimed. □

Exercise 83. (Cumbersome). Let $A, B \in \widehat{S}$, and let $\mathscr{F}(A, B)$ denote the system of all functions f with $\mathrm{dom}(f) \subseteq A$ and $\mathrm{rng}(f) \subseteq B$. Prove that

$$ {}^*\mathscr{F}(A, B) = \{f \in \mathscr{F}({}^*A, {}^*B) : f \text{ is internal}\}. $$

Exercise 84. Let $\mathscr{A}, \mathscr{B} \in \widehat{S}$ be systems of entities. Prove that

$$ {}^*\{A \times B : A \in \mathscr{A}, B \in \mathscr{B}\} = \{A \times B : A \in {}^*\mathscr{A},\ B \in {}^*\mathscr{B}\}. $$

Appendix B

Solutions to the Exercises

The following solutions to the exercises are not complete, but all important ideas are sketched.

Exercise 1: X is not Dedekind complete: Let $f(t) := \sqrt{t}$, and $A \subseteq X$ consist of all $x \in X$ such that there is some $t_0 > 0$ with $x(t) \leq f(t)$ for $t > t_0$. Then A is bounded (e.g. by $x(t) := t$), but A has no maximal element. To see this, we will prove that any $x \in A$ satisfies $S := \limsup_{t\to\infty} x(t) < \infty$. Then the constant function $y(t) := S + 1$ belongs to A and is strictly larger than x.

We have for sufficiently large t that $x(t^2) \leq f(t^2) = t$, and so the rational function $y(t) = x(t^2)/t$ satisfies $\limsup_{t\to\infty} y(t) \leq 1$. Since $y(t) = p(t)/q(t)$ with polynomials p, q this means that either $y(t) \to -\infty$ or that the degree $\deg p$ of the polynomial p is at most as large as the degree $\deg q$ of q. In the first case, $x(t) \leq 0$ for sufficiently large t. In the second case, observe that $\deg p$ is even and $\deg q$ is odd by the definition of y. Hence, $\deg p \leq \deg q - 1$ which implies that $x(t^2) = ty(t)$ converges as $t \to \infty$. Hence, $\limsup_{t\to\infty} x(t) < \infty$ in both cases, as claimed.

X is not Archimedean: The function $x(t) := t$ belongs to X and satisfies $x \geq n$ for any $n \in \mathbb{N}$. $\star$

Exercise 2: The set is not a field, since $x_{0,1}$ has no inverse with respect to multiplication. It is also not Dedekind complete, since the set $\{x_{0,b} : b > 0\}$ has no least upper bound. X is Archimedean, for if $x_{a,b} \in X$, then we find some $n \in \mathbb{N}$ with $n > a$, and so $x_{n,0} \in X_{\mathbb{N}}$ satisfies $x_{n,0} > x_{a,b}$. $\star$

Exercise 3: It is not totally ordered, since e.g. the equivalence class of the sequence $x : n \mapsto (-1)^n$ cannot be compared with 0. It is not a field, because there is some $[x] \in X$ such that $[x] \neq 0$ and x contains infinitely many 0's (e.g. $x : n \mapsto$

$1+(-1)^n$); then $[x]$ is not invertible. X is not Dedekind complete: Let A be the set of all equivalence classes of sequences converging to 0. Then A is bounded from above (e.g. by 1). However, A has no least upper bound: If $[x] \in X$ were such a bound, then $[x] > 0$ (because the equivalence class of $n \mapsto n^{-1}$ belongs to X). Note that $[y] \in A$ implies $\lambda[y] \in A$ for any $\lambda \in \mathbb{R}$. Hence, in the case $[x] \in A$, we find $2[x] \in A$, and so $[x]$ is no upper bound. But if $[x] \notin A$, then $[x]/2$ provides a strictly smaller upper bound for A. In both cases, we found a contradiction to the fact that $[x]$ is the smallest upper bound. X fails to be Archimedean, since for the sequence $x : n \mapsto n$, we even have $[x] > n$ for any n. ⋆

Exercise 4: The function $x(t) = t$ is an example of an "infinitesimal". With the exception of Dedekind completeness, the answers to the other questions are always "no" (with a similar proof as in the previous exercise). However, the space X *is* Dedekind complete. (In fact, an analogous statement holds for each σ-finite measure space. This observation goes back to Riesz and Kantorovich, see e.g. [Zaa67] for a general proof). Let $M \subseteq X$ be bounded from above. We have to show that M has a least upper bound. If M is countable, the least upper bound is evidently $s(t) := \sup\{x(t) : x \in M\}$ where we identify x with *some* function of its equivalence class (hence, s is only defined up to equivalence classes). However, if M is uncountable, s would not be well-defined in this way. To avoid this difficulty, we assume in addition that the functions in M are uniformly bounded. This is no loss of generality, because we can replace M by $\{\arctan\circ x : x \in M\}$; and if y is a least upper bound for this set, then $\tan\circ y$ is a least upper bound for M. Now, if the functions in M are uniformly bounded, then for each countable subset $A \subseteq M$ the least upper bound $x_A := \sup A$ is integrable, and the set $\{\int_0^1 x_A(t)\,dt : A \subseteq M \text{ countable}\}$ is bounded and thus has a least upper bound S. Choose a sequence of countable sets $A_n \subseteq M$ such that $\int_0^1 x_{A_n}(t)\,dt \to S$. The union A_∞ of these sets A_n is countable, and so $x_{A_\infty} = \sup A_\infty$ satisfies $\int_0^1 x_{A_\infty}(t)\,dt = S$. Moreover, for each $x \in M$ we must have $x \le x_{A_\infty}$ almost everywhere, since otherwise

$$\int_0^1 x_{A_\infty\cup\{x\}}(t)\,dt > \int_0^1 x_{A_\infty}(t)\,dt = S,$$

contradicting the definition of S. Hence, x_{A_∞} is an upper bound for M, and in view of $A_\infty \subseteq M$, it must be the least upper bound for M. ⋆

Exercise 5: If such a predicate would exist, then $\{c\} = \{\underline{x} \in {}^*A : \alpha(\underline{x})\}$ would be a standard entity by the standard definition principle. But this means that there is some standard entity B with $\{c\} = {}^*B$. Since ${}^*B \subseteq {}^*A$, we have $B \subseteq A$ (Lemma 3.5), and so ${}^\sigma B \subseteq {}^\sigma A$. By Corollary 3.11, we have ${}^\sigma B \subseteq {}^*B = \{c\}$.

Since ${}^{\sigma}B \subseteq {}^{\sigma}A$ and $c \notin {}^{\sigma}A$, this implies $B = \varnothing$, and Lemma 3.5 then gives the contradiction $\{c\} = {}^*B = {}^*\varnothing = \varnothing$. ⋆

Exercise 6: By Proposition 3.16 and Lemma 3.14, we find some index k with $x_1, \dots, x_k \in {}^*S_n$. Then

$$\{x_1, \dots, x_n\} = \{\underline{x} \in {}^*S_k : \underline{x} = x_1 \vee \cdots \vee \underline{x} = x_n\}$$

and

$$(\underline{x}_1, \dots, \underline{x}_n) = \{(\underline{x}_1, \dots, \underline{x}_n) \in {}^*S_k \times \cdots \times {}^*S_k : \underline{x}_1 = x_1 \wedge \cdots \wedge \underline{x}_n = x_n\}$$

are internal by the internal definition principle resp. by the internal definition principle for relations. If B is an internal entity, and $A \subseteq B$ is external, then A must be infinite, since otherwise $A = \{x_1, \dots, x_n\}$ with $x_i \in B$ would be internal by what we just proved (and since each x_i is internal, because $\mathscr{I}$ is transitive). ⋆

Exercise 7: We have

$$g \circ f = \{(\underline{x}, \underline{y}) \in \operatorname{dom}(f) \times \operatorname{rng}(g) \mid \exists \underline{z} \in \operatorname{dom}(g) : (\underline{x}, \underline{z}) \in f \wedge (\underline{z}, \underline{y}) \in g\}.$$

This set is internal by the internal definition principle for relations. ⋆

Exercise 8: 1. By the internal definition principle, the sets

$$g_i := \{\underline{z} \in f_i \mid \exists \underline{x} \in A_i, \underline{y} \in \operatorname{rng}(f_i) : \underline{z} = (\underline{x}, \underline{y})\}$$

are internal. Hence, $f = g_1 \cup \cdots \cup g_n$ is internal.
2. Fix some $b \in B$, and let $f_1 : A \to B$ be the constant function $f_1(x) = b$, i.e. $f_1 = A \times \{b\}$. Then f_1 is internal by the internal definition principle. Now define $F : A \to B$ by $F(x) = f(x)$ for $x \in A_0$ and $F(x) = f_1(x)$ for $x \in A \setminus A_0$. By 1., it follows that F is internal (observe that $A_0 = \operatorname{dom}(f)$ and $A \setminus A_0$ are internal by Theorem 3.19).
3. Let $f : A \to B$ and $A_0 \subseteq A$ be internal. Then

$$f|_{A_0} = \{(\underline{x}, \underline{y}) \in f : \underline{x} \in A_0\}$$

is internal by the internal definition principle for relations. ⋆

Exercise 9: The answer is negative: Let A be some external set. Then the set $\mathscr{A} = \{A\}$ cannot be a subset of some internal set B, since then we would have $A \in B$, and so A is internal because $\mathscr{I}$ is transitive (Proposition 3.16). ⋆

Exercise 10: If $\mathscr{U}$ has the above form, then $j_0 \in \bigcap \mathscr{U}$, and so $\mathscr{U}$ is not free. Conversely, if $\mathscr{U}$ is not free, then there is some $j_0 \in J$ with $j_0 \in \bigcap \mathscr{U}$. For any $U \subseteq J$ precisely one of the sets U and $J \setminus U$ belongs to $\mathscr{U}$, because $\mathscr{U}$ is an ultrafilter. If $j_0 \notin U$, then U cannot belong to $\mathscr{U}$ by our choice of j_0. Conversely, if $j_0 \in \mathscr{U}$, then $j_0 \notin J \setminus U$, and so $J \setminus U$ cannot belong to $\mathscr{U}$; but this implies $U \in \mathscr{U}$. Hence, $\mathscr{U}$ contains precisely those sets $U \subseteq J$ with $j_0 \in \mathscr{U}$. ⋆

Exercise 11: If $\mathscr{U}$ contains the filter of Example 4.2, then we have for any $j_0 \in J$ that $J \setminus \{j_0\} \in \mathscr{U}$; hence $\mathscr{U}$ is free. Conversely, assume that $\mathscr{U}$ is free but that $J \setminus J_0 \notin \mathscr{U}$ for some finite set $J_0 \subseteq J$. Since $\mathscr{U}$ is an ultrafilter, we must have $J_0 \in \mathscr{U}$. For each $j \in J_0$ there is some $U_j \in \mathscr{U}$ such that $j \notin U_j$ (otherwise $\mathscr{U}$ would not be free). The finite intersection $J_0 \cap \bigcap_{j \in J_0} U_j$ belongs to $\mathscr{U}$, because $\mathscr{U}$ is a filter. But this intersection is empty, a contradiction. ⋆

Exercise 12: Note first that the relation $[f] \notin \mathscr{I}_0$ means that $f(j) \in S_0$ holds not almost everywhere. Since $\mathscr{U}$ is an ultrafilter, this means that $f(j) \in S_0$ holds almost nowhere. By choosing an appropriate representative f, we may thus assume that $f(j) \notin S_0$ for all j. Similarly, we may assume that $g(j) \notin S_0$ for all j.

Assume now that $[f] \neq [g]$, i.e. $f(j) = g(j)$ does not hold almost everywhere. Since $\mathscr{U}$ is an ultrafilter, we have $f(j) = g(j)$ almost nowhere, i.e. $f(j) \neq g(j)$ for all $j \in U$ where $U \in \mathscr{U}$. Let U_1 denote the set of all $j \in U$ for which $f(j) \setminus g(j) \neq \varnothing$, and $U_2 := U \setminus U_1$. For all $j \in U_2$, we have $g(j) \setminus f(j) \neq \varnothing$. Since $\mathscr{U}$ is an ultrafilter, one of the sets U_1 or U_2 must belong to $\mathscr{U}$. Without loss of generality, we assume that $U_1 \in \mathscr{U}$. For all $j \in U_1$, we find a function $h : J \to \widehat{S}$ such that $h(j) \in f(j) \setminus g(j)$. Then $[h] \in_{\mathscr{U}} [f]$ but not $[h] \in_{\mathscr{U}} [g]$, a contradiction to (4.3). ⋆

Exercise 13: The condition is $f_x(j) = f_y(j)$ for all except at most finitely many j, i.e. the set $D = \{j : f_x(j) = f_y(j)\}$ has a finite complement.

Indeed, if $J \setminus D$ is finite, then we have $D \in \mathscr{U}$ for any δ-incomplete ultrafilter (because $\mathscr{U}$ is free by Corollary 4.13, and so $D \in \mathscr{U}$ by Exercise 11). Hence, $f_x(j) = f_y(j)$ holds for almost all j.

Conversely, if $J \setminus D$ is infinite, we "construct" a δ-incomplete ultrafilter $\mathscr{U}$ with $D \notin \mathscr{U}$ as follows: Let $\mathscr{F}_0$ denote the filter of Example 4.2, i.e. $F \in \mathscr{F}_0$ if and only if $J \setminus M$ is finite. The system $\mathscr{B} = \mathscr{F}_0 \cup \{J \setminus D\}$ has the finite intersection property, since for any $F \in \mathscr{F}_0$ the set $F \cap D$ is infinite (and in particular nonempty). Hence, $\mathscr{B}$ generates a filter $\mathscr{F}$. Let $\mathscr{U}$ be an ultrafilter containing $\mathscr{F}$ (Theorem 4.9). Exercise 11 implies that $\mathscr{U}$ is free. Hence, $\mathscr{U}$ is δ-incomplete by Proposition 4.11. Since $J \setminus D \in \mathscr{B} \subseteq \mathscr{U}$, we have $D \notin \mathscr{U}$, as claimed. Thus, $f_x(j) \neq f_y(j)$ for almost all j, and so $x \neq y$. ⋆

Exercise 14: 1. Let $A = \{a_1, \ldots, a_n\}$. By Proposition 4.19, we have ${}^*A = \varphi([F])$ where $F : J \to \widehat{S}$ is the constant function $F(j) := A$. By Proposition 4.19 and

Theorem 4.18, we have $x \in {}^*A$ if and only if $x = \varphi([f])$ where $[f] \in_{\mathscr{U}} [F]$. Of course, the constant functions $f_i(j) := a_i$ $(i = 1, \ldots, n)$ satisfy $[f_i] \in_{\mathscr{U}} [F]$, and so $\varphi([f_i]) \in {}^*A$, i.e. ${}^*a_i \in {}^*A$ (by Proposition 4.19, we have ${}^*a_i = \varphi([f_i])$). It remains to prove that any other function f satisfying $[f] \in_{\mathscr{U}} [F]$ must satisfy $[f] =_{\mathscr{U}} [f_i]$ for some $i \in \{1, \ldots, n\}$, i.e. that $M_i := \{j : f(j) = i\}$ belongs to $\mathscr{U}$ for some i. But since J is the disjoint union of the sets $M_1, \ldots, M_n$ and since $\mathscr{U}$ is an ultrafilter, precisely one of these sets must belong to $\mathscr{U}$: Otherwise their complements would belong to $\mathscr{U}$ and thus also the intersection of these complements which is empty, a contradiction to $\varnothing \notin \mathscr{U}$.
2. As in 1., we have $x \in {}^*A$ if and only if $x = \varphi([f])$ where $f : J \to A$. So we have to prove that there are uncountable many equivalence classes of functions $f : J \to A$, if A is infinite. Assume that $[f_1], [f_2], \ldots$ is an enumeration of *all* such equivalence classes. Since $\mathscr{U}$ is δ-incomplete, we find by Proposition 4.12 countably many pairwise disjoint $J_1, J_2, \ldots \subseteq J$ with $\bigcup J_n = J$. Define now $f : J \to A$ such that $f(j) \in A \setminus (f_1(j) \cup \cdots \cup f_n(j))$ for $j \in J_n$ (this is possible, since A is infinite). Then $f(j) = f_n(j)$ holds at most on the set $J_1 \cup \cdots \cup J_n$, i.e. almost nowhere. Hence, $[f] \neq [f_n]$ for all n, and so the equivalence class of $f : J \to A$ is not contained in the given enumeration. Hence, the set of all equivalence classes is not countable. $\star$

Exercise 15: The first statement follows immediately from the transfer principle. If x is finite, then $n \geq |x| \geq h = |h|$, and so h is finite and belongs to ${}^\sigma\mathbb{N}$ by Proposition 5.9. If x is infinite, then $h > |x| - 1 \geq n - 1$ for each $n \in {}^\sigma\mathbb{N}$ which is not possible for $h \in {}^\sigma\mathbb{N}$. $\star$

Exercise 16: The equation $x^2 = 2$ has no solution: This follows by the transfer principle from $\forall \underline{x} \in \mathbb{Q} : \underline{x}^2 \neq 2$. However, there is some $x \in {}^*\mathbb{Q}$ with $x^2 \approx 2$: We have

$$\forall \underline{n} \in \mathbb{N} : \exists \underline{x} \in \mathbb{Q} : |\underline{x}^2 - 2| \leq \underline{n}^{-1},$$

and the transfer principle implies

$$\forall \underline{n} \in {}^*\mathbb{N} : \exists \underline{x} \in {}^*\mathbb{Q} : |\underline{x}^2 - 2| \leq \underline{n}^{-1}.$$

Fixing some $h \in \mathbb{N}_\infty$, we thus find some $x \in {}^*\mathbb{Q}$ with $|x^2 - 2| \leq h^{-1}$, i.e. $x^2 \approx 2$. $\star$

Exercise 17: Consider the transitively bounded sentence

$$\forall \underline{x}, \underline{y} \in \mathbb{R} : (\underline{y} > 0 \implies \exists \underline{z} \in \mathbb{Q} : |\underline{x} - \underline{z}| < \underline{y}).$$

Since $\mathbb{Q}$ is dense in $\mathbb{R}$, this sentence is true, and so the transfer principle implies

$$\forall \underline{x}, \underline{y} \in {}^*\mathbb{R} : (\underline{y} > 0 \implies \exists \underline{z} \in {}^*\mathbb{Q} : |\underline{x} - \underline{z}| < \underline{y}).$$

Hence, for each $x \in {}^*\mathbb{R}$ and some infinitesimal $y > 0$ (even for all), we find some $z \in {}^*\mathbb{Q}$ with $|x - z| < y$, i.e. $x \approx z$. ⋆

Exercise 18: If this set (denote it by M_x) were internal, then $\inf({}^*\mathbb{R}) = M_x - x$ would be internal by the internal definition principle. ⋆

Exercise 19: The statement follows by transfer of

$$\forall \underline{x} \in \mathbb{N} : ((\exists \underline{n} \in \mathbb{N} : \underline{x} = 2\underline{n}) \iff \neg(\exists \underline{n} \in \mathbb{N} : \underline{x} = 2\underline{n} - 1)).$$

Concerning the map $*$ of Theorem 4.20, the elements of ${}^*\mathbb{N}$ correspond to equivalence classes of functions $f : J \to \mathbb{N}$ (recall Proposition 4.19). Either the set

$$M = \{j : f(j) \text{ is even}\}$$

or its complement belongs to $\mathscr{U}$. The first case occurs if and only if f may for almost all j be written in the form $f(j) = 2g(j)$ where $g : J \to \mathbb{N}$, and the second case occurs if and only if f may for almost all j be written in the form $f(j) = 2g(j) - 1$ where $g : J \to \mathbb{N}$. Thus, the statement follows in view of Example 5.6. ⋆

Exercise 20: We have $\lim_{j \to j_0} f(j) = x$ if and only if for any open neighborhood U of x we find some open neighborhood J_U of j_0 with $f(j) \in U$ for $j \in J_U \setminus \{j_0\}$. The latter means $f^{-1}(U) = \{j : f(j) \in U\} \in \mathscr{F}$ which by Lemma 5.27 is equivalent to $U \in f(\mathscr{F})$. Hence, $\lim_{j \to j_0} f(j) = x$ if and only if any open neighborhood U of x is contained in $f(\mathscr{F})$, i.e. if and only if $\lim_{j \to \mathscr{F}} f(j) = x$. ⋆

Exercise 21: 1. Let $\beta(\underline{x})$ be the internal formula $(\underline{x} \leq h_0 \implies \alpha(\underline{x}))$. Then $\beta(h)$ holds for all infinite $h \in {}^*\mathbb{N}_\infty$. By the permanence principle, we thus find some $n_0 \in {}^\sigma\mathbb{N}$ such that $\beta(n)$ holds for all $n \in {}^*\mathbb{N}$ with $n_0 \leq n \leq h$. In particular, $\beta(n)$ and thus $\alpha(n)$ holds for any *finite* $n \in {}^\sigma\mathbb{N}$ with $n \geq n_0$.
2. The proof is analogously reduced to the permanence principle for $\mathbb{R}$: Let $\beta(\underline{x})$ denote the internal formula $(\underline{x} > c \implies \alpha(\underline{x}))$. Then $\beta(d)$ holds for all infinitesimals $d \in \inf({}^*\mathbb{R})$, $d > 0$. By the permanence principle for $\mathbb{R}$, we thus find some $\varepsilon_0 \in {}^\sigma\mathbb{R}_+$ such that $\beta(\varepsilon)$ holds for all standard or nonstandard $\varepsilon \in {}^*\mathbb{R}$ with $0 < \varepsilon \leq \varepsilon_0$. In particular, $\alpha(\varepsilon)$ holds for all $\varepsilon \in {}^*\mathbb{R}$ with $c < \varepsilon \leq \varepsilon_0$. ⋆

Exercise 22: Let $\alpha(\underline{n})$ be the internal formula $|x_{\underline{n}}| \leq \underline{n}^{-1}$. Then $\alpha(n)$ holds for all $n \in {}^\sigma\mathbb{N}$. By the permanence principle, there is some $h \in \mathbb{N}_\infty$ such that $\alpha(\underline{n})$ holds for all $\underline{n} \in {}^*\mathbb{N}$ with $\underline{n} \leq h$. In particular, $|x_n| \leq n^{-1}$ for all $n \in \mathbb{N}_\infty$ with $n \leq h$ which implies that $x_n \approx 0$ for those n. ⋆

Exercise 23: The assumptions imply in both cases in view of Theorem 3.19 that A and B are internal, i.e. there are sets $C, D \in \widehat{S}$ with $A \in {}^*C$ and $B \in {}^*D$. In view of Corollary A.3, we may assume that all elements of C and D are entities. Let $C_0 \subseteq C$ and $D_0 \subseteq D$ be the subsets of entities of C resp. D, and let $U := \bigcup C_0$ and $V := \bigcup D_0$. By Theorem 2.1, we have $U, V \in \widehat{S}$. Let $\mathscr{F}$ denote the system of all functions f with $\operatorname{dom}(f) \subseteq U$ and $\operatorname{rng}(f) \subseteq V$ and recall (Exercise 83) that ${}^*\mathscr{F}$ consists of all internal functions f with $\operatorname{dom}(f) \subseteq {}^*U$ and $\operatorname{rng}(f) \subseteq {}^*V$.

1. By the hint, the sentence

$$\forall \underline{x} \in \mathscr{F} : \exists \underline{y} \in \mathscr{F} : \operatorname{dom}(\underline{y}) = \operatorname{rng}(\underline{x}) \wedge \operatorname{rng}(\underline{y}) = \operatorname{dom}(\underline{y}) \wedge \text{“}\underline{y} \text{ is one-to-one”}$$

is true. Here, the shortcuts rng and dom use quantifiers over U and V. The $*$-transform implies the statement for the choice $\underline{x} = f$ and $g = \underline{y}$.

2. By the Schröder-Bernstein theorem, the sentence

$$\begin{aligned} \forall \underline{x}_1, \underline{x}_2 \in \mathscr{F} : ((\text{“}\underline{x}_1, \underline{x}_2 \text{ one-to-one”} \wedge \operatorname{rng}(\underline{x}_1) \subseteq \operatorname{dom}(\underline{x}_2) \wedge \operatorname{rng}(\underline{x}_2) \subseteq \operatorname{dom}(\underline{x}_1)) \\ \implies \exists \underline{y} \in \mathscr{F} : (\text{“}\underline{y} \text{ one-to-one”} \wedge \operatorname{dom}(\underline{y}) = \operatorname{dom}(\underline{x}_1) \wedge \operatorname{rng}(\underline{y}) = \operatorname{dom}(\underline{x}_2))) \end{aligned}$$

is true (as before, rng and dom use quantifiers over U and V). The $*$-transform implies the statement for the choice $\underline{x}_1 = f_1$, $\underline{x}_2 = f_2$, $g = \underline{y}$. ⋆

Exercise 24: The sets $I := A \cap B$, and $A_0 := A \setminus B$ are internal subsets of the *-finite set A and thus *-finite by Theorem 6.13. Moreover, $A = A_0 \cup I$ and $A \cup B = A_0 \cup B$ where the sets on the right-hand side are disjoint. Hence, Theorem 6.14 implies ${}^{\#}A = {}^{\#}A_0 + {}^{\#}I$ and ${}^{\#}(A \cup B) = {}^{\#}A_0 + {}^{\#}B$. Combining these two equations, the statement follows. ⋆

Exercise 25: We have

$$X^{<\mathbb{N}} = \{\underline{x} \in \mathcal{P}(\mathbb{N} \times X) \mid \exists \underline{n} \in \mathbb{N} : (\underline{x} : \{1, \ldots, \underline{n}\} \to X \wedge \operatorname{dom}(\underline{x}) \subseteq \{1, \ldots, \underline{n}\})\}$$

where $\underline{k} \notin \operatorname{dom}(\underline{x})$ can be formulated in a bounded form, if we use that $\underline{x} \subseteq \mathbb{N} \times X$. Note now that ${}^*\mathcal{P}(\mathbb{N} \times X)$ consists of all internal subsets of ${}^*(\mathbb{N} \times X) = {}^*\mathbb{N} \times {}^*X$. The standard definition principle implies

$$\begin{gathered} {}^*(X^{<\mathbb{N}}) = \{\underline{x} \in {}^*\mathcal{P}(\mathbb{N} \times X) \mid \exists \underline{n} \in {}^*\mathbb{N} : \\ (\underline{x} : \{1, \ldots, \underline{n}\} \to {}^*X \wedge \operatorname{dom}(\underline{x}) \subseteq \{1, \ldots, \underline{n}\})\}, \end{gathered}$$

where $\operatorname{dom}(\underline{x}) \subseteq \{1, \ldots, \underline{n}\}$ has the expected meaning in view of the fact that we have $\underline{x} \subseteq {}^*\mathbb{N} \times {}^*X$. Thus, the first statement follows. For the second statement, note that

$$\begin{gathered} \#_X(\cdot) = \{\underline{z} \in \mathcal{P}(X^{<\mathbb{N}} \times \mathbb{N}) \mid \exists \underline{x} \in {}^*\mathcal{P}(\mathbb{N} \times X), \exists \underline{n} \in {}^*\mathbb{N} : \\ \underline{x} : \{1, \ldots, \underline{n}\} \to {}^*X \wedge \operatorname{dom}(\underline{x}) \subseteq \{1, \ldots, \underline{n}\} \wedge \underline{z} = (\underline{x}, \underline{n})\}, \end{gathered}$$

and apply similarly as above the standard definition principle, using the fact that we already proved that ${}^*(X^{<\mathbb{N}}) = {}^*X^{<{}^*\mathbb{N}}$ and that ${}^*\mathcal{P}(X^{<\mathbb{N}} \times \mathbb{N})$ consists of all internal subsets of ${}^*(X^{<\mathbb{N}}) \times \mathbb{N} = {}^*X^{<{}^*\mathbb{N}} \times {}^*\mathbb{N}$. ⋆

Exercise 26: There is some n such that R and the order relation on R both belong to ${}^*S_n$. By Proposition 3.16, there is some n such that $R \in {}^*S_n$. Consider the sentence

$$\forall \underline{x} \in S_n, \underline{y} \in S_n, \underline{z} \in \mathcal{P}(S_n) : (\alpha(\underline{x}, \underline{y}, \underline{z}) \implies \exists \underline{w} \in \mathcal{P}(\mathcal{P}(S_n) \times S_n) : \beta(\underline{x}, \underline{y}, \underline{z}, \underline{w}))$$

where $\alpha(\underline{x}, \underline{y}, \underline{z}, S_n)$ is a transitively bounded sentence with the meaning "$\underline{y}$ is a total order on $\underline{x}$, and the elements of $\underline{z}$ are finite subsets of $\underline{x}$" and $\beta(\underline{x}, \underline{y}, \underline{z}, \underline{w})$ means

$$\underline{w} : \underline{z} \to \underline{x} \wedge \forall \underline{u} \in \underline{z} : \underline{w}(\underline{u}) = \max\{\underline{u}\},$$

where the order $\underline{y}$ is used to define $\max\{\underline{u}\}$. Note that in order to formalize α, it suffices to quantify over S_n and $\mathbb{N}$. The $*$-transform of α then becomes "$\underline{y}$ is a total order on $\underline{x}$, and the elements of $\underline{z}$ are *-finite subsets of $\underline{x}$". Hence the $*$-transform of the above true sentence implies the statement. ⋆

Exercise 27: Put $U := \bigcup \mathscr{A}$. Then ${}^*U = \bigcup {}^*\mathscr{A}$ by Theorem A.4. Using the notation of Exercise 25, we have

$$\forall \underline{x} \in \mathscr{A} : (c(\underline{x}) = \infty \dot\vee \exists \underline{y} \in U^{<\mathbb{N}} : (\text{“}\underline{y} \text{ is a bijection onto } \underline{x}\text{”} \wedge c(\underline{x}) = \#_U(\underline{y})))$$

Hence, in view of Exercise 25,

$$\forall \underline{x} \in {}^*\mathscr{A} : ({}^*c(\underline{x}) = \infty \dot\vee \exists \underline{y} \in {}^*U^{<{}^*\mathbb{N}} :$$
$$(\text{“}\underline{y} \text{ is a bijection onto } \underline{x}\text{”} \wedge {}^*c(\underline{x}) = \#_{{}^*U}(\underline{y}))).$$

This implies that for any $\underline{x} = B \in {}^*\mathscr{A}$ one of the following alternatives holds: Either ${}^*c(B) = \infty$, or ${}^*c(\underline{x}) = h$ where $y : \{1, \ldots, h\} \to B$ is an internal bijection. But this means ${}^*c(B) = {}^\#B$. For the second statement, note that A is finite if and only if ${}^{\#*}A = {}^*c({}^*A) = {}^*(c(A)) \neq \infty$. ⋆

Exercise 28: If the sequence x_n is bounded, all ${}^*x_h$ with $h \in \mathbb{N}_\infty$ are finite by Theorem 7.2. Corollary 7.3 thus shows that the sets of accumulation points is given by $\{\operatorname{st}({}^*x_h) : h \in \mathbb{N}_\infty\}$. The supremum and infimum of this set are $\limsup x_n$ and $\liminf x_n$, respectively. If the sequence is unbounded, ${}^*x_h$ is infinite for some $h \in \mathbb{N}_\infty$, and so $\operatorname{st}({}^*x_h)$ is not defined. However, if one chooses the natural notation $\operatorname{st}({}^*x_h) = \pm\infty$ if ${}^*x_h$ is infinite and $\pm{}^*x_h > 0$, then the formula still holds: Indeed, $\pm\infty$ is an accumulation point of x_n if and only if $\operatorname{st}({}^*x_n) = \pm\infty$ for some n: For positive sign, the proof is analogous to Theorem 7.2 (just drop the absolute values in the proof). For negative sign, the proof is similar or may be reduced to the case of positive sign by considering the sequence $-x_n$. ⋆

Exercise 29: (One of many solutions): Put $x_n := (-1)^n$. Then ${}^*x_h = 1$ if h is even, and ${}^*x_h = -1$ if h is odd (recall Exercise 19). For the map $*$ of Theorem 4.20, the map $h \mapsto {}^*x_h$ associates to each equivalence class $[f_h]$ with $f_h : J \to \mathbb{N}$ the value ^{*1} if the set $\{j : f_h(j) \text{ even}\}$ belongs to $\mathscr{U}$, and ^{*0} otherwise. ⋆

Exercise 30: If x_n is a Cauchy sequence, then we find for any $\varepsilon \in \mathbb{R}_+$ some n_0 such that, in view of the transfer principle,

$$\forall \underline{n}, \underline{m} \in {}^*\mathbb{N} : (\underline{n}, \underline{m} \geq n_0 \implies |{}^*x_{\underline{n}} - {}^*x_{\underline{m}}| \leq {}^*\varepsilon).$$

In particular, we have for any $h, k \in \mathbb{N}_\infty$ that $|{}^*x_h - {}^*x_k| < \varepsilon$. Since $\varepsilon > 0$ is arbitrary here, it follows that ${}^*x_h \approx {}^*x_k$.

Conversely, if ${}^*x_h \approx {}^*x_k$ for each $h, k \in \mathbb{N}_\infty$, we have for any $\varepsilon \in \mathbb{R}_+$ that the internal predicate

$$\forall \underline{n}, \underline{m} \in {}^*\mathbb{N} : (\underline{n}, \underline{m} \geq \underline{y} \implies |{}^*x_{\underline{n}} - {}^*x_{\underline{m}}| < {}^*\varepsilon)$$

holds for all $\underline{y} \in \mathbb{N}_\infty$. By the permanence principle, it also holds for some $\underline{y} \in {}^\sigma\mathbb{N}$, i.e. for some $\underline{y} = {}^*n_0$ with $n_0 \in \mathbb{N}$. The converse direction of the transfer principle implies that x_n is a Cauchy sequence. ⋆

Exercise 31: If x is an interior point of A, we find some $\varepsilon \in \mathbb{R}_+$ such that

$$\forall \underline{y} \in \mathbb{R} : (|x - \underline{y}| < \varepsilon \implies \underline{y} \in A).$$

The transfer principle implies that ${}^*A$ contains all points $y \in {}^*\mathbb{R}$ which satisfy $|x - y| < {}^*\varepsilon$, in particular all points with $y \approx x$. Conversely, if $\operatorname{mon}(x) \subseteq {}^*A$, then the internal predicate

$$\forall \underline{y} \in {}^*\mathbb{R} : (|x - \underline{y}| < \underline{\varepsilon} \implies \underline{y} \in {}^*A)$$

holds for any $\underline{\varepsilon} \in \inf({}^*\mathbb{R})$, $\underline{\varepsilon} > 0$. By the Cauchy principle (permanence principle), the predicate holds also for some standard $\underline{\varepsilon} = {}^*\varepsilon$, $\varepsilon \in \mathbb{R}_+$. The inverse form of the transfer principle implies that A contains all $y \in \mathbb{R}$ with $|x - y| < \varepsilon$, i.e. x is an interior point of A. ⋆

Exercise 32: Evidently, $A = \varnothing$ and $A = \mathbb{R}$ have this property. One might suspect that all open sets have this property, but actually $A = \varnothing$ and $A = \mathbb{R}$ are the only sets: Assume that A has this property. Then the internal formula

$$\forall \underline{x} \in {}^*A, \underline{y} \in {}^*\mathbb{R} : (|\underline{x} - \underline{y}| \leq \underline{z} \implies \underline{y} \in {}^*A)$$

holds for any $\underline{z} \in \inf({}^*\mathbb{R})$, $\underline{z} > 0$. By the Cauchy principle (permanence principle), this formula also holds for some $\underline{z} = {}^*\varepsilon$ with $\varepsilon \in \mathbb{R}_+$. The inverse direction of the transfer principle implies

$$\forall \underline{x} \in A, \underline{y} \in \mathbb{R} : (|\underline{x} - \underline{y}| \leq \varepsilon \implies \underline{y} \in A).$$

Thus, if A contains some point x, then an induction by n shows that A contains the intervals $[x - n\varepsilon, x + n\varepsilon]$ which means that $A = \mathbb{R}$.

There is no contradiction to Exercise 31, because not any $x \in {}^*A$ is infinitely close to some standard point (otherwise A would be compact, but compact sets are not open unless $A = \varnothing$). ⋆

Exercise 33: A must be the empty set: Indeed, by Exercise 32, we must have either $A = \varnothing$ or $A = \mathbb{R}$. But the union of all monads is $\mathrm{fin}({}^*\mathbb{R}) \neq {}^*\mathbb{R}$, so that $A = \mathbb{R}$ is not possible. ⋆

Exercise 34: These are precisely those sets which have no accumulation points, or, equivalently, those sets A which have the property that the intersection of A with any bounded set is finite (this equivalence follows immediately from the Bolzano-Weierstraß theorem (Corollary 7.5)).

Indeed, if $x \in \mathbb{R}$ is an accumulation point of A, choose a sequence $x_n \in A$, $x_n \neq x$, such that $x_n \to x$. Then ${}^*x_h \approx {}^*x$ for some $h \in \mathbb{N}_\infty$ by Theorem 7.1. The transfer principle implies ${}^*x_h \neq {}^*x$. Hence, ${}^*x_h \in {}^*A$ is finite with ${}^*x_h \neq {}^*x = {}^*(\mathrm{st}({}^*x_h))$.

Conversely, assume that the intersection of A with any bounded set is finite. Then, for $K_n := \{\underline{x} \in \mathbb{R} : |\underline{x}| \leq n\}$, the set $A \cap K_n$ is finite, and so ${}^\sigma(A \cap K_n) = {}^*(A \cap K_n) = {}^*A \cap {}^*K_n$. By the standard definition principle, ${}^*K_n = \{\underline{x} \in {}^*\mathbb{R} : |\underline{x}| \leq {}^*n\}$. Consequently,

$$\{\underline{x} \in {}^*A : |\underline{x}| \leq {}^*n\} = \{{}^*\underline{y} : \underline{y} \in A \wedge |\underline{y}| \leq n\}.$$

Taking the union over all n, we then find for any $x \in {}^*A \cap \mathrm{fin}({}^*\mathbb{R})$ some $y \in A$ with $x = {}^*y$. Since then $y = \mathrm{st}(x)$, the set A has the required property. ⋆

Exercise 35: The set A is perfect if and only if for any $x \in A$ the monad $\mathrm{mon}(x)$ contains some point of ${}^*A$ which is different from ${}^*x$.

Indeed, if A is perfect and $x \in A$, we find a sequence $x_n \in A$ such that $x_n \neq x$ and $x_n \to x$. By Theorem 7.1, we have ${}^*x_h \approx {}^*x$ for some $h \in \mathbb{N}_\infty$. Hence ${}^*x_h \in \mathrm{mon}(x)$, and the transfer principle implies ${}^*x_h \in {}^*A$ and ${}^*x_h \neq {}^*x$. Conversely, assume that A is not perfect. Then there is some $x \in A$ and some $\varepsilon \in \mathbb{R}_+$ such that

$$\forall \underline{y} \in A : (\underline{y} \neq x \implies |x - \underline{y}| > \varepsilon).$$

The transfer principle implies that all $y \in {}^*A$ with $y \neq {}^*x$ satisfy $|{}^*x - y| > {}^*\varepsilon$ and so in particular $y \not\approx {}^*x$, i.e. $y \notin \mathrm{mon}(x)$. ⋆

Exercise 36: Note first that $x_n := a + n(b - a)/h$ is an internal sequence as a composition of internal functions (Exercise 7). Hence, the set $\{n \in {}^*\mathbb{N} : {}^*f(x_n) > {}^*c\}$ is an internal subset of ${}^*\mathbb{N}$ by the internal definition principle. Thus, it contains a smallest element by Theorem 5.12. ⋆

Exercise 37: Assume that $c := f'(0)$ exists. Then we have for any $0 \neq x \approx 0$ that $|x|/x \approx {}^*c$, i.e. ${}^*c \approx 1$ (for $x > 0$) and simultaneously ${}^*c \approx -1$ (for $x < 0$), a contradiction. ⋆

Exercise 38: Put $F(x) := f(x)g(x)$. Given $h \approx 0$, put $dx := h$, $df := {}^*f({}^*x_0 + dx) - {}^*f({}^*x_0)$, $dg := {}^*g({}^*x_0 + dx) - {}^*g({}^*x_0)$, and $dF := {}^*F({}^*x_0 + h) - {}^*F({}^*x_0)$. Then

$$df\,{}^*g({}^*x_0 + dx) + {}^*f({}^*x_0)dg = dF,$$

and so

$$\frac{dF}{dx} = \frac{df}{dx}{}^*g({}^*x_0 + dx) + {}^*f({}^*x_0)\frac{dg}{dx} \approx {}^*(f'(x_0)g(x_0) + f(x_0)g'(x_0)),$$

where we used Proposition 5.17 and the fact that ${}^*g({}^*x_0 + dx) \approx {}^*g({}^*x_0)$ which in turn follows from $dg = g'(x_0)dx \approx 0$. ⋆

Exercise 39: The statement follows from transfer of

$$\forall \underline{x}, \underline{y} \in [a,b] : (\underline{x} < \underline{y} \implies \exists \underline{z} \in [a,b] : \underline{x} < \underline{z} < \underline{y} \wedge \frac{f(\underline{x}) - f(\underline{y})}{\underline{x} - \underline{y}} = f'(\underline{z}))$$

which is true by the classical mean value theorem. ⋆

Exercise 41: By Proposition 4.11, $\mathscr{U}$ is δ-incomplete over $J := \mathbb{N}$. Hence, the corresponding ultrafilter model of Theorem 4.20 provides a nonstandard map $*$. An element $h \in \mathbb{N}_\infty$ in this model is given by $h := \varphi([h_0])$ where $h_0(n) := n$. Note that $g(n,x) := |[2^n x] - 2[2^{n-1}x]|$ defines a relation on $\mathbb{N}\times\mathbb{R}\times\mathbb{R}$. By Example 4.22, we have for $x \in \mathbb{R}$ and $y \in {}^*\mathbb{R}$, $y = \varphi([f_y])$, $f_y : J \to \mathbb{R}$ that

$$(h, {}^*x, y) \in {}^*g \iff (h_0(j), x, f_y(j)) \in g \text{ for almost all } j,$$

i.e. ${}^*g(h,x) = \varphi([F])$ where $F(j) := g(h_0(j), x)$, i.e. $F(n) = g(n,x)$. By Theorem 5.32, we thus have

$$f(x) = \operatorname{st}(g(h,x)) = \operatorname{st}({}^*g(h,x)) = \lim_{n\to\mathscr{U}} g(n,x),$$

where f is the function from Theorem 7.28. Since this function attains only the values 0 and 1 and is nonmeasurable on $[0,1]$, the statement follows. ⋆

Exercise 42: Let $\mathscr{A}$ be a nonempty system of entities $A \in \widehat{S}$ which has the finite intersection property and at most the cardinality of κ. By Lemma 8.7, we may assume that $\mathscr{A}$ is an entity. Put $U := \bigcup \mathscr{A}$, and consider the relation

$$\varphi := \{(\underline{x}, \underline{y}) \in \mathscr{A} \times U \mid \underline{y} \in \underline{x}\},$$

i.e. $(A, \underline{y}) \in \varphi$ if and only if $\underline{y} \in A$. Note that $\varphi \in \widehat{S}$. Since $\mathscr{A}$ has the finite intersection property, φ is concurrent on $\mathscr{A}$, and so there is some b which satisfies ${}^*\varphi$ on ${}^\sigma\mathscr{A}$. By the standard definition principle for relations, we have

$$ {}^*\varphi = \{(\underline{x}, \underline{y}) \in {}^*\mathscr{A} \times {}^*U \mid \underline{y} \in \underline{x}\}. $$

Since b satisfies ${}^*\varphi$ on ${}^\sigma\mathscr{A}$, we have $b \in \bigcap {}^\sigma\mathscr{A}$. ⋆

Exercise 43: Since S is infinite, we may assume that $\mathbb{N} \in \widehat{S}$ is an entity. By Proposition 4.19, we have $x \in {}^*\mathbb{N}$ if and only if $x = \varphi([f])$ for some $f : J \to \mathbb{N}$. Since J is countable, the system of all functions $f : J \to \mathbb{N}$ has at most the cardinality of $\mathbb{N}^\mathbb{N}$ which is the cardinality of $\mathcal{P}(\mathbb{N})$ (this can be seen e.g. by the estimate $|\mathbb{N}^\mathbb{N}| \le |(2^\mathbb{N})^\mathbb{N}| = |2^{\mathbb{N}\times\mathbb{N}}| = |2^\mathbb{N}| = |\mathcal{P}(\mathbb{N})|$). Thus, ${}^*\mathbb{N}$ has at most the cardinality of $\mathcal{P}(\mathbb{N})$. Hence, Proposition 8.11 implies that $*$ is not a κ-enlargement when κ has a strictly larger cardinality than $\mathcal{P}(\mathbb{N})$. In particular, $*$ is not an enlargement. ⋆

Exercise 44: By Theorem 8.10, there is some *-finite $\mathscr{B} \subseteq {}^*\mathscr{A}$ with ${}^\sigma\mathscr{A} \subseteq \mathscr{B}$. Then ${}^*A_0 \subseteq \bigcup {}^\sigma\mathscr{A} \subseteq \bigcup \mathscr{B}$, and so

$$ \exists \underline{x} \in {}^*\mathcal{P}(\mathscr{A}) : (\text{“}\underline{x} \text{ is } {}^*\text{-finite”} \wedge {}^*A_0 \subseteq \bigcup \underline{x}). $$

The inverse form of the transfer principle implies that there is some finite $\underline{x} = \mathscr{A}_0 \in \mathcal{P}(\mathscr{A})$ with $A_0 \subseteq \bigcup \mathscr{A}_0$.

A completely different solution proceeds as follows: Let $\mathscr{B}$ denote the system of all sets of the form $A_0 \setminus A$ ($A \in \mathscr{A}$). If one cannot find a finite $\mathscr{A}_0 \subseteq \mathscr{A}$ with $A_0 \subseteq \bigcup \mathscr{A}_0$, then $\mathscr{B}$ has the finite intersection property, and so $\bigcap {}^\sigma\mathscr{B} \neq \varnothing$. This means ${}^*A_0 \setminus \bigcup {}^\sigma\mathscr{A} \neq \varnothing$, a contradiction to the assumption. ⋆

Exercise 45: By Theorem 8.10, there is some *-finite set R_0 with ${}^\sigma\mathbb{R} \subseteq R_0 \subseteq {}^*\mathbb{R}$. The transfer of the statement in the hint implies that

$$ \forall \underline{y} \subseteq {}^*\mathcal{P}(\mathbb{R}), \underline{\varepsilon} \in \mathbb{R}_+ : (\text{“}\underline{y}\ {}^*\text{-finite”} \implies \exists \underline{n} \in {}^*\mathbb{N} : \forall \underline{x} \in \underline{y} : {}^*\alpha(\underline{x}, \underline{n}, \underline{\varepsilon})) $$

where ${}^*\alpha(\underline{x}, \underline{n}, \underline{\varepsilon})$ is a shortcut for

$$ \exists \underline{z} \in {}^*\mathbb{Z} : |\underline{n}\underline{x} - \underline{z}| < \underline{\varepsilon}. $$

Let $c \in \inf({}^*\mathbb{R})$, $c > 0$. By the transfer principle, we find some $\varepsilon \in {}^*\mathbb{R}_+$ such that

$$ \forall \underline{x} \in {}^*\mathbb{R} : ((\exists \underline{n} \in {}^*\mathbb{N} : {}^*\alpha(\underline{x}, \underline{n}, \varepsilon)) \implies |\sin(\pi \underline{n}\underline{x})| < c). $$

Applying the former sentence with this $\underline{\varepsilon} = \varepsilon$ and $\underline{y} = R_0$, we find some $h = \underline{n} \in {}^*\mathbb{N}$ such that $|\sin(\pi h x)| < c$ for all $x \in R_0$. This implies $\sin(\pi h x) \approx 0$ for all $x \in {}^\sigma\mathbb{R}$. ⋆

Exercise 46: Necessity is proved as in Theorem 8.12. For sufficiency, let $\mathscr{B}$ be an internal nonempty system of entities with the finite intersection property and at most the cardinality of κ. Observe that $U := \bigcup \mathscr{B}$ and $P := \mathscr{B} \times U$ are internal by Theorem 3.19. Consider the relation

$$\varphi := \{\underline{z} \in P \mid \exists \underline{x} \in \mathscr{B} : \exists \underline{y} \in \underline{x} : \underline{z} = (\underline{x}, \underline{y})\}$$

which is internal by the internal definition principle. Note that $\operatorname{dom}(\varphi) \subseteq \mathscr{B}$ has at most the cardinality of κ. We have $(B, \underline{y}) \in \varphi$ for some $B \in \mathscr{B}$ if and only if $\underline{y} \in B$. Since $\mathscr{B}$ has the finite intersection property, the relation φ is concurrent on $\mathscr{B}$. Hence, it is satisfied on $\mathscr{B}$, i.e. $\bigcap \mathscr{B} \neq \varnothing$. $\star$

Exercise 47: Assume contrary that $A \setminus \bigcup \mathscr{A}_0 \neq \varnothing$ for each finite $\mathscr{A}_0 \neq \varnothing$. Let $\mathscr{B}$ be the system of all sets of the form $A_0 \setminus A$ with $A \in \mathscr{A}$. Then $\mathscr{B}$ has the finite intersection property, and so $\bigcap \mathscr{B} \neq \varnothing$. But this means that $A \not\subseteq \bigcup \mathscr{A}$, a contradiction. $\star$

Exercise 49: The proof is analogous to Theorem 10.1 with p in place of $\|\cdot\|$. The only difference is in the proof of the analogue to Lemma 10.2: We have to prove that there is a constant $c \in \mathbb{R}$ such that the functional $F(x_0 + \lambda x_1) := f(x_0) + \lambda c$ $(x \in X_0,\ \lambda \in \mathbb{R})$ satisfies $F(x_0 + \lambda x_1) \le p(x_0 + \lambda x_1)$, i.e. $f(x_0) + \lambda c \le p(x_0 + \lambda x_1)$. In the case $\lambda > 0$, we may divide by λ and need the estimate $f(x) + c \le p(x + x_1)$ $(x \in X_0)$. In case $\lambda < 0$, we divide by $-\lambda$ and need $f(x) - c \le p(x - x_1)$ $(x \in X_0)$. The case $\lambda = 0$ is trivial. Thus, we have to find some c satisfying

$$f(x) - p(x + x_1) \le c \le p(x + x_1) - f(x) \qquad (x \in X_0).$$

Since for $x, y \in X_0$ the relation

$$f(x) + f(y) = f(x + y) \le p(x + y) = p((x + x_1) + (y - x_1)) \le p(x + x_1) + p(y - x_1)$$

holds, we find

$$\sup_{x \in X_0} (f(x) - p(x + x_1)) \le \inf_{y \in X_0} (p(y - x_1) - f(y)),$$

so that we may just choose some c in between these two quantities.

Theorem 10.1 is for $\mathbb{K} = \mathbb{R}$ a special case of the above result, because if $f \in X_0^*$, we may put $p(x) := \|f\|\,\|x\|$, and then find some linear $F : X \to \mathbb{R}$ extending f such that $F(x) \le p(x)$, i.e. $F(x) \le \|f\|\,\|x\|$ $(x \in X)$. Replacing x by $-x$, we find also $-F(x) = F(-x) \le \|f\|\,\|-x\| = \|f\|\,\|x\|$, so that we actually have $|F(x)| \le \|f\|\,\|x\|$, i.e. $\|F\| \le \|f\|$, as required. $\star$

Exercise 50: Assume $C = \{x_1, x_2, \ldots\}$. We may successively extend f to a linear functional F_n on the subspace $U_n := \operatorname{span}(X_0 \cup \{x_1, \ldots, x_n\})$ such that F_n is an extension of F_{n-1} with $\|F_n\| = \|f\|$ (trivial induction by n). Put $U := \bigcup U_n$, and $F_0(x) := F_n(x)$ for $x \in U_n$. Clearly, U is a subspace, and F_0 is well-defined and linear and satisfies $|F_0(u)| \le \|f\| \, \|u\|$ $(u \in U)$.

Given some $x \in X$ choose a sequence $u_n \in U$ with $u_n \to x$ (such a sequence exists, since $\overline{U} \supseteq \overline{C} = X$). Then u_n is a Cauchy sequence which implies that also $F_0(u_n)$ is a Cauchy sequence, because

$$|F_0(u_n) - F_0(u_j)| = |F_0(u_n - u_j)| \le \|f\| \, \|u_n - u_j\| \,.$$

Thus $F(u_n) \to c$ converges. Moreover, c is independent of the particular choice of the sequence u_n, since if $v_n \in U$ is another sequence with $v_n \to x$, then

$$|F_0(u_n) - F_0(v_n)| \le \|f\| \, \|u_n - v_n\| \to 0.$$

Thus, we may put $F(x) := c$. The function $F : X \to \mathbb{R}$ defined in this way is linear, because if $u_n, v_n \in U$ satisfy $u_n \to x$ and $v_n \to y$, then $w_n := u_n + v_n \in U$ converges to $x + y$, and $F_0(w_n) = F_0(u_n) + F_0(v_n) \to F(x) + F(y)$ which implies $F(x+y) = F(x) + F(y)$; similarly, $F_0(\lambda u_n) = \lambda F_0(u_n) \to \lambda F(x)$ implies $F(\lambda x) = \lambda F(x)$. Moreover, since $|F_0(u_n)| \le \|f\| \, \|u_n\| \to \|f\| \, \|x\|$, we find $|F(x)| \le \|f\| \, \|x\|$, i.e. $F \in X^*$ satisfies $\|F\| \le \|f\|$. ⋆

Exercise 51: Let $\mathscr{F}$ denote the system of all maps $A : X \to Y$ in the form (10.1) with finitely many $f_n \in X^*$ and $y_n \in Y$, i.e.

$$\mathscr{F} := \{\underline{z} \in Y^X \mid \exists \underline{y} \in Y^{<\mathbb{N}}, \underline{f} \in (X^*)^{<\mathbb{N}} : (\#_Y(\underline{y}) = \#_{X^*}(\underline{f}) \wedge \forall \underline{x} \in X : \underline{z}(\underline{x}) = \sum_{\underline{n}=1}^{\#_Y(\underline{y})} (\underline{f}(\underline{n}))(\underline{x})\underline{y}(\underline{n}))\}.$$

The standard definition principle implies in view of Theorem 3.21 and Exercise 25 that ${}^*\mathscr{F}$ consists of all internal maps $Z : {}^*X \to {}^*Y$ for which there are *-finite sequences $y_1, \ldots, y_h \in {}^*Y$ and $f_1, \ldots, f_h \in {}^*(X^*)$ such that

$$Z(x) = \sum_{n=1}^{h} f_n(x) y_n \qquad (x \in {}^*X).$$

Consider now the binary relation

$$\varphi := \{(\underline{x}, \underline{y}) \in X \times \mathscr{F} \mid \underline{y}(\underline{x}) = A(\underline{x})\}.$$

Then φ is concurrent on X: Given $x_1, \ldots, x_n \in X$, let X_0 be the linear hull of X. The restriction $A : X_0 \to Y$ may be written in the form (10.1), i.e. we find some

$F \in \mathscr{F}$ with $A(x) = F(x)$ for all $x \in X_0$. This means $(x_1, F), \dots, (x_n, F) \in \varphi$. Since $*$ is an X-enlargement, ${}^*\varphi$ is satisfied on ${}^\sigma X$, i.e. we find some $Z \in {}^*\mathscr{F}$ such that $({}^*x, Z) \in {}^*\varphi$ for any $x \in X$. The standard definition principle for relations implies that $Z({}^*x) = {}^*A({}^*x) = {}^*(A(x))$ for any $x \in X$. Since $Z \in {}^*\mathscr{F}$, A has the desired form. ⋆

Exercise 52: We first prove an analogue of Lemma 10.4: For any finite number of elements $x_1, \dots, x_K \in X$, $x_k = (\xi_{k,n})_n$ and any $\varepsilon > 0$ there exist real numbers $\eta_1, \dots, \eta_N \in \mathbb{R}$ such that

$$f(x_k) = \sum_{n=1}^{N} \eta_n \xi_{k,n} \qquad (k = 1, \dots, K).$$

Indeed, as in the proof of Lemma 10.4, we may assume that $x_1, \dots, x_K$ are linearly independent. Choose N such that the truncated vectors $y_k := (\xi_{k,1}, \dots, \xi_{k,N})$ are linearly independent. Defining

$$g(\lambda_1 y_1 + \cdots + \lambda_K y_K) := f(\lambda_1 x_1 + \cdots + \lambda_K y_K)$$

and extending g to $\mathbb{R}^N$, we find as in the proof of Lemma 10.4 the required numbers $\eta_n := g(e_n)$ where $e_1, \dots, e_N$ are the canonical base vectors of $\mathbb{R}^N$.

Consider now the relation

$$\varphi := \{(\underline{x}, \underline{y}) \in X \times \mathbb{R}^{<\mathbb{N}} \mid f(\underline{x}) = \sum_{\underline{n}=1}^{\#_{\mathbb{R}}(\underline{y})} \underline{y}(\underline{n})\underline{x}(\underline{n})\}.$$

By what we just proved, φ is concurrent on X. Since $*$ is an enlargement, Theorem 8.10 implies that there is some $y \in {}^*\mathbb{R}^{<\mathbb{N}}$ which satisfies ${}^*\varphi$ on ${}^\sigma\mathrm{dom}(\varphi)$, i.e. $({}^*x, y) \in {}^*\varphi$ for any $x \in X$. By Exercise 25, y is a *-finite internal sequence $\eta_1, \dots, \eta_h$ where $h := ({}^*\#_X(\cdot))(y)$. ⋆

Exercise 53: Assume there is some $y = (\eta_n)_n$ such that (10.2) is a Hahn-Banach limit. Given some n, consider the sequence defined by $\xi_k := 0$ $(k \neq n)$ and $\xi_n := 1$. Since f_y is a Hahn-Banach limit, its value for this sequence (which converges to 0) must be 0. On the other hand, by (10.2), this value must be η_n. Hence $\eta_n = 0$. Since this argument holds for any n, we would have $f_y(x) = 0$ for any $x \in \ell_\infty$ by (10.2), contradicting the fact that f_y is a Hahn-Banach limit.

Let $c \subseteq \ell_\infty$ be the subspace of all convergent sequences. For $x = (\xi_n)_n \in c$, define $f_0(x) = \lim \xi_n$. Then $f_0 \in c^*$ (with $\|f_0\| = 1$), and so we may use the Hahn-Banach extension theorem to extend f_0 to an element $f \in \ell_\infty^*$ which thus is a Hahn-Banach limit. ⋆

Exercise 54: Fixing some $h \in \mathbb{N}_\infty$, the functional $f((\xi_n)_n) := \mathrm{st}({}^*\xi_h)$ is such a Hahn-Banach limit. Indeed, Theorem 10.6 implies that f is in fact a Hahn-Banach limit (put $\eta_1 = \cdots = \eta_{h-1} = 0$ and $\eta_h = 1$). Moreover, $\mathrm{st}({}^*\xi_h)$ is always an accumulation point of a bounded sequence $x = (\xi_n)_n$ by Corollary 7.3 (note that ${}^*\xi_h$ is finite, because $|\xi_n| \le \|x\|_\infty$ implies by the transfer principle that $|{}^*\xi_h| \le {}^*(\|x\|_\infty)$). ⋆

Exercise 55: Given $x = (\xi_n)_n \in \ell_\infty$, put $l := \liminf \xi_n$ and $L := \limsup \xi_n$. Given $\varepsilon \in \mathbb{R}_+$, there is some $N \in \mathbb{N}$ such that $l - \varepsilon \le \xi_n \le L + \varepsilon$ holds for each $n \ge N$. Putting $\eta_n := \xi_n - (l - \varepsilon)$, we thus have $\eta_{n+N} \ge 0$, and so $0 \le f((\eta_{n+N})_n) = f((\eta_n)_n) = f(x) - (l - \varepsilon)$ which implies $f(x) \ge l - \varepsilon$. The estimate $f(x) \le L + \varepsilon$ follows analogously by putting $\eta_n := (L + \varepsilon) - \xi_n$ in view of $0 \le f((\eta_{n+N})_n) = f((\eta_n)_n) = L + \varepsilon - f(x)$. We thus have proved $l - \varepsilon \le f(x) \le L + \varepsilon$. Let $\varepsilon \to 0$ to obtain the estimate in the statement. This estimate then implies that f is a Hahn-Banach limit: On the one hand, this estimate implies that the functional f is bounded with $\|f\| \le 1$, because $|f(x)| \le \max\{|l|, |L|\} \le \|x\|_\infty$. On the other hand, if $\xi_n \to c$ converges, then $l = L = c$, and so $f(x) = c$. To see that even $\|f\| = 1$, observe that $f(x) = 1$ where x is the constant sequence $\xi_n = 1$. ⋆

Exercise 56: Since $(\xi_{n+1})_n = -x$, we must have $f(x) = f((\xi_{n+1})_n) = f(-x) = -f(x)$, and so $f(x) = 0$. Since 0 is not an accumulation point of the sequence x, a Banach-Mazur limit with the required properties does not exist. ⋆

Exercise 57: The set X_0 of all x for which ζ_n converges is a linear subspace of $X := \ell_\infty$. Since ζ_n depends linearly on x, it follows that f is linear and p is sublinear. Hence, there is a linear functional $F \in \ell_\infty^*$ which extends f and satisfies $F(x) \le p(x)$ for all $x \in \ell_\infty$. We claim that each such functional F is a Banach-Mazur limit: If x is the constant sequence $\xi_n = c$, then $\zeta_n = c$, and so $F(x) = f(x) = c$. The estimate $F(x) \le p(x)$ implies that F is positive: Indeed, $-F(x) = F(-x) \le p(-x)$ shows $F(x) \ge -p(-x)$. Thus, if $x = (\xi_n)_n$ with $\xi_n \ge 0$, we have $F(x) \ge -p(-x) \ge 0$, since the definition immediately implies $p(-x) \le 0$. Finally, F is shift invariant: If $x = (\xi_n)_n$ and $y = (\xi_{n+1})_n$, then

$$F(x) - F(y) = F(x - y) \le p(x - y) = \limsup_{n\to\infty} (\xi_1 - \xi_n)/n = 0,$$

and analogously

$$F(y) - F(x) = F(y - x) \le p(y - x) = \limsup_{n\to\infty} (\xi_n - \xi_1)/n = 0$$

which together implies $F(x) = F(y)$, as required. ⋆

Exercise 58: If x is almost convergent to c, then we find for any $\varepsilon \in \mathbb{R}_+$ some $n_0 \in \mathbb{N}$ such that, in view of the transfer principle,

$$\forall \underline{n} \in {}^*\mathbb{N} : \left(\underline{n} \geq {}^*n_0 \implies \forall \underline{k} \in {}^*\mathbb{N} : \left| \frac{1}{\underline{n}} \sum_{\underline{m}=1}^{\underline{n}} {}^*\xi_{\underline{m}+\underline{k}} - {}^*c \right| < {}^*\varepsilon \right).$$

Given $h \in \mathbb{N}_\infty$ and $k \in {}^*\mathbb{N}$, we thus find that

$$\left| \frac{1}{h} \sum_{n=1}^{h} {}^*\xi_{n+k} - {}^*c \right| < {}^*\varepsilon.$$

Since this holds for any $\varepsilon \in \mathbb{R}_+$, (10.10) follows.

Conversely, let (10.10) be satisfied. Given $\varepsilon \in \mathbb{R}_+$, the internal formula

$$\forall \underline{k} \in {}^*\mathbb{N} : \left| \frac{1}{\underline{n}} \sum_{\underline{m}=1}^{\underline{n}} {}^*\xi_{\underline{m}+\underline{k}} - {}^*c \right| < {}^*\varepsilon$$

holds for all $\underline{n} \in \mathbb{N}_\infty$. By the permanence principle, we find some $n_0 \in \mathbb{N}$ such that this formula holds for all $\underline{n} = {}^*n$ where $n \in \mathbb{N}$ satisfies $n \geq n_0$. By the inverse form of the transfer principle, we find

$$\forall \underline{k} \in \mathbb{N} : \left| \frac{1}{n} \sum_{\underline{m}=1}^{n} \xi_{\underline{m}+\underline{k}} - c \right| < \varepsilon$$

for any $n \in \mathbb{N}$ with $n \geq n_0$. But this means that x is almost convergent to c.

For the second statement, fix $h \in \mathbb{N}_\infty$. Applying (10.10) two times, we find that

$$\frac{1}{h} {}^*\xi_k \approx {}^*c - \frac{h-1}{h} \frac{1}{h-1} \sum_{n=1}^{h-1} {}^*\xi_{n+k+1} \approx {}^*c - \frac{h-1}{h} {}^*c = \frac{{}^*c}{h}.$$

Hence, $|{}^*\xi_k| / h \leq |{}^*c| / h + 1$ which implies $|{}^*\xi_k| \leq |{}^*c| + h$. In particular,

$$\exists \underline{n} \in {}^*\mathbb{N} : \forall \underline{k} \in {}^*\mathbb{N} : |{}^*\xi_{\underline{k}}| \leq |{}^*c| + \underline{n}.$$

The inverse form of the transfer principle implies that $(\xi_k)_k$ is bounded. ⋆

Exercise 59: The answer is negative: Choose A such that χ_A corresponds to the sequence $(0, 1, 0, 0, 1, 1, 0, 0, 0, 1, 1, 1, \ldots)$. Then A has the density $d = 1/2$, but the sequence $\chi_A = (a_n)_n$ is not almost convergent to $1/2$, because for any n we find some k such that $a_{1+k} = \cdots = a_{n+k} = 0$, and so $\frac{1}{n} \sum_{m=1}^{n} a_{m+k}$ cannot converge to $1/2$ uniformly in k. Theorem 10.13 implies that there is some Banach-Mazur limit f with $f(\chi_A) \neq 1/2$, and by Proposition 10.11, we find even a Banach-Mazur limit of Cesàro type with this property. ⋆

Exercise 60: If $\mathscr{U}$ is not an ultrafilter, then there is some filter $\mathscr{F} \supsetneq \mathscr{U}$ (Proposition 4.7). Then $\varnothing \neq \operatorname{mon}(\mathscr{F}) \subsetneq \operatorname{mon}(\mathscr{U})$ by Theorem 12.6, and so $\mathscr{F}$ satisfies neither $\operatorname{mon}(\mathscr{U}) \subseteq \operatorname{mon}(Fcl)$ nor $\operatorname{mon}(\mathscr{U}) \cap \operatorname{mon}(\mathscr{F}) = \varnothing$.

Conversely, let $\mathscr{U}$ be an ultrafilter and $\mathscr{F}$ be some filter with $\operatorname{mon}(\mathscr{U}) \cap \operatorname{mon}(\mathscr{F}) \neq \varnothing$. Then we have for any $U \in \mathscr{U}$, $F \in \mathscr{F}$ that ${}^*U \cap {}^*F \neq \varnothing$ and so $U \cap F \neq \varnothing$. Let $\mathscr{F}_0$ denote the system of all sets of the form $U \cap F$ with $U \in \mathscr{U}$, $F \in \mathscr{F}$. By what we proved, $\varnothing \neq \mathscr{F}_0$. It follows readily that $\mathscr{F}_0$ is a filter, and for the choice $F = X$ resp. $U = X$, we find $\mathscr{U}, \mathscr{F} \subseteq \mathscr{F}_0$. Since $\mathscr{U}$ is an ultrafilter, we have $\mathscr{U} = \mathscr{F}_0$, and so $\mathscr{F} \subseteq \mathscr{F}_0 = \mathscr{U}$, i.e. $\operatorname{mon}(\mathscr{U}) \subseteq \operatorname{mon}(\mathscr{F})$. Moreover, if $\mathscr{F} = \mathscr{U}'$ is an ultrafilter, then the relation $\mathscr{U}' = \mathscr{F} \subseteq \mathscr{F}_0$ implies $\mathscr{U}' = \mathscr{F}_0 = \mathscr{U}$. ⋆

Exercise 61: 1. Let $A \subseteq X$ be compact, and $A_0 \subseteq X$ be closed with $A_0 \subseteq A$. By Theorem 12.39, we have to prove that for each $y \in {}^*A_0$ there is some $x \in A_0$ with $y \approx_{\mathscr{O}} {}^*x$. Since A is compact and $y \in {}^*A_0 \subseteq {}^*A$, we find some $x \in A_0$ with $y \approx_{\mathscr{O}} {}^*x$, i.e. $x \in \operatorname{st}(y)$. But since A_0 is closed and $y \in {}^*A_0$, Theorem 12.35 implies $x \in A_0$.
2. If $A \subseteq X$ is compact, then each $y \in {}^*A$ is infinitely close to some $x \in {}^\sigma A$ by Theorem 12.39. Corollary 12.36 implies that A is closed. ⋆

Exercise 62: If A is compact, then Theorem 12.39 implies that $\operatorname{st}^{-1}(A) \cap {}^*A = {}^*A$ is standard and thus internal. Conversely, let $B := {}^*A \cap \operatorname{st}^{-1}(A)$ be internal, and let $\mathscr{C}$ be an open cover of A. For any $y \in B$ there is some $x \in A$ with $y \approx_{\mathscr{O}} {}^*x$. There is some $O \in \mathscr{C}$ with $x \in O$. Then $y \in {}^*O$, and so $y \in \bigcup {}^\sigma\mathscr{C}$. This proves $B \subseteq \bigcup {}^\sigma\mathscr{C}$. Theorem 8.16 implies that there is some finite $\mathscr{C}_0 \subseteq \mathscr{C}$ with $B \subseteq \bigcup {}^\sigma\mathscr{C}_0$. In particular, ${}^\sigma A \subseteq B \subseteq \bigcup {}^\sigma\mathscr{C}_0$ which implies $A \subseteq \bigcup \mathscr{C}_0$. Hence, A is compact. ⋆

Exercise 63: Let $\mathscr{C}$ be an open cover of $A_0 := \operatorname{st}(A)$. For any $a \in A$, we find in view of $A \subseteq \operatorname{ns}({}^*X)$ some $a_0 \in {}^*X$ with $a \approx_{\mathscr{O}} {}^*a_0$. We have $a_0 \in \operatorname{st}(a) \subseteq \operatorname{st}(A) = A_0$. Hence, there is some $O \in \mathscr{C}$ with $a_0 \in O$, and so $a \in {}^*O$. This proves $A \subseteq \bigcup {}^\sigma\mathscr{C}$. By Theorem 8.16, there is a finite $\mathscr{C}_0 \subseteq \mathscr{C}$, with $A \subseteq \bigcup {}^\sigma\mathscr{C}_0$. Since $\mathscr{C}_0 = \{O_1, \ldots, O_n\}$ is finite, we have

$$\bigcup {}^\sigma\mathscr{C}_0 = {}^*O_1 \cup \cdots \cup {}^*O_n = {}^*(O_1 \cup \cdots \cup O_n) = {}^*\Big(\bigcup \mathscr{C}_0\Big),$$

and so $A \subseteq {}^*(\bigcup \mathscr{C}_0)$. Theorem 12.35 thus implies $A_0 = \operatorname{st}(A) \subseteq \operatorname{st}\left({}^*(\bigcup \mathscr{C}_0)\right) = \overline{\bigcup \mathscr{C}_0}$, and so A_0 is compact by Lemma 12.43. ⋆

Exercise 64: Let X be compact, and $\mathscr{F}$ be an ultrafilter. Choose some $y \in \operatorname{mon}(\mathscr{F})$. Since X is compact, there is some $x \in X$ with $y \approx_{\mathscr{O}} {}^*x$, i.e. $y \in \operatorname{mon}(x) = \operatorname{mon}(\mathscr{U}(x))$. Hence, $\operatorname{mon}(\mathscr{F}) \cap \operatorname{mon}(\mathscr{U}(x)) \neq \varnothing$ which implies $\operatorname{mon}(\mathscr{F}) \subseteq \operatorname{mon}(\mathscr{U}(x)) = \operatorname{mon}(x)$ by Exercise 60. ⋆

Exercise 65: Let $f : X \to Y$, and $K \subseteq X$ be compact. We have to prove that $I := f(K)$ is compact. Let $y \in {}^*I = {}^*f({}^*K)$ (Theorem 3.13). Then we find some $x \in {}^*K$ with $y = {}^*f(x)$. Since K is compact, there is some $x_0 \in K$ with $x \approx_{\mathscr{O}} {}^*x_0$ (Theorem 12.39). Hence, Theorem 12.57 implies that $y = {}^*f({}^*x) \approx_{\mathscr{O}} {}^*f({}^*x_0) =: y_0$ Since $y_0 = {}^*(f(x_0)) \in {}^\sigma I$, Theorem 12.39 implies that I is compact. ⋆

Exercise 66: If $O \subseteq Y$ is open, then we have for any $y \in O$ that $\operatorname{mon}(y) \subseteq {}^*O$ (Theorem 12.34). If $x \in P := f^{-1}(O)$, then ${}^*f(\operatorname{mon}(x)) \subseteq \operatorname{mon}(f(x)) \subseteq {}^*O$ by Theorem 12.57. Hence, we have in view of Theorem 3.13 that $\operatorname{mon}(x) \subseteq ({}^*f)^{-1}({}^*O) = {}^*P$, and so P is open by Theorem 12.34. ⋆

Exercise 67: Given $U \in \mathscr{U}$, we find some $\varepsilon \in \mathbb{R}_+$ such that

$$B_\varepsilon := \{(\underline{x}, \underline{y}) \in X \times X \mid d(\underline{x}, \underline{y}) < \varepsilon\}$$

is contained in U. The standard definition principle for relations implies

$${}^*B_\varepsilon = \{(\underline{x}, \underline{y}) \in {}^*X \times {}^*X \mid {}^*d(\underline{x}, \underline{y}) < {}^*\varepsilon\}.$$

Hence, if we find for any $\varepsilon \in \mathbb{R}_+$ some $x \in {}^*X$ with ${}^*d({}^*x, y) < {}^*\varepsilon$, we find in particular for any $U \in \mathscr{U}$ some $x \in {}^*X$ with $({}^*x, y) \in {}^*B_\varepsilon \subseteq {}^*U$, and so $y \in \operatorname{pns}({}^*X)$. Conversely, if $y \in \operatorname{pns}({}^*X)$ and $\varepsilon \in \mathbb{R}_+$ are given, we find in view of $B_\varepsilon \in \mathscr{U}$ some $x \in X$ with $({}^*x, y) \in {}^*B_\varepsilon$ which means ${}^*d({}^*x, y) < {}^*\varepsilon$. ⋆

Exercise 68: Necessity has been proved in Corollary 13.13. For sufficiency, suppose that any Cauchy sequence converges. By Theorem 13.16, we have to prove that $\operatorname{pns}({}^*X) \subseteq \operatorname{ns}({}^*X)$. Thus, let $y \in \operatorname{pns}({}^*X)$ be given. By Exercise 67, we find for each n some $x_n \in X$ with ${}^*d({}^*(x_n), y) < 1/{}^*n$. Using the triangle inequality for ${}^*d$ (which holds by the transfer principle), we thus find for any $\varepsilon \in \mathbb{R}_+$ that ${}^*d({}^*(x_n), {}^*(x_m)) < {}^*\varepsilon$ for all $n, m \in \mathbb{N}$ with $n, m \geq 2/\varepsilon$. But since this estimate means by the inverse form of the transfer principle that $d(x_n, x_m) < \varepsilon$, we may conclude that x_n is a Cauchy sequence (x_n is a sequence by a countable form of the axiom of choice). By assumption, $x_n \to x$ for some $x \in X$.

We claim that $y \approx_{\mathscr{U}} {}^*x$. Given $\varepsilon \in \mathbb{R}_+$, we have to prove that ${}^*d({}^*x, y) < {}^*\varepsilon$ (Proposition 13.6). But choosing some $n \in \mathbb{N}$ with $n > 2/\varepsilon$ such that $d(x, x_n) < \varepsilon/2$, we find

$${}^*d({}^*x, y) \leq {}^*d({}^*x, {}^*(x_n)) + {}^*d({}^*(x_n), y) \leq {}^*\varepsilon/2 + 1/{}^*n = {}^*\varepsilon,$$

as desired. ⋆

Exercise 69: If $U \in \mathscr{U}$ and $V \subseteq Y \times Y$ satisfies $V \supseteq U \cap (Y \times Y)$, put $W := U \cup V$. Then $W \in \mathscr{U}$ (because $W \supseteq U$), and $V = W \cap (Y \times Y)$. Hence $V \in \mathscr{U}_Y$. If $U_1, U_2 \in \mathscr{U}$, then $U := U_1 \cap U_2 \in \mathscr{U}$, and thus $(U_1 \cap (Y \times Y)) \cap (U_2 \cap (Y \times Y)) =$

$U \cap (Y \times Y) \in \mathscr{U}_Y$. Moreover, $\Delta_Y := \{(y, y) : y \in Y\} \subseteq U \cap (Y \times Y)$ for any $U \in \mathscr{U}$, i.e. $\Delta_Y \subseteq V$ for any $V \in \mathscr{U}_Y$. In particular, $\mathscr{U}_Y$ is a filter which satisfies the first property of Definition 13.1.

If $U \in \mathscr{U}$, then $U^{-1} \in \mathscr{U}$, and so $(U \cap (Y \times Y))^{-1} = U^{-1} \cap (Y \times Y) \in \mathscr{U}_Y$. Moreover, there is some $V \in \mathscr{U}$ with $V^2 \subseteq U$. Then $W := V \cap (Y \times Y) \in \mathscr{U}$, and $W^2 \subseteq V^2 \subseteq U$ and $W^2 \subseteq (Y \times Y)^2 = Y \times Y$. Hence, $W^2 \subseteq U \cap (Y \times Y)$. ⋆

Exercise 70: Let X be complete, and $Y \subseteq X$ be closed with the inherited uniform structure. By Theorem 13.16, we have to prove that any $y \in \mathrm{pns}({}^*Y)$ belongs to $\mathrm{ns}({}^*Y)$. It follows from the definition (and Proposition 13.19) that $\mathrm{pns}({}^*Y) \subseteq \mathrm{pns}({}^*X)$. Since X is complete, Theorem 13.16 implies that there is some $x \in X$ with $y \approx_{\mathscr{O}} {}^*x$. Since Y is closed, we find in view of Theorem 12.35 and $y \in {}^*Y$ that $x \in Y$. We are done if we can prove that $y \in \mathrm{ns}({}^*Y)$, i.e. $y \in {}^*V({}^*x)$ for any $V \in \mathscr{U}_Y$. By definition, $V = U \cap (Y \times Y)$ for some $U \in \mathscr{U}$. Since $y \approx_{\mathscr{O}} {}^*x$, we have $(y, {}^*x) \in {}^*U \cap ({}^*Y \times {}^*Y) = {}^*V$, and so $y \in {}^*V({}^*x)$, as desired. ⋆

Exercise 71: If X is precompact, then Theorem 13.22 implies that $\mathrm{pns}({}^*X) = {}^*X$ is even standard. Conversely, if $\mathrm{pns}({}^*X)$ is internal, consider the set $\mathscr{A} := \{U(x) : x \in X\}$. For any $y \in \mathrm{pns}({}^*X)$, there is some $x \in X$ with $y \in {}^*U({}^*x) = {}^*(U(x))$, and so $\mathrm{pns}({}^*X) \subseteq \bigcup \mathscr{A}$. Since $\mathrm{pns}({}^*X)$ is internal, Theorem 8.16 implies that there is some finite $\mathscr{A}_0 \subseteq \mathscr{A}$ with $\mathrm{pns}({}^*X) \subseteq \bigcup {}^\sigma\mathscr{A}_0$. Hence, there are points $x_1, \ldots, x_n \in X$ with $\mathrm{pns}({}^*X) \subseteq {}^*(U(x_1)) \cup \cdots \cup (U(x_n))$. This implies $X \subseteq U(x_1) \cup \cdots \cup U(x_n)$, since for any $x \in X$, we have ${}^*x \in \mathrm{pns}({}^*X)$, and so we find some k with ${}^*x \in {}^*U(x_k)$, i.e. $x \in U(x_k)$. ⋆

Exercise 72: The only place in the proof of Theorem 13.26 where the saturation property was used is in the proof that $\bigcap \mathscr{A} \neq \varnothing$. However, if $*$ is comprehensive, we may extend the function $f : {}^\sigma\mathscr{U} \to {}^*X$, defined by $f({}^*U) := x_U$ $(U \in \mathscr{U})$ (axiom of choice!), to some internal function $F : {}^*\mathscr{U} \to {}^*X$. Consider the internal binary relation

$$\varphi := \{(\underline{x}, \underline{y}) \in {}^*\mathscr{U} \times {}^*X \mid \underline{y} \in \underline{x}(F(\underline{x}))\}.$$

For $U \in \mathscr{U}$, we have $({}^*U, y) \in \varphi$ if and only if $y \in {}^*U(x_U)$. Thus, since $\mathscr{A}$ has the finite intersection property, φ is concurrent on ${}^\sigma\mathscr{U}$. Since $*$ is a compact $\mathcal{P}(X)$-enlargement, φ is satisfied on ${}^\sigma\mathscr{U}$, i.e. there is some $y \in {}^*X$ with $({}^*U, y) \in \varphi$ for any $U \in \mathscr{U}$. But this means $y \in \bigcap \mathscr{A}$. ⋆

Exercise 73: Let $x \in \mathrm{fin}({}^*X)$ and $O \subseteq X$ be open with $0 \in O$. By Lemma 14.6, there is a balanced neighborhood $U \subseteq O$ of 0. We find some n with $x \in n{}^*U = {}^*(nU)$. Since U is balanced, we find by the transfer principle that

$$\forall \underline{x} \in {}^*(nU) : \forall \underline{y} \in {}^*\mathbb{R} : (\underline{y} \geq {}^*n \implies \underline{x} \in \underline{y}{}^*U).$$

This implies $x \in y^*U$ for any $y \in {}^*\mathbb{R}$ with $y \geq {}^*n$. Hence, $cx \in {}^*U \subseteq {}^*O$ for any $c \in {}^*\mathbb{R}_+$ with $c \leq 1/{}^*n$. In particular, $cx \in {}^*O$ for any $c \in \inf({}^*\mathbb{R})$ with $c > 0$.

Conversely, assume that $cx \in \inf({}^*X)$ for any $c \in \inf({}^*\mathbb{R})$, $c > 0$. If U is a neighborhood of 0, then the internal formula $\underline{\varepsilon}x \in {}^*O$ holds for any $\underline{\varepsilon} \in \inf({}^*\mathbb{R})$, $\underline{\varepsilon} > 0$. By the permanence principle (Cauchy principle), we have ${}^*\varepsilon x \in {}^*O$ for some $\varepsilon \in \mathbb{R}_+$, and so $x \in \lambda^*O$ for $\lambda := 1/\varepsilon$. $\star$

Exercise 74: If A is bounded and U is a neighborhood of 0, then there is some $n \in \mathbb{N}$ with $A \subseteq nU$, and so ${}^*A \subseteq n^*U$. In particular, for any $x \in {}^*A$ we have $x \in n^*U$, i.e. $x \in \operatorname{fin}({}^*X)$.

Conversely, let ${}^*A \subseteq \operatorname{fin}({}^*X)$ and a neighborhood U of 0 be given. By Exercise 73, the internal formula $\underline{\varepsilon}{}^*A \subseteq {}^*U$ holds for any $\underline{\varepsilon} \in \inf({}^*\mathbb{R})$, $\underline{\varepsilon} > 0$. By the permanence principle (Cauchy principle), we have ${}^*\varepsilon{}^*A \subseteq {}^*U$ for some $\varepsilon \in \mathbb{R}_+$. The inverse form of the transfer principle implies $A \subseteq \varepsilon^{-1}U$, and so A is bounded. $\star$

Exercise 75: Let $y \in \operatorname{pns}({}^*X)$, and U be a neighborhood of 0. By Theorem 14.2 and Lemma 14.6, we find a balanced neighborhood O of 0 with $O + O \subseteq U$. By Theorem 14.3, the set

$$U_O := \{(\underline{x}, \underline{y}) \in X \times X : \underline{x} - \underline{y} \in O\}$$

belongs to $\mathscr{U}$, and so we find some $x \in X$ with $({}^*x, y) \in {}^*U_O$ which by the standard definition principle for relations means ${}^*x - y \in {}^*O$. By Theorem 14.2, we find some $\lambda \neq 0$ with $\lambda x \in O$. Since O is balanced, it is no loss of generality to assume that $\lambda \in (0, 1)$. We have ${}^*x \in \lambda^{-1*}O$. Since ${}^*O$ is balanced and $\lambda \in (0, 1)$, we thus find

$$y \in \lambda^{-1*}O - {}^*O \subseteq \lambda^{-1*}O + \lambda^{-1*}O = \lambda^{-1}({}^*O + {}^*O) \subseteq \lambda^{-1*}U.$$

Hence, $y \in \operatorname{fin}({}^*X)$.

If $A \subseteq X$ is precompact, then Theorem 13.22 implies ${}^*A \subseteq \operatorname{pns}({}^*X) \subseteq \operatorname{fin}({}^*X)$ which means that A is bounded by Exercise 74. $\star$

Exercise 76: $\inf({}^*X)$ is a vector subspace of ${}^*X$ by Lemma 14.8. To see that $\inf({}^*X)$ is closed, let $x \notin \inf({}^*X)$. Then there is some neighborhood U of 0 with $x \notin {}^*U$. We find a balanced neighborhood O of 0 with $O + O \subseteq U$. Then $V = x + {}^*O$ is a neighborhood of x (Proposition 14.4), and $V \cap \inf({}^*X) = \varnothing$: Indeed, if $y \in V \cap \inf({}^*X)$, we have $y \in x + {}^*O$ and $y \in {}^*O$, and so we find $o_1, o_2 \in {}^*O$ with $o_1 = x + o_2$, i.e. $x = o_1 - o_2 \in {}^*O - {}^*O = {}^*O + {}^*O \subseteq {}^*U$, a contradiction. To prove $\inf({}^*X) \subseteq \operatorname{fin}({}^*X)$, let $x \in \inf({}^*X)$ be given. Then we have for any neighborhood U of 0 that $x \in {}^*U$, and so $x \in \operatorname{fin}({}^*X)$. $\star$

Exercise 77: In view of Proposition 13.2 and Theorem 14.3, a topological vector space X is Hausdorff if and only if the relation $x-y \in O$ for any open neighborhood O of 0 implies that $x = y$.

The relation $[x] - [y] \in O$ for any open neighborhood $O \subseteq X/U$ of $[0]$ is equivalent to the fact that for any neighborhood $O_0 \subseteq X$ of 0 there is some $u \in U$ with $x - y + u \in O_0$. Putting $z := y - x$, this means that we find for any neighborhood $O_0 \subseteq X$ of 0 some u with $u \in z + O$, i.e. any neighborhood of z contains some element of U, i.e. $z \in \overline{U}$. We thus have proved that $[x] - [y] \in O$ for any open neighborhood $O \subseteq X/U$ if and only if $y - x \in \overline{U}$.

Together with the assumption in the beginning, we obtain: X/U is Hausdorff if and only if for any $[x], [y] \in X/U$ the relation $y - x \in \overline{U}$ implies $[x] = [y]$. Since the latter means $y - x \in U$, we have that X/U is Hausdorff if and only if the relation $y - x \in \overline{U}$ for some $y, x \in X$ implies $y - x \in U$. This is the case if and only if $\overline{U} \subseteq U$, i.e. if and only if U is closed.

For the second claim, note that $X/\{0\}$ can in a canonical way be identified with X such that also the open sets are in correspondence. ⋆

Exercise 78: Let $\mathscr{F}$ denote the system of all functions from subsets of $\mathcal{P}(S_0)$ into $[0, \infty]$ (recall Exercise 83). Consider the sentence

$$\forall \underline{x} \in \mathcal{P}(S_0), \underline{y} \in \mathscr{F}, \underline{z} \in \mathcal{P}(S_0)^{<\mathbb{N}} : \alpha \implies \beta$$

where α and β are shortcuts with the meaning

$$\text{“}\underline{x} \text{ is a set algebra”} \wedge \text{“}\underline{y} : \underline{x} \to [0, \infty] \text{ is a measure”} \wedge \operatorname{rng}(\underline{z}) \subseteq \underline{x}$$

and

$$\bigcup \operatorname{rng} \underline{z} \in \Sigma \wedge (\text{“Range of } \underline{z} \text{ pairwise disjoint”} \implies \underline{y}(\bigcup \operatorname{rng} \underline{z}) = \sum_{\underline{n}} \underline{z}(\underline{n})$$

respectively. The transfer of this sentence implies the statement for the choice $\underline{x} := \Sigma$, $\underline{y} := \mu$, $\underline{z}(n) := A_n$ $(n = 1, \ldots, h)$. ⋆

Exercise 79: Let $c > 0$ be infinitesimal with $\delta(x) = 0$ for $|x| > c$. In view of the transfer principle, we may estimate analogously as for standard integrals

$$\left| \int \delta(x)^* f(x)\, dx - {}^*f(0) \right| = \left| \int \delta(x) \left({}^*f(x) - {}^*f(0) \right) dx \right| \leq \int \delta(x) M \, dx = M$$

where $M := \sup\{|{}^*f(x) - {}^*f(0)| : x \in {}^*\mathbb{R},\ \delta(x) \neq 0\}$ (note that the supremum exists, since the considered set is internal by the internal definition principle). We have

$$M \leq \sup\{|{}^*f(x) - {}^*f(0)| : x \in {}^*\mathbb{R} \wedge |x| \leq c\} =: M_0$$

with some infinitesimal $c > 0$. Since f is continuous at 0, we have $M_0 \approx 0$, and so also $M \approx 0$ which implies the statement. ⋆

Exercise 80: Putting $U := \bigcup_{i \in I} X_i$ and $X := \prod_{i \in I} X_i$, we have

$$X = \{\underline{y} \in U^I \mid \forall \underline{x} \in I : \underline{y}(\underline{x}) \in f(\underline{x})\},$$

where f denotes the mapping $i \mapsto X_i$. The standard definition principle implies

$${}^*X = \{\underline{y} \in {}^*(U^I) \mid \forall \underline{x} \in {}^*I : \underline{y}(\underline{x}) \in {}^*f(\underline{x})\}.$$

By Theorem 3.21, we find

$${}^*X = \{\underline{y} : {}^*I \to {}^*U \mid \forall \underline{x} \in {}^*I : \underline{y} \in {}^*f(\underline{x}) \wedge \underline{y} \text{ internal}\}.$$

By Corollary A.6, we have ${}^*U = \bigcup_{i \in {}^*I} {}^*X_i$. Since $f(i) =: {}^*X_i$ $(i \in {}^*I)$, we have

$${}^*X = \{y : {}^*I \to {}^*\bigcup_{i \in {}^*I} {}^*X_i \mid y(i) \in {}^*X_i \text{ for all } i \in {}^*I \text{ and } y \text{ is internal}\}.$$

This implies the statement. ⋆

Exercise 81: By Proposition 3.16, X is a subset of some standard entity, say $X \subseteq {}^*B$ where $B \in \widehat{S}$. By Theorem 3.21, the system ${}^*\mathcal{P}(B)$ consists precisely of all internal subsets of ${}^*B$. Hence, ${}^*\mathcal{P}(B)$ contains all internal subsets of X, and conversely, all elements of ${}^*\mathcal{P}(B)$ which are subsets of X are internal. Hence, the system of all internal subsets of X can be written in the form $\{\underline{x} \in {}^*\mathcal{P}(B) : \underline{x} \subseteq X\}$ which is internal by Corollary 3.18. ⋆

Exercise 82: Observe that $A \times B$ is internal. Hence, by Exercise 81, the system P of all internal subsets of $A \times B$ is internal. Hence,

$$F = \{\underline{x} \in P \mid \underline{x} : A \to B\}$$

is internal by the internal definition principle (note that $\underline{x} : A \to B$ is an internal predicate in view of Proposition 3.6). Moreover, by Exercise 81, the system P_A of all internal subsets of A is internal. Note that if $f : A_0 \to B$ is internal, then $A_0 = \operatorname{dom}(f)$ is internal by Theorem 3.19. Hence, as above,

$$\mathscr{F} = \{\underline{x} \in P \mid \exists \underline{y} \in P_A : (\underline{x} : \underline{y} \to B)\}$$

is internal by the internal definition principle. ⋆

Exercise 83: The set $C := \mathscr{F}(A, B)$ belongs to $\widehat{S}$. Hence, we find some index n with $C \subseteq S_n$ (since S_n is transitive), and so ${}^*C \subseteq {}^*S_n$ (Lemma 3.5). The sentence $\forall \underline{x} \in S_n : (\underline{x} \in C \iff \alpha(\underline{x}))$ is true, where $\alpha(\underline{x})$ is the predicate with the meaning that $\underline{x} \subseteq A \times B$, and that $\underline{x}$ is a function (to formalize the latter quantify over A and B). The transfer principle implies that ${}^*C$ consists of all elements $\underline{x} \in {}^*S_n$ for which ${}^*\alpha(\underline{x})$ is true. The condition $\underline{x} \in {}^*S_n$ may be dropped, since we already know ${}^*C \subseteq {}^*S_n$, and ${}^*\alpha(\underline{x})$ becomes: "$\underline{x} \subseteq {}^*(A \times B) = {}^*A \times {}^*B$ is true, and $\underline{x}$ is a function". ⋆

Exercise 84: Put $\mathscr{C} := \{A \times B : A \in \mathscr{A}, B \in \mathscr{B}\}$, $U := \bigcup \mathscr{A}$ and $V := \bigcup \mathscr{B}$. Then $A \times B \subseteq U \times V$ for each $A \in \mathscr{A}$ and each $B \in \mathscr{B}$. Putting $P := \mathcal{P}(U \times V)$, we thus have

$$\mathscr{C} = \{\underline{z} \in P \mid \exists \underline{x} \in \mathscr{A}, \underline{y} \in \mathscr{B} : \underline{z} = \underline{x} \times \underline{y}\}.$$

The standard definition principle implies

$${}^*\mathscr{C} = \{\underline{z} \in {}^*P \mid \exists \underline{x} \in {}^*\mathscr{A}, \underline{y} \in {}^*\mathscr{B} : \underline{z} = \underline{x} \times \underline{y}\}.$$

Now note that ${}^*P = {}^*\mathcal{P}(U \times V)$ consists by Theorem 3.21 of all internal subsets of ${}^*(U \times V) = {}^*U \times {}^*V = (\bigcup {}^*\mathscr{A}) \times (\bigcup {}^*\mathscr{B})$ (for the equalities we used that $*$ is a superstructure monomorphism and Theorem A.4). Since by Theorem 3.19 each of the sets $A \times B$ is internal for $A \in {}^*\mathscr{A}$ and $B \in {}^*\mathscr{B}$, the statement follows. ⋆

Bibliography

[AFHKL86] Albeverio, S., Fenstad, J. E., Høegh-Krohn, R., and Lindstrøm, T., *Nonstandard methods in stochastic analysis and mathematical physics*, Academic Press, Orlando, San Diego, New York, 1986.

[Ban87] Banach, S., *Introduction to linear operations*, North-Holland, Amsterdam, New York, Oxford, 1987, translation of *Théorie des Opérations Linéaires*, Polish Scientific Publ., Warsaw, 1979.

[BH02] Borde, P. and Haddad, L., *Etude de la mesurabilité de la function* $^o\cos(2\pi\omega x)$, Ital. J. Pure Appl. Math. **11** (2002), 19–31.

[CK90] Chang, C. C. and Keisler, H. J., *Model theory*, 3rd ed., North-Holland Publ. Company, Amsterdam, London, New York, Tokyo, 1990.

[Cut88] Cutland, N. (ed.), *Nonstandard analysis and its applications*, Cambridge Univ. Press, Cambridge, 1988.

[Dav77] Davis, M., *Applied nonstandard analysis*, John Wiley & Sons, New York, London, Sidney, 1977.

[DS88] Diener, F. and Stroyan, K. D., *Syntactical methods in infinitesimal analysis*, In Cutland [Cut88], 258–281.

[FL92] Fuchssteiner, B. and Luxemburg, W. A. J. (eds.), *Analysis and geometry. Trends in research and teaching*, Mannheim, Leipzig, Wien, Zürich, BI Wissenschaftsverlag, 1992, dedicated to D. Laugwitz.

[FW91] Foreman, M. and Wehrung, F., *The Hahn-Banach theorem implies the existence of a non-Lebesgue measurable set*, Fund. Math. **138** (1991), no. 1, 13–19.

[Gol98] Goldblatt, R., *Lectures on the hyperreals*, Springer, New York, Berlin, Heidelberg, 1998.

[Hen72a] Henson, C. W., *The nonstandard hulls of a uniform space*, Pacific J. Math. **43** (1972), 115–137.

[Hen72b] Henson, C. W., *On the nonstandard representation of measures*, Trans. Amer. Math. Soc. **172** (1972), 437–446.

[HL85] Hurd, A. and Loeb, P., *An introduction to nonstandard real analysis*, Academic Press, Orlando, San Diego, New York, 1985.

[HM72] Henson, C. W. and Moore, L. C., Jr., *The nonstandard theory of topological vector spaces*, Trans. Amer. Math. Soc. **172** (1972), 405–435.

[HM74] Henson, C. W. and Moore, L. C., Jr., *Subspaces of the nonstandard hull of a normed space*, Trans. Amer. Math. Soc. **197** (1974), 131–143.

[HM83] Henson, C. W. and Moore, L. C., Jr., *Nonstandard analysis and the theory of Banach spaces*, Nonstandard Analysis and the Theory of Banach Spaces (Heidelberg) (Hurd, A., ed.), Lect. Notes Math., no. 983, Springer, Heidelberg, 1983, 27–112.

[HS69] Hewitt, E. and Stromberg, K., *Real and abstract analysis*, 2nd ed., Springer, Berlin, Heidelberg, New York, 1969.

[Jec97] Jech, T. J., *Set theory*, 2nd ed., Springer, Berlin, Heidelberg, New York, 1997.

[Kel50] Kelley, J. L., *The Tychonoff product theorem implies the axiom of choice*, Fund. Math. **37** (1950), 75–76.

[KM92] Keller, G. and Moore, L. C., Jr., *Invariant means on the group of integers*, In Fuchssteiner and Luxemburg [FL92], dedicated to D. Laugwitz, 1–18.

[Lau86] Laugwitz, D., *Zahlen und Kontinuum*, Wissensch. Buchgesellsch. Darmstadt, Zürich, 1986.

[LG81] Lutz, R. and Goze, M., *Nonstandard analysis. A practical guide with applications*, Lect. Notes Math., no. 881, Springer, Berlin, New York, 1981.

[Loe75] Loeb, P. A., *Conversion from nonstandard to standard measure spaces and applications in probability theory*, Trans. Amer. Math. Soc. **211** (1975), 113–122.

[Lor48] Lorentz, G. G., *A contribution to the theory of divergent sequences*, Acta Math. **80** (1948), 167–190.

[LR94] Landers, D. and Rogge, L., *Nichtstandard Analysis*, Springer, Berlin, Heidelberg, New York, 1994.

[LRN51] Loś, J. and Ryll-Nardzewski, C., *On the applications of Tychonoff's theorem in mathematical proofs*, Fund. Math. **38** (1951), 233–237.

[LRN54] Loś, J. and Ryll-Nardzewski, C., *Effectiveness of the representation theory for Boolean algebras*, Fund. Math. **41** (1954), 49–56.

[Lux62] Luxemburg, W. A. J., *Two applications of the method of construction of ultrapowers in analysis*, Bull. Amer. Math. Soc. (N.S.) **68** (1962), 416–419.

[Lux63] Luxemburg, W. A. J., *Addendum to "On the measurability of a function which occurs in a paper by A. C. Zaanen"*, Proc. Netherl. Acad. Sci. (A) **66** (1963), 587–590, (Indag. Math. **25** (1963), 587–590).

[Lux69a] Luxemburg, W. A. J., *A general theory of monads*, In *(Intern. Symp. on the) Applications of Model Theory to Algebra, Analysis, and Probability (1967)* [Lux69b], 18–86.

[Lux69b] Luxemburg, W. A. J. (ed.), *(Intern. symp. on the) Applications of model theory to algebra, analysis, and probability (1967)*, Toronto, Holt, Rinehart and Winston, 1969.

[Lux69c] Luxemburg, W. A. J., *Reduced powers of the real number system and equivalents of the Hahn-Banach extension theorem*, In *(Intern. Symp. on the) Applications of Model Theory to Algebra, Analysis, and Probability (1967)* [Lux69b], 123–137.

[Lux73] Luxemburg, W. A. J., *What is nonstandard analysis?*, Amer. Math. Monthly **80** (1973), no. 6 part II, 38–67.

[Lux92] Luxemburg, W. A. J., *Nonstandard hulls, generalized limits and almost convergence*, In Fuchssteiner and Luxemburg [FL92], dedicated to D. Laugwitz, 19–45.

[LZ71] Luxemburg, W. A. J. and Zaanen, A. C., *Riesz spaces*, vol. I, North-Holland Publ. Company, Amsterdam, London, 1971.

[Nel77] Nelson, E., *Internal set theory, a new approach to nonstandard analysis*, Bull. Amer. Math. Soc. **83** (1977), 1165–1198.

[Paw91] Pawlikowski, J., *The Hahn-Banach theorem implies the Banach-Tarski paradox*, Fund. Math. **138** (1991), no. 1, 20–22.

[Pin73] Pincus, D., *The strength of the Hahn-Banach theorem*, Victoria Symposium on Nonstandard Analysis (Univ. of Victoria 1972) (Dold, A. and Eckmann, B., eds.), Lect. Notes Math., no. 369, Springer, 1973, 203–248.

[PS77] Pincus, D. and Solovay, R. M., *Definability of measures and ultrafilters*, J. Symbolic Logic **42** (1977), no. 2, 179–190.

[Ric82] Richter, M. M., *Ideale Punkte, Monaden und Nichtstandard-Methoden*, Vieweg, Braunschweig, Wiesbaden, 1982.

[RK04] Reeken, M. and Kanovei, V., *Nonstandard analysis: Axiomatically*, Springer, Berlin, Heidelberg, New York, 2004.

[Rob64] Robinson, A., *On generalized limits and linear functionals*, Pacific J. Math. **14** (1964), 269–283.

[Rob70] Robinson, A., *Non-standard analysis*, 2nd ed., North-Holland Publ. Company, Amsterdam, London, 1970.

[Rob88] Robert, A., *Nonstandard analysis*, John Wiley & Sons, New York, Chichester, Brisbane, 1988.

[Rud87] Rudin, W., *Real and complex analysis*, 3rd ed., McGraw-Hill, Singapore, 1987.

[Rud90] Rudin, W., *Functional analysis*, 14th ed., McGraw-Hill, New Delhi, New York, 1990.

[RZ69] Robinson, A. and Zakon, E., *A set-theoretical characterization of enlargements*, In Luxemburg [Lux69b], 109–122.

[SB86] Stroyan, K. D. and Bayod, J. M., *Foundations of infinitesimal stochastic analysis*, North-Holland, Amsterdam, 1986.

[Sie38] Sierpiński, W., *Fonctions additives non complètement additives et fonctions non mesurables*, Fund. Math. **30** (1938), 96–99.

[Sik64] Sikorski, R., *Boolean algebras*, Springer, Berlin, Heidelberg, New York, 1964.

[SL76] Stroyan, K. D. and Luxemburg, W. A. J., *Introduction to the theory of infinitesimals*, Academic Press, New York, San Francisco, London, 1976.

[Sol70] Solovay, R. M., *A model of set-theory in which every set of reals is Lebesgue measurable*, Ann. of Math. (2) **92** (1970), 1–56.

[Tay69] Taylor, R. F., *On some properties of bounded internal functions*, In Luxemburg [Lux69b], 167–170.

[vdB87] van den Bergh, I., *Nonstandard asymptotic analysis*, Lect. Notes Math., no. 1249, Springer, Berlin, New York, 1987.

[vQ79] von Querenburg, B., *Mengentheoretische Topologie*, 2nd ed., Springer, Berlin, Heidelberg, New York, 1979.

[Wag86] Wagon, S., *The Banach-Tarski paradox*, 2nd ed., Cambridge Univ. Press, Cambridge, 1986.

[Zaa67] Zaanen, A. C., *Integration*, North-Holland Publ. Company, Amsterdam, 1967.

[Zak69] Zakon, E., *Remarks on the nonstandard real axis*, In Luxemburg [Lux69b], 195–227.

Index