Non-standard Analysis

David Pierce

August 1, 2014
Miman Sinan Güzel Sanatlar Üniversitesi
http://mat.msgsu.edu.tr/~dpierce/

Preface

Here are notes of my course, called Rudiments of Nonstandard Analysis, given at the Nesin Mathematical Village, in Şirince, July 14–27, 2014. I started lecturing in Turkish, switching to English when it transpired that everybody understood this. After every lecture, I wrote out a record what had happened. The present document is a highly edited version of the daily records. I have added some references, and cross-references, and I have made some corrections and amplifications, though without changing the content of particular days.

The experience of the students ranged from one year of university to some years of graduate school. Prerequisites of the course had been given as

> Calculus and algebra (the theorem that maximal ideals are prime),

though in the event, most of the students did not know much algebra. The "abstract" of the course was

The axiom of choice, ultrafilters, ultraproducts. Connections with algebra. The rudiments of nonstandard analysis.

What actually happened can be seen in the main text. A few EXERCISES are made explicit in the text; also, formally stated theorems without proofs can be considered as exercises. In the actual course, some of these were proved by students at the board.

Contents

List of Figures

Summaries of the days

1\.

1. Calculus with infinitesimals (the limit of the sum is the sum of the limits). Non-Archimedean ordered fields.

2. More calculus with infinitesimals (the limit of the product is the product of the limits). Rings and their ideals. Power sets as Boolean rings.

3. Why $\mathbb{R}$ had to be rigorously defined. Equivalence of standard and non-standard definitions of limits. Dedekind's construction of $\mathbb{R}$. The Cauchy-sequence construction of $\mathbb{R}$. Valuation rings.

4. Ideals of power sets. The Maximal Ideal Theorem and its proof by Zorn's Lemma. Maximal ideals $\mathfrak{m}$ of $\mathscr{P}(\omega)$ that contain all finite subsets of ω. The Sorites Paradox. The ordering of the ultrapower $\mathbb{R}^\omega/M$, where M is the maximal ideal $\{x \in \mathbb{R}^\omega \colon \operatorname{supp}(x) \in \mathfrak{m}\}$.

5. Ultraproducts $\prod_{i\in\omega} K_k \big/ M$ of fields, as for example finite fields. Logical formulas. The Prime Ideal Theorem. Łoś's Theorem for fields, and the Transfer Principle.

6. Łoś's Theorem (and the Compactness Theorem) in an arbitrary signature. ${}^*\mathbb{R}$ as the [ultrapower] $\mathbb{R}^\omega/M$. ${}^*\mathbb{N}$. Non-standard analysis of sequences. Filters and ultrafilters.

2.
1. Ultrafilters on ω. More non-standard analysis of sequences. Standard parts of finite non-standard real numbers. Why the Transfer Principle does not apply to second-order properties.
2. Non-standard analysis of bounded sets and limit points. The Bolzano–Weierstrass Theorem.
3. Closed sets and open sets. Monads. Logical and topological compactness. The Heine–Borel Theorem.
4. Topological spaces. The compactness of the spectrum of a ring. Non-standard characterization of topological compactness. The logical Compactness Theorem implies the Prime Ideal Theorem. Boolean algebras and the Stone Representation Theorem.
5. Logical truth and Lindenbaum algebras. Stone spaces and the proof of the Stone Representation Theorem. Bases of topological spaces. Proof of the logical Compactness Theorem from the Prime Ideal Theorem.
6. The Axiom of Choice is equivalent to Łoś's Theorem and the Prime Ideal Theorem together.

1 First week

1.1 Monday

Suppose a function f is defined by

$$f(x) = \begin{cases} x\sin(1/x), & \text{if } x \neq 0, \\ 0, & \text{if } x = 0. \end{cases}$$

Is f continuous at 0? Why or why not?

By the standard definition, for an arbitrary function f from $\mathbb{R}$ to $\mathbb{R}$, the expression

$$\lim_{x \to a} f(x) = L$$

means*

$$\forall \varepsilon \Bigl(\varepsilon > 0 \implies \exists \delta \bigl(\delta > 0 \;\&\; \forall x \\ (0 < |x - a| < \delta \implies |f(x) - L| < \varepsilon)\bigr)\Bigr). \quad (1)$$

Then f is continuous at a if and only $\lim_{x \to a} f(x) = f(a)$.

Theorem 1. *If* $\lim_{x \to a} f(x) = L$ *and* $\lim_{x \to a} g(x) = M$, *then*

$$\lim_{x \to a} (f + g)(x) = L + M$$

("the limit of the sum is the sum of the limits").

Standard proof. By the triangle inequality,

$$|(f + g)(x) - (L + M)| \leqslant |f(x) - L| + |g(x) - M|.$$

Suppose $\varepsilon > 0$. For some positive δ_1 and δ_2,

$$0 < |x - a| < \delta_1 \implies |f(x) - L| < \frac{\varepsilon}{2},$$
$$0 < |x - a| < \delta_2 \implies |g(x) - M| < \frac{\varepsilon}{2}.$$

Let $\delta = \min(\delta_1, \delta_2)$. Then

$$0 < |x - a| < \delta \implies |f(x) - L| + |g(x) - M| < \varepsilon,$$

*Commonly the defining sentence is written as $(\forall \varepsilon > 0)(\exists \delta > 0)\forall x(0 < |x - a| < \delta \Rightarrow |f(x) - L| < \varepsilon)$.

and therefore

$$0 < |x - a| < \delta \implies |(f + g)(x) - (L + M)| < \varepsilon.$$

Thus $\lim_{x\to a}(f + g)(x) = L + M$. □

By the standard, "ε-δ" definition, $\lim_{x\to a} f(x) = L$ means that $f(x)$ is **relatively** close to L, or as close as we like to L, provided x is **sufficiently** close (but not equal) to a. Here the variable x ranges over the *ordered field* $\mathbb{R}$ of real numbers.

We are going to develop a notion of being "absolutely" close, denoted by $\simeq$. Then $\lim_{x\to a} f(x) = L$ will mean that $f(x)$ is absolutely close to L, provided x is absolutely close (but not equal) to a: in symbols,

$$\forall x \ (x \simeq a \ \ \& \ \ x \neq a \implies f(x) \simeq L). \tag{2}$$

However, the variable x here will range over an ordered field larger than $\mathbb{R}$.

If $x \simeq a$, then $x - a \simeq 0$, and $x - a$ will be called **infinitesimal** (or **infinitely small**). The sum of two infinitesimals will be infinitesimal. Then for the theorem above, we shall have:

Non-standard proof. If $x \simeq a$, but $x \neq a$, then $f(x) \simeq L$ and $g(x) \simeq M$, so $f(x) - L$ and $g(x) - M$ are infinitesimal, and therefore so is their sum, which is equal to $(f + g)(x) - (L + M)$; thus

$$(f + g)(x) \simeq (L + M).$$ □

The simplest example of an ordered field that includes $\mathbb{R}$ and contains infinitesimals is the field $\mathbb{R}(X)$ of rational functions in the variable X over $\mathbb{R}$; an arbitrary nonzero element of this field can be written as

$$\frac{a_n X^n + a_{n-1} X^{n-1} + \cdots + a_0 X^0}{b_m X^m + b_{m-1} X^{m-1} + \cdots + b_0 X^0}$$

for some nonnegative integers n and m, for some a_i and b_j in $\mathbb{R}$, where $a_n b_m \neq 0$. We define the element of $\mathbb{R}(X)$ to be **positive** if $a_n b_m > 0$. Hence for all positive integers n,

$$X - n = \frac{1 \cdot X^1 - nX^0}{1 \cdot X^0} > 0,$$

so $X > n$, and therefore $0 < 1/X < 1/n$. Thus X is **infinite,** while $1/X$ is **infinitesimal.**

To be precise, there are two equivalent ways to define an ordering of a field K. If $P \subset K$ and is closed under addition and multiplication, while

$$P \sqcup \{0\} \sqcup \{-x \colon x \in P\} = K,$$

then P is the set of **positive** elements of K with respect to an ordering of K, given by

$$x < y \iff y - x \in P.$$

Alternatively, $<$ is linear ordering of K such that

$$x < y \implies x + z < y + z,$$
$$x < y \ \& \ z > 0 \implies xz < yz.$$

Because it has infinite elements, the ordered field $\mathbb{R}(X)$ is **non-Archimedean.** It is not rich enough for doing analysis. We are going to do analysis in a non-Archimedean ordered field denoted by

$$^*\mathbb{R},$$

which will be the *quotient* $\mathbb{R}^\omega/M$, where

- $\omega = \{0, 1, 2, \dots\}$, the set of nonnegative integers, and
- M is a *nonprincipal maximal ideal* of the *ring* $\mathbb{R}^\omega$ of functions from ω to $\mathbb{R}$.

1.2 Tuesday

Theorem 2. *If* $\lim_{x\to a} f(x) = L$ *and* $\lim_{x\to a} g(x) = M$*, then*

$$\lim_{x\to a} (fg)(x) = LM$$

("the limit of the product is the product of the limits").

Standard proof. Since

$$\begin{aligned}|(fg)(x) - LM| &= |f(x)\cdot g(x) - LM|\\ &= |f(x)\cdot g(x) - f(x)\cdot M + f(x)\cdot M - LM|\\ &\leqslant |f(x)|\cdot|g(x)-M| + |f(x)-L|\cdot|M|,\end{aligned}$$

it is enough to let δ_1 be such that

$$\begin{aligned}0 < |x-a| < \delta_1 &\implies |f(x) - L| < \frac{\varepsilon}{2|M|}\\ &\implies |f(x)| < |L| + \frac{\varepsilon}{2|M|},\end{aligned}$$

then let δ_2 be such that

$$0 < |x-a| < \delta_2 \implies |g(x) - M| < \frac{\varepsilon}{2\left(|L| + \frac{\varepsilon}{2|M|}\right)},$$

then let $\delta = \max(\delta_1, \delta_2)$. □

For the non-standard proof, taking place in the non-Archimedean ordered field $^*\mathbb{R}$ mentioned above, we have noted that the sum of infinitesimals is infinitesimal; we also need that the product of a **finite** (that is, non-infinite) number and an infinitesimal is infinitesimal.

Nonstandard proof. If $x \simeq a$, then $f(x) \simeq L$ and $g(x) \simeq M$, so $f(x) - L \simeq 0$ and $g(x) - M \simeq 0$, and therefore

$$\begin{aligned}(fg)(x) - LM &= f(x) \cdot g(x) - LM \\ &= f(x) \cdot g(x) - f(x) \cdot M + f(x) \cdot M - LM \\ &= f(x) \cdot \bigl(g(x) - M\bigr) + \bigl(f(x) - L\bigr) \cdot M \\ &\simeq 0.\end{aligned}$$ □

What we now call simply calculus was once called *infinitesimal* calculus because it was conceived in terms of infinitesimals. When calculus was made rigorous, infinitesimals were not involved, but today's ε-δ definitions were developed.

Why make calculus rigorous? Well, why do we believe that there is a number called $\sqrt{2}$? For centuries or even millenia, we have been able to compute approximations to this number. An algorithm for computing *decimal* approximations used to be learned in school (by my father, for example). However, Dedekind asserts (and I agree) that, before he gave a rigorous definition of the field $\mathbb{R}$ of real numbers, the equation

$$\sqrt{2} \cdot \sqrt{3} = \sqrt{6}$$

was not a theorem, because there can be no algorithm for multiplying infinite decimal expansions.

Let us review precisely the axioms for fields. First, an **abelian group** is a structure $(G, 0, -, +)$, where

- addition $(+)$ is a commutative, associative binary operation on the set G,
- 0 is an identity with respect to addition, and
- $-$ is inversion with respect to addition and 0.

A **field** is a structure $(K, 0, 1, -, +, \times)$, where

- $(K, 0, -, +)$ and $(K \smallsetminus \{0\}, 1, {}^{-1}, \times)$ are abelian groups (for some singulary operation ${}^{-1}$ on $K \smallsetminus \{0\}$), and
- multiplication $(\times)$ distributes over addition.

If we do not require the existence of a multiplicative inverse ${}^{-1}$ on $K \smallsetminus \{0\}$ (but we still require that $\times$ be commutative and associative, with identity 1), then what we have is a **ring.*** In a ring, by distributivity,

$$x \cdot 0 = x \cdot (0 + 0) = x \cdot 0 + x \cdot 0.$$

Fields, like $\mathbb{Q}$ and $\mathbb{R}$, are rings, but $\mathbb{Z}$ is a ring that is not a field. Let

$$\mathbb{N} = \{1, 2, 3, \dots\};$$

this is not even a ring. But if $n \in \mathbb{N}$, then there is a ring denoted by

$$\mathbb{Z}/n\mathbb{Z},$$

consisting of the elements $x + n\mathbb{Z}$, where

$$n\mathbb{Z} = \{ny \colon y \in \mathbb{Z}\},$$

so $x + n\mathbb{Z} = \{x + ny \colon y \in \mathbb{Z}\}$. The ring $\mathbb{Z}/n\mathbb{Z}$ is a field if and only if n is a prime p; and then this field is denoted by

$$\mathbb{F}_p.$$

Can this be made into an *ordered* field? No, because in an ordered field, always $1 + \dots + 1 > 0$. Such a field has **characteristic** 0; but $\operatorname{char}(\mathbb{F}_p) = p$, since in $\mathbb{F}_p$,

$$\underbrace{1 + \dots + 1}_{p} = 0.$$

*In some books, what we have defined as a ring would be called more precisely a nontrivial commutative ring with (multiplicative) identity; but we shall consider no other rings than these.

In general then, if K is an ordered field, we may assume $\mathbb{Q} \subseteq K$. Let

$$R = \{\text{finite elements of } K\},$$
$$I = \{\text{infinitesimal elements of } K\}.$$

Then R is a ring, and I is an **ideal** of R, because it is an additive subgroup of R that is closed under multiplication by elements of R.

If R is an arbitrary ring, and $I \subseteq R$, then I is an ideal of R if and only if

$$\left.\begin{array}{c} x, y \in I \implies x + y \in I, \\ r \in R \;\&\; x \in I \implies rx \in I, \\ 0 \in I. \end{array}\right\} \tag{3}$$

A ring is the **improper ideal** of itself; every other ideal of the ring is a **proper** ideal. A **maximal ideal** of R is a proper ideal I that is maximal as such, that is, for all ideals J of R,

$$I \subset J \implies J = R.$$

Then $n\mathbb{Z}$ is an ideal of $\mathbb{Z}$; it is a proper ideal, if $n > 1$; it is a maximal ideal, if n is prime.

An ideal is proper if and only if it does not contain 1. The ideal I of infinitesimal elements of the ring R of finite elements of a non-Archimedean field is a maximal ideal, because if $x \in R \smallsetminus I$, then $x^{-1} \in R$, so any ideal of R containing x contains also 1.

The **quotient ring** R/I is defined like $\mathbb{Z}/n\mathbb{Z}$; and when I is a maximal ideal of R, then R/I is a field.

The ordered field $\mathbb{R}$ is **complete:** every nonempty subset with an upper bound has a least upper bound, that is, a **supremum.** If again R is the ring of finite elements of a non-Archimedean ordered field, and I is the ideal of infinitesimal elements, then R/I is an Archimedean ordered field, and therefore it embeds in $\mathbb{R}$. Indeed, since we may assume $\mathbb{Q} \subseteq R$, there is a map

$$x + I \mapsto \sup\{u \in \mathbb{Q} \colon u < x\} \tag{4}$$

from R/I to $\mathbb{R}$, and this is an embedding.

An arbitrary element a of the power $\mathbb{R}^\omega$ can be written out as one of

$$(a_0, a_1, a_2, \dots), \qquad (a_k \colon k \in \omega),$$

where the a_k are in $\mathbb{R}$. Then the ring operations on $\mathbb{R}^\omega$ are given by

$$a + b = (a_k + b_k \colon k \in \omega),$$
$$ab = (a_k b_k \colon k \in \omega).$$

What are the ideals of $\mathbb{R}^\omega$? On $\mathbb{R}$, let us define

$$x^* = \begin{cases} x^{-1}, & \text{if } x \neq 0, \\ 0, & \text{if } x = 0; \end{cases}$$

and then on $\mathbb{R}^\omega$,

$$x^* = ({x_0}^*, {x_1}^*, \dots) = ({x_k}^* \colon k \in \omega).$$

There is a bijection $I \mapsto \bar{I}$ from $\{\text{ideals of } \mathbb{R}^\omega\}$ to $\{\text{ideals of } {\mathbb{F}_2}^\omega\}$, where

$$\bar{I} = \{x^* x \colon x \in I\}.$$

Given a in $\mathbb{R}^\omega$ or ${\mathbb{F}_2}^\omega$, we define

$$\operatorname{supp}(a) = \{i \in \omega \colon a_i \neq 0\},$$

the **support** of a. Then $x \mapsto \operatorname{supp}(x)$ is a bijection from ${\mathbb{F}_2}^\omega$ to $\mathscr{P}(\omega)$, and for x and y in ${\mathbb{F}_2}^\omega$ we have

$$\operatorname{supp}(xy) = \operatorname{supp}(x) \cap \operatorname{supp}(y),$$
$$\operatorname{supp}(x + y) = \operatorname{supp}(x) \bigtriangleup \operatorname{supp}(y),$$

where $\triangle$ denotes **symmetric difference:**

$$X \triangle Y = (X \smallsetminus Y) \cup (Y \smallsetminus X) = (X \cup Y) \smallsetminus (X \cap Y).$$

So $\mathscr{P}(\omega)$ is a ring, with $\triangle$ as addition and $\cap$ as multiplication. Then the maximal ideal M of $\mathbb{R}^\omega$ corresponds to a maximal ideal $\mathfrak{m}$ of $\mathscr{P}(\omega)$, namely $\{\operatorname{supp}(x) \colon x \in M\}$.

The ring $\mathscr{P}(\omega)$ is a **Boolean ring,** that is,

$$x^2 = x.$$

From this we have

$$2x = (2x)^2 = 4x^2 = 4x,$$
$$0 = 2x.$$

Thus the characteristic of a Boolean ring is 2. Hence also

$$x \cdot (1 + x) = 0,$$

so the only Boolean field is $\mathbb{F}_2$. Hence $\mathscr{P}(\omega)/\mathfrak{m} \cong \mathbb{F}_2$, and

$$X \in \mathfrak{m} \iff \omega \smallsetminus X \notin \mathfrak{m}.$$

1.3 Wednesday

Why was it not a theorem before Dedekind that $\sqrt{2} \cdot \sqrt{3} = \sqrt{6}$? We can approximate roots with decimals, but have no algorithm for multiplying (or adding) infinite decimal expansions. For example, what is

$$(2.333\dots) \times (2.\overline{142857})?$$

We have

$$2.3 \times 2.1 = 4.83,$$
$$2.33 \times 2.14 = 4.9862,$$
$$2.333 \times 2.142 = 4.997286,$$
$$2.3333 \times 2.1428 = 4.99979524,$$

but

$$2.3334 \times 2.1429 = 5.00024286.$$

So we cannot know the first digit of the product, just by computing with finitely many digits.

Dedekind gives us $\mathbb{R}$ as an ordered field. Using this, we shall construct an ordered field ${}^*\mathbb{R}$ (namely $\mathbb{R}^\omega/M$ as before) where $\mathbb{R} \subset {}^*\mathbb{R}$. If $f \colon \mathbb{R} \to \mathbb{R}$, then f will extend to a function ${}^*f$ from ${}^*\mathbb{R}$ to ${}^*\mathbb{R}$ (so ${}^*f \restriction \mathbb{R} = f$). The equivalence between the standard and non-standard definitions of limits (in (1) and (2), pages 5 and 6, respectively) will be established as follows.

Theorem 3. *In* $\mathbb{R}$, (1) *holds, namely*

$$\forall\varepsilon\Big(\varepsilon > 0 \implies \exists\delta\,\big(\delta > 0 \;\&\; \forall x \\ (0 < |x-a| < \delta \implies |f(x) - L| < \varepsilon)\big)\Big),$$

if and only if in ${}^*\mathbb{R}$,

$$\forall x\,(x \simeq a \;\&\; x \neq a \implies {}^*f(x) \simeq L).$$

Proof. ($\Rightarrow$). Suppose $b \simeq a$ and $b \neq a$. Then for all positive δ in $\mathbb{R}$, $0 < |b-a| < \delta$. Let $\varepsilon \in \mathbb{R}$ and $\varepsilon > 0$. By (1), for some positive δ in $\mathbb{R}$,

$$\mathbb{R} \vDash \forall x\,(0 < |x-a| < \delta \implies |f(x) - L| < \varepsilon)$$

(here $\vDash$ is the truth relation), and so (as we shall show)

$$^*\mathbb{R} \vDash \forall x\ (0 < |x - a| < \delta \implies |^*f(x) - L| < \varepsilon),$$

and therefore $|^*f(b) - L| < \varepsilon$. This being true for all positive ε in $\mathbb{R}$, we have $^*f(b) \simeq L$.

($\Leftarrow$). Suppose (1) fails, so that, for some positive ε in $\mathbb{R}$,

$$\mathbb{R} \vDash \forall\delta\ \big(\delta > 0 \implies \exists x\ (0 < |x - a| < \delta\ \ \&\ \ |f(x) - L| < \varepsilon)\big).$$

Then $^*\mathbb{R} \vDash$ [same thing, with $^*f$ for f]:

$$^*\mathbb{R} \vDash \forall\delta\ \big(\delta > 0 \implies \exists x\ (0 < |x - a| < \delta\ \ \&\ \ |^*f(x) - L| < \varepsilon)\big).$$

In particular, for all positive infinitesimal δ, there is b in $^*\mathbb{R}$ such that $0 < |b - a| < \delta$ (so $b \simeq a$, but $b \neq a$), and $|^*f(b) - L| \geqslant \varepsilon$, so $^*f(b) \not\simeq L$. □

We could do non-standard analysis (as in the non-standard proofs of Theorems 1 and 2), without proving Theorem 3; this theorem just assures us that we would not be doing anything new.

How do we get the property of $^*\mathbb{R}$ used in the last proof? First it will be useful to review how we obtain $\mathbb{R}$. Recall from page 9 that $(K, 0, 1, -, +, \times)$ is a **field** if

- $(K, 0, -, +)$ and $(K \smallsetminus \{0\}, 1, {}^{-1}, \times)$ (for some ${}^{-1}$) are abelian groups,
- $+$ distributes over $\times$.

If $(K \smallsetminus \{0\}, 1, \times)$ is only a **monoid** ($\times$ is commutative and associative, 1 is an identity), then $(K, 0, 1, -, +, \times)$ is a **ring.**

If K is a field, then $(K, <)$ is an **ordered field** if $\{x \in K : x > 0\}$ is closed under $+$ and $\times$, and

$$K = \{x \in K : x < 0\} \sqcup \{0\} \sqcup \{x \in K : x > 0\}.$$

Adapting Dedekind's definition, let us say that a **cut** of $\mathbb{Q}$ is a nonempty proper subset A of $\mathbb{Q}$ such that*

$$y \in A \ \& \ x < y \implies x \in A,$$
$$x \in A \implies \exists y \, (y \in A \ \& \ x < y).$$

Let

$$\mathbb{R} = \{\text{cuts of } \mathbb{Q}\}.$$

Then $x \mapsto \{y \in \mathbb{Q} \colon y < x\}$ is an embedding of $(\mathbb{Q}, <)$ in $(\mathbb{R}, \subset)$, and addition and multiplication on $\mathbb{Q}$ extend to *continuous* operations on $\mathbb{R}$, and then (by continuity) $\mathbb{R}$ becomes an ordered field.

Note that Dedekind's construction relies only on the *ordering* of $\mathbb{Q}$. Another construction of $\mathbb{R}$ may be more useful for us. For this, we let

$$R = \{\text{Cauchy sequences of } \mathbb{Q}\},$$
$$I = \{\text{sequences of } \mathbb{Q} \text{ converging to } 0\}.$$

Recall that $a \in R$ means

$$\forall \varepsilon \, \big(\varepsilon > 0 \implies \exists k \, \forall m \, \forall n$$
$$(m \in \mathbb{N} \ \& \ k \leqslant m \ \& \ n \in \mathbb{N} \ \& \ k \leqslant n \implies |a_m - a_n| < \varepsilon)\big).$$

Then $I \subset R$.

Theorem 4. *I is a maximal ideal of R.*

Proof. This is an EXERCISE; but assuming I is an ideal, let us show it is maximal. Say $a \in R \smallsetminus I$. We must show that every ideal J containing

*For Dedekind, the cut is $(A, \mathbb{Q} \smallsetminus A)$, and if A has a supremum in $\mathbb{Q}$, it does not matter whether this supremum belongs to A or not.

a and including I is R. Say $b \in R$ and $c \in I$. Then $ab + c \in J$. We want to show $1 \in J$. So we want to find b in $\mathbb{R}$ and c in I such that

$$ab + c = 1.$$

Given that $a = (a_0, a_1, \dots)$, we could let $b = ({a_0}^{-1}, {a_1}^{-1}, \dots)$, so $ab = (1, 1, 1, \dots) = 1$. But possibly $a_k = 0$. Recall from page 12 the definition on $\mathbb{R}$,

$$x^* = \begin{cases} x^{-1}, & \text{if } x \neq 0, \\ 0, & \text{if } x = 0. \end{cases}$$

So let

$$b = a^* = ({a_0}^*, {a_1}^*, \dots).$$

Then ab and $1 - ab$ are in $\in \{0, 1\}^\omega$. We want to show $1 - ab \in I$. That is, we want to show that, for some k, if $n \in \mathbb{N}$ and $k \leqslant n$, then $a_n \neq 0$.

The desired conclusion holds because a is Cauchy, but does not converge to 0. To be precise, for some positive ε, for all k, for some n in $\mathbb{N}$, we have $k \leqslant n$, but $|a_n| \geqslant \varepsilon$; but also, for some k, for all m and n in $\mathbb{N}$, if $k \leqslant m$ and $k \leqslant n$, then $|a_n - a_m| < \varepsilon/2$. We may assume $|a_n| \geqslant \varepsilon$, and so

$$\frac{\varepsilon}{2} \leqslant |a_n| - \frac{\varepsilon}{2} < |a_m|;$$

in particular, $a_m \neq 0$. □

Theorem 5. *[In the notation above,] the function* $x + I \mapsto \lim(x)$ *is a well-defined bijection from* R/I *to* $\mathbb{R}$*, and it preserves addition and multiplication: it is an isomorphism of fields.*

Proof. Exercise. □

Now let K be a non-Archimedean ordered field, so it has infinite elements. Let

$$\begin{aligned} R &= \{\text{finite elements of } K\}, \\ I &= \{\text{infinitesimal elements of } K\} \\ &= \{x^{-1} \colon x \in K \smallsetminus R\} \cup \{0\}. \end{aligned}$$

Then R is a **valuation ring** of K, because

$$\forall x\ (x \in K \smallsetminus R \implies x^{-1} \in R).$$

Let

$$R^{\times} = \{x \in R \smallsetminus \{0\} \colon x^{-1} \in R\},$$

the **group of units** of R. Then

$$I = R \smallsetminus R^{\times}.$$

We asserted on page 11 that I is an ideal; but this follows, simply because R is a valuation ring of K. Indeed, suppose $x, y \in I$. Either x/y or y/x is a well-defined element of R; assume $y/x \in R$. Then $1 + y/x \in R$, that is,

$$\frac{x+y}{x} \in R.$$

Since $x^{-1} \notin R$, also $1/(x+y) \notin R$, so $x + y \in I$. Similarly, if $x \in I$ and $y \in R$, then $xy \in I$ (EXERCISE).

So I is an ideal; automatically it is a maximal ideal, as we showed on page 11: if $a \in R \smallsetminus I$, then $a^{-1} \in R$, so every ideal containing a contains 1.

For another example, let $K = \mathbb{Q}(X)$ and

$$\begin{aligned} R &= \{f \in K \colon f(0) \text{ is defined}\}, \\ I &= \{f \in K \colon f(0) = 0\}. \end{aligned}$$

Then R is a valuation ring of K with maximal ideal I.

We want to let $K = \mathbb{R}^\omega/M$, where M is a maximal ideal of R^ω. We know that M is determined by a maximal ideal of the Boolean ring $(\mathscr{P}(\omega), 0, \omega, \triangle, \cap)$.

Theorem 6. *A subset I of $\mathscr{P}(\omega)$ is an ideal if and only if*

$$X, Y \in I \implies X \cup Y \in I,$$
$$Y \in I \ \& \ X \subseteq Y \implies X \in I,$$
$$\varnothing \in I.$$

Proof. EXERCISE: an adaptation of the definition (3) of ideals on page 11. □

For example, $\mathscr{P}(\omega)$ has (at least) two kinds of ideals:

- {finite subsets of ω}, and
- $\mathscr{P}(A)$, if $A \subseteq \omega$.

1.4 Friday

We want to understand the ideals of $\mathscr{P}(\omega)$, and more generally of $\mathscr{P}(\Omega)$ for arbitrary sets Ω. We have the characterization of ideals of $\mathscr{P}(\omega)$ in Theorem 6; it still holds when ω is replaced by Ω. If $A \subseteq \mathscr{P}(\Omega)$, then the ideal $\mathscr{P}(A)$ of $\mathscr{P}(\Omega)$ is called a **principal ideal.** If Ω is infinite, then the set of finite subsets of $\mathscr{P}(\Omega)$, which we can denote by

$$\mathscr{P}_\omega(\Omega),$$

is a **nonprincipal ideal,** because there is no subset A of Ω such that $\mathscr{P}_\omega(\Omega) = \mathscr{P}(A)$.

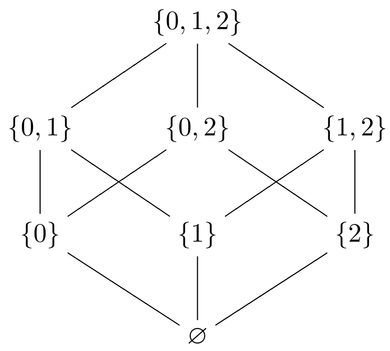

Figure 1. The elements of $\mathscr{P}(3)$

However, if Ω is finite, then all ideals of $\mathscr{P}(\Omega)$ are principal. Consider for example the case where Ω is 3, where $3 = \{0, 1, 2\}$, as in Figure 1: every ideal consists of an element and the elements below it in the diagram.

Theorem 7. *Suppose I is an ideal of $\mathscr{P}(\Omega)$ and $A \subseteq \Omega$. The smallest ideal of $\mathscr{P}(\Omega)$ that includes I and contains A is*

$$\{X \cup Y \colon X \in I \;\;\&\;\; Y \subseteq A\}. \tag{5}$$

Proof. If $X_0, X_1 \in I$ and $Y_0, Y_1 \subseteq A$, then

$$\begin{gathered}(X_0 \cup Y_0) \cup (X_1 \cup Y_1) = (X_0 \cup X_1) \cup (Y_0 \cup Y_1),\\ X_0 \cup X_1 \in I, \qquad\qquad Y_0 \cup Y_1 \subseteq A.\end{gathered}$$

Thus the set indicated in (5) is closed under $\cup$.

If $X \in I$ and $Y \subseteq A$ and $Z \subseteq X \cup Y$, then

$$\begin{gathered}Z = Z \cap (X \cup Y) = (Z \cap X) \cup (Z \cap Y),\\ Z \cap X \in I, \qquad\qquad Z \cap Y \subseteq A.\end{gathered}$$

Thus the indicated set is closed under taking subsets.

Obviously the indicated set is included in every ideal of $\mathscr{P}(\Omega)$ that includes I and contains A. $\square$

Denote the new ideal in the theorem by

$$I + (A).$$

If this is $\mathscr{P}(\Omega)$, then it contains Ω, so for some X in I and Y in $\mathscr{P}(A)$,

$$\Omega = X \cup Y,$$
$$X \supseteq \Omega \smallsetminus Y \supseteq \Omega \smallsetminus A,$$
$$\Omega \smallsetminus A \in I.$$

From this we obtain:

Theorem 8. *Let I be a proper ideal of $\mathscr{P}(\Omega)$. Then I is a maximal ideal if and only if, for all X in $\mathscr{P}(\Omega)$,*

$$X \notin I \implies \Omega \smallsetminus X \in I. \tag{6}$$

Proof. If I is maximal, then for all A in $\mathscr{P}(\Omega) \smallsetminus I$, we have $I + (A) = \mathscr{P}(\Omega)$, so $\Omega \smallsetminus A \in I$, as above.

Suppose conversely (6) holds. If $A \in \mathscr{P}(\Omega) \smallsetminus I$, then $\Omega \smallsetminus A \in I$, so $I + (A)$ contains both A and $\Omega \smallsetminus A$ and therefore Ω. $\square$

Principal maximal ideals of $\mathscr{P}(\omega)$ exist, namely $\mathscr{P}(\omega \smallsetminus \{k\})$. Do nonprincipal maximal ideals exist? They will, by the Maximal Ideal Theorem (Theorem 11) below.

A subset $\mathscr{C}$ of $\mathscr{P}(\Omega)$ is a **chain** if it is linearly ordered by $\subset$, that is, for all X and Y in $\mathscr{C}$,

$$X \subseteq Y \text{ OR } Y \subseteq X.$$

In the following, note well that

$$\bigcup \mathscr{C} = \bigcup_{Y \in \mathscr{C}} Y = \{x\colon \exists Y\ (Y \in \mathscr{C}\ \ \&\ \ x \in Y)\}.$$

Theorem 9 (Zorn's Lemma)**.** *Let $\mathscr{A} \subseteq \mathscr{P}(\Omega)$. Suppose that, for every subset $\mathscr{C}$ of $\mathscr{A}$ that is a chain, $\bigcup \mathscr{C} \in \mathscr{A}$. Then $\mathscr{A}$ has a maximal element X (so that, for all Y in A, if $X \subseteq Y$, then $X = Y$).*

Actually we shall treat Zorn's Lemma as an *axiom.* It is equivalent to the *Axiom of Choice* (see page 50).

Theorem 10. *The union of a nonempty chain $\mathscr{C}$ of proper ideals of a ring R is a proper ideal.*

Proof. If $x, y \in \bigcup \mathscr{C}$, then for some I and J in $\mathscr{C}$, $x \in I$ and $y \in J$. We may assume $I \subseteq J$. Then $x \in J$, so $x + y \in J$, and hence $x + y \in \bigcup \mathscr{C}$.

Similarly if $x \in \bigcup \mathscr{C}$ and $y \in R$, then $xy \in \bigcup \mathscr{C}$. So $\bigcup \mathscr{C}$ is an ideal.

If $\bigcup \mathscr{C}$ is the improper ideal, then $1 \in \bigcup \mathscr{C}$, so for some I in $\mathscr{C}$, $1 \in I$; but I is proper, so $1 \notin I$. □

Theorem 11 (Maximal Ideal Theorem)**.** *Every proper ideal I of a ring R is included in a maximal ideal of R.*

Proof. This maximal ideal is a maximal element of the set of proper ideals of R that include I; it exists, by Zorn's Lemma and the last theorem. □

In particular, there is a maximal ideal $\mathfrak{m}$ of $\mathscr{P}(\omega)$ that includes the ideal $\mathscr{P}_\omega(\omega)$. Let

$$M = \{x \in \mathbb{R}^\omega\colon\ \operatorname{supp}(x) \in \mathfrak{m}\},$$

where

$$\operatorname{supp}(x) = \{k \in \omega \colon x_k \neq 0\}.$$

As noted on page 12, M is an ideal, indeed a maximal ideal, of $\mathbb{R}^\omega$. Then $\mathbb{R}^\omega/M$ consists of the elements $a + M$, where $a \in \mathbb{R}^\omega$; and

$$\begin{aligned} a + M = b + M &\iff a - b \in M \\ &\iff \operatorname{supp}(a - b) \in \mathfrak{m} \\ &\iff \{k \in \omega \colon a_k \neq b_k\} \in \mathfrak{m}. \end{aligned}$$

Let us refer to elements of $\mathfrak{m}$ as **small,** and to elements of $\mathscr{P}(\omega) \smallsetminus \mathfrak{m}$ as **large.** Then:

- the union of two small sets is small;
- a subset of a small set is small;
- the complement of a small set is large.

Note that, by this definition, all finite sets are small, although some infinite sets will be small as well. The definition is a kind of resolution of the **Sorites Paradox,** or Paradox of the Heap, attributed to Eubulides of Miletus, 4th century BCE. If from a heap of sand, one grain is removed, still a heap of sand remains; but all of the grains in the heap can be removed one by one, so that, paradoxically, even one grain can constitute a heap.

A heap is a *large* pile of sand, and one grain is a *small* pile; but these are only relative terms, rather in the sense of page 6. Elements of $\mathfrak{m}$ are *absolutely* small; of $\mathscr{P}(\omega) \smallsetminus \mathfrak{m}$, absolutely large.

That $\mathbb{R}^\omega/M$ is a field follows from basic ring theory: the quotient of a ring by a maximal ideal is a field. We want $\mathbb{R}^\omega/M$ to be an ordered field in which $\mathbb{R}$ embeds. The embedding will be $x \mapsto (x, x, x, \dots) + M$; this map is indeed injective, since $(x, x, x, \dots)$ is just $(x \colon k \in \omega)$, and

$$\begin{aligned} (x \colon k \in \omega) + M = (y \colon k \in \omega) + M &\iff \{k \in \omega \colon x \neq y\} \in \mathfrak{m} \\ &\iff x = y. \end{aligned}$$

We define the ordering by

$$a + M < b + M \iff \{k \in \omega \colon a_k \geqslant b_k\} \in \mathfrak{m}.$$

This is a valid definition, because if $a+M = a'+M$ and $b+M = b'+M$, then

$$\begin{aligned}&\{k \in \omega \colon a'_k \geqslant b'_k\} \\ &\qquad \subseteq \{k \in \omega \colon a_k \geqslant b_k\} \cup \{k \in \omega \colon a'_k \neq a_k\} \cup \{k \in \omega \colon b'_k \neq b_k\},\end{aligned}$$

so if $\{k \in \omega \colon a_k \geqslant b_k\}$ is small, then so is $\{k \in \omega \colon a'_k \geqslant b'_k\}$, and conversely by symmetry.

Similarly all of the axioms of ordered fields hold in $\mathbb{R}^\omega/M$, as we shall show generally tomorrow. Meanwhile, note that $(1, 2, 3, \dots) + M$ is an infinite element of $\mathbb{R}^\omega/M$, and so $(1, 1/2, 1/3, \dots) + M$ is an infinitesimal element. Of the elements

$$(0, 1, 0, 1, 0, 1, \dots) + M, \qquad (1, 0, 1, 0, 1, 0, \dots) + M.$$

Which one is greater? We don't know. Either the set of odd numbers or the set of even numbers is in $\mathfrak{m}$. If $\{\text{odd numbers}\} \in \mathfrak{m}$, then

$$\begin{aligned}(0, 1, 0, 1, \dots) + M &= (0, 0, 0, 0, \dots) + M \\ &< (1, 1, 1, 1, \dots) + M \\ &= (1, 0, 1, 0, \dots) + M.\end{aligned}$$

1.5 Saturday

We want to understand $\mathbb{R}^\omega/M$, where M is a maximal ideal of the ring $\mathbb{R}^\omega$. This is a special case of

$$\prod_{i \in \Omega} K_i \Big/ M,$$

where Ω is an arbitrary set, and each K_i is a field, so $\prod_{i\in\Omega} K_i$ is the ring consisting of $(a_i : i \in \Omega)$, where $a_i \in K_i$; and M is a maximal ideal of this ring.

For example, Ω could be the set of primes, and if $p \in \Omega$, K_p could be $\mathbb{F}_p$, that is, $\mathbb{Z}/p\mathbb{Z}$. If $\ell \in \Omega$, let $\ell = 0$ be the equation

$$\underbrace{1 + \cdots + 1}_{\ell} = 0.$$

This is true in $\mathbb{F}_p$ if and only if $p = \ell$. Then

$$\{p \in \Omega : \ell = 0 \text{ in } \mathbb{F}_p\} = \{\ell\}.$$

Let

$$\mathfrak{m} = \{\operatorname{supp}(x) : x \in M\}.$$

Assume $\mathscr{P}_\omega(\Omega) \subseteq \mathfrak{m}$; then $\{\ell\} \in \mathfrak{m}$, so $\ell = 0$ is not true in $\prod_{p\in\Omega} \mathbb{F}_p$. In particular, this field has characteristic 0. But it is a kind of "average" of the fields $\mathbb{F}_p$, each of which has characteristic p. It is an example of a **pseudofinite field.**

In general, let

$$\mathscr{K} = (K_i : i \in \Omega), \qquad \prod \mathscr{K} = \prod_{i\in\Omega} K_i.$$

There are homomorphisms

- $x \mapsto x + M$ from $\prod \mathscr{K}$ to $\prod \mathscr{K}/M$,
- $x \mapsto x_i$ from $\prod \mathscr{K}$ to K_i for each i in Ω.

If $a \in \prod \mathscr{K}$, we let the **interpretation** of a in $\prod \mathscr{K}/M$ be $a + M$; in K_i, a_i. Then a polynomial equation $f = g$ with **parameters** from $\prod \mathscr{K}$ has a meaning in $\prod \mathscr{K}/M$ and in each K_i. For example, $f = g$ could be

$$ab^2 + 5c = a^3c - 6d^4$$

for some a, b, c, and d in $\prod\mathscr{K}$. **Truth** is denoted by $\vDash$. Then

$$\prod\mathscr{K}/M \vDash f = g \iff \{i \in \Omega\colon K_i \vDash f \neq g\} \in \mathfrak{m},$$

where again $\mathfrak{m} = \{\operatorname{supp}(x)\colon x \in M\}$. We can understand the equation $f = g$ as $\varphi(a^0, \dots, a^{n-1})$ for some polynomial equation $\varphi(x^0, \dots, x^{n-1})$ in no parameters, in the signature $\{0, 1, -, +, \times\}$ of rings. (The superscripts are not exponents, but indices.) In the example above, φ is

$$x^0(x^1)^2 + 5x^2 = (x^0)^3x^2 - 6(x^3)^4.$$

In general, φ is an **atomic formula.** Other **formulas** are obtained from these by use of $\wedge$ (and), $\neg$ (not), and $\exists x$ (there exists x). A **sentence** is a formula with no **free variables.** In the formula

$$\forall x\ (0 < |x - a| < \delta \Rightarrow |f(x) - L| < \varepsilon),$$

- x is a bound variable (not free),
- a and L are parameters,
- δ and ε are free variables,
- $<$, $|\cdot|$, and f are from a larger signature than that of plain rings.

A formula $\sigma \Rightarrow \tau$ is an abbreviation for $\neg\sigma \vee \tau$; and $\sigma \vee \tau$, for $\neg(\neg\sigma \wedge \neg\tau)$.

For any ring R and sentences σ and τ,

$$\begin{aligned} R \vDash \neg\sigma &\iff R \nvDash \sigma, \\ R \vDash \sigma \wedge \tau &\iff R \vDash \sigma\ \&\ R \vDash \tau, \\ R \vDash \exists x\ \varphi(x) &\iff \text{for some } a \text{ in } R,\ R \vDash \varphi(a). \end{aligned}$$

Here $\iff$ and & are abbreviations for the English expressions "if and only if" and "and" respectively; and $R \nvDash \sigma$ means σ is not true in R.

Given a sentence σ with parameters from $\prod \mathscr{K}$, let

$$\|\sigma\| = \{i \in \Omega \colon K_i \vDash \sigma\}.$$

Following Wilfrid Hodges in his *Model Theory* [1], we may call $\|\sigma\|$ the **Boolean value** of σ. Then

$$\begin{aligned} \|\sigma \wedge \tau\| &= \|\sigma\| \cap \|\tau\|, \\ \|\neg\sigma\| &= \Omega \smallsetminus \|\sigma\|, \\ \|\sigma \vee \tau\| &= \|\sigma\| \cup \|\tau\|. \end{aligned}$$

Because $\mathfrak{m}$ is an ideal,

$$\|\sigma \vee \tau\| \in \mathfrak{m} \iff \|\sigma\| \in \mathfrak{m} \ \& \ \|\tau\| \in \mathfrak{m}. \tag{7}$$

Because $\mathfrak{m}$ is a *maximal* ideal,

$$\|\sigma\| \notin \mathfrak{m} \iff \|\neg\sigma\| \in \mathfrak{m}, \tag{8}$$

$$\|\sigma \wedge \tau\| \in \mathfrak{m} \iff \|\sigma\| \in \mathfrak{m} \text{ OR } \|\tau\| \in \mathfrak{m}. \tag{9}$$

Why? Recall that

$$\mathfrak{m} \subseteq \mathscr{P}(\Omega), \tag{10}$$

$$X, Y \in \mathfrak{m} \iff X \cup Y \in \mathfrak{m}, \tag{11}$$

$$Y \in \mathfrak{m} \ \& \ X \subseteq Y \implies X \in \mathfrak{m}, \tag{12}$$

$$X \notin \mathfrak{m} \iff \Omega \smallsetminus X \in \mathfrak{m}. \tag{13}$$

But (12) follows from the other three conditions. We have that (7) follows from (11), and (8) from (13). Finally, (9) follows from (7) and (8), because

$$\begin{aligned} \|\sigma \wedge \tau\| \in \mathfrak{m} &\iff \|\neg\sigma \vee \neg\tau\| \notin \mathfrak{m} && \text{[by (8)]} \\ &\iff \|\neg\sigma\| \notin \mathfrak{m} \text{ OR } \|\neg\sigma\| \notin \mathfrak{m} && \text{[by (7)]} \\ &\iff \|\sigma\| \in \mathfrak{m} \text{ OR } \|\tau\| \in \mathfrak{m}. && \text{[by (8)]} \end{aligned}$$

Note that (9) means $\mathfrak{m}$ is a *prime ideal.* In general, a *proper* ideal I of a ring R is a **prime ideal** if

$$xy \in I \iff x \in I \text{ OR } y \in I.$$

The improper ideal is not counted as prime, just as 1 is not counted as a prime number. In $\mathbb{Z}$, the prime ideals are (p) and (0); the ideals (p) are also maximal ideals, but (0) is not, since

$$(0) \subset (p) \subset \mathbb{Z}.$$

In every ring, all maximal ideals are prime. Indeed, if I is a maximal ideal of R, then R/I is a field, hence an **integral domain:** this means

$$xy = 0 \implies x = 0 \text{ OR } y = 0,$$

which in the present context means

$$xy + I = I \implies x + I = I \text{ OR } y + I = I,$$

that is, $xy \in I \implies x \in I$ OR $y \in I$. The following then is a corollary of the Maximal Ideal Theorem:

Theorem 12 (Prime Ideal Theorem)**.** *Every proper ideal of a ring is included in a prime ideal of the ring.*

A reason for stating this result separately is that it is strictly weaker than the Maximal Ideal Theorem: see page 50. However, the two theorems are equivalent for *Boolean* rings, where all prime ideals are maximal (Exercise).

The following is a special case of *Łoś's Theorem:*

Theorem 13. *If $\mathscr{K}$ is an indexed family $(K_i : i \in \Omega)$ of fields, and M is a maximal ideal of $\prod \mathscr{K}$, then for all sentences σ with parameters from $\prod \mathscr{K}$,*

$$\prod \mathscr{K}/M \vDash \sigma \iff \|\sigma\| \notin \mathfrak{m}. \tag{14}$$

Proof. We know this is true when σ is atomic. Suppose it is true when σ is ρ or τ. Then

$$\begin{aligned}\prod\mathscr{K}/M \vDash \neg\rho &\iff \prod\mathscr{K}/M \nvDash \rho\\ &\iff \|\rho\| \in \mathfrak{m}\\ &\iff \|\neg\rho\| \notin \mathfrak{m},\end{aligned}$$

and similarly

$$\begin{aligned}\prod\mathscr{K}/M \vDash \rho\wedge\tau &\iff \prod\mathscr{K}/M \vDash \rho \;\;\&\;\; \prod\mathscr{K}/M \vDash \tau\\ &\iff \|\rho\| \notin \mathfrak{m} \;\;\&\;\; \|\tau\| \notin \mathfrak{m}\\ &\iff \|\rho\wedge\tau\| \notin \mathfrak{m}.\end{aligned}$$

Finally, suppose for some formula $\varphi(x)$, for all a in $\prod\mathscr{K}$, (14) holds when σ is $\varphi(a)$. Let a be $(a_i\colon i\in\Omega)$ in $\prod\mathscr{K}$ such that, for all i in Ω, if $K_i \vDash \exists x\ \varphi(x)$, then $K_i \vDash a_i$. Then

$$\begin{aligned}\|\exists x\ \varphi(x)\| &= \{i\in\Omega\colon K_i \vDash \exists x\ \varphi(x)\}\\ &= \{i\in\Omega\colon K_i \vDash \varphi(a_i)\}\\ &= \{i\in\Omega\colon K_i \vDash \varphi(a)\}\\ &= \|\varphi(a)\|.\end{aligned}$$

Moreover, for all b in $\prod\mathscr{K}$,

$$\|\varphi(b)\| \subseteq \|\varphi(a)\|.$$

Hence

$$\begin{aligned}\prod\mathscr{K}/M \vDash \exists x\ \varphi(x) &\iff \text{for some } b \text{ in } \prod\mathscr{K},\ \prod\mathscr{K}/M \vDash \varphi(b)\\ &\iff \text{for some } b \text{ in } \prod\mathscr{K},\ \|\varphi(b)\| \notin \mathfrak{m}\\ &\iff \|\varphi(a)\| \notin \mathfrak{m}\\ &\iff \|\exists x\ \varphi(x)\| \notin \mathfrak{m}.\end{aligned}$$

□

In the proof, we use the Axiom of Choice to find a; see the final lecture (page 60). The whole proof uses pure logic: nothing of ring theory (except for the algebra of $\mathscr{P}(\Omega)$). The fields K_i could be replaced by groups, linear orders, ordered fields, or something else. Then we get *Łoś's Theorem,* namely Theorem 1.6.

In case each K_i is $\mathbb{R}$, and the parameters of σ are from $\mathbb{R}$ only, not $\mathbb{R}^\omega$, we get $\|\sigma\| \in \{\varnothing, \Omega\}$, so

$$\mathbb{R}^\Omega/M \vDash \sigma \iff \mathbb{R} \vDash \sigma.$$

We used this to prove Theorem 3 (page 14).

1.6 Sunday

We have proved:

Łoś's Theorem. *For every indexed family* $(\mathfrak{A}_i \colon i \in \Omega)$ *of* structures *having a common* signature $\mathscr{S}$*, for every maximal ideal* $\mathfrak{m}$ *of* $\mathscr{P}(\Omega)$*, there is a structure* $\mathfrak{B}$ *of* $\mathscr{S}$ *such that, for all sentences* σ *of* $\mathscr{S}$*, if*

$$\|\sigma\| = \{i \in \Omega \colon \mathfrak{A} \vDash \sigma\},$$

then

$$\mathfrak{B} \vDash \sigma \iff \|\sigma\| \notin \mathfrak{m}.$$

(This much is the ***Compactness Theorem****; again, see the final lecture, page 60.) The universe of* $\mathfrak{B}$ *can be taken as* $\prod_{i\in\Omega} A_i/{\sim}$*, where*

$$a \sim b \iff \{i \in \Omega \colon a_i \neq b_i\} \in \mathfrak{m}.$$

Here a **structure** $\mathfrak{A}$ is a set A, called the **universe** of $\mathfrak{A}$, together with some operations and relations on A, that is, functions from A^n to A and subsets of A^n for various n in ω. The **signature** of the structure consists of a symbol for each of the operations and relations. For example,

- $\mathfrak{G} = (G, 0, -, +)$, an abelian group;
- $\mathfrak{R} = (R, 0, 1, -, +, \times)$, a ring;
- $\mathfrak{L} = (L, <)$, a linear order.

Note that $A^0 = \{\varnothing\} = \{0\} = 1$. Hence a **nullary** operation on A, that is, a function from A^0 to A, is determined by an element of a. In Łoś's Theorem, let

$$A = \prod_{i \in \Omega} A_i.$$

Then each $\mathfrak{A}_i$ can be considered as a structure of $\mathscr{S}(A)$, the elements of A being treated as constants, a being interpreted in $\mathfrak{A}_i$ as a_i.

We are interested in Łoś's Theorem when

- Ω is ω,
- $\mathfrak{m}$ is a nonprincipal maximal ideal of $\mathscr{P}(\omega)$, that is, $\mathscr{P}_\omega(\omega) \subseteq \mathfrak{m}$, and
- each $\mathfrak{A}_i$ is $\mathbb{R}$ with its full structure, that is, there is a symbol for *every* operation and relation on $\mathbb{R}$.

Let

$$^*\mathbb{R} = \mathbb{R}^\omega / M,$$

where $M = \{x \in \mathbb{R}^\omega \colon \operatorname{supp}(x) \in \mathfrak{m}\}$. We treat the embedding $x \mapsto (x, x, x, \dots) + M$ of $\mathbb{R}$ in $^*\mathbb{R}$ as an inclusion. If $f \colon \mathbb{R} \to \mathbb{R}$, then, as a symbol, f is interpreted in $^*\mathbb{R}$ as a function $^*f$ from $^*\mathbb{R}$ to $^*\mathbb{R}$ given by

$$^*f(a + M) = (f(a_k) \colon k \in \omega) + M,$$

as in Figure 2. Then $^*f$ can be called the **extension** of f. One must show that $^*f$ is well defined (or take this as a consequence of Łoś's Theorem). Likewise,

$$^*\mathbb{N} = \{x + M \colon x \in \mathbb{N}^\omega\}.$$

$$\begin{array}{ccccccc} a+M & = & (a_0, & a_1, & a_2, & \dots)+M \\ \downarrow & & & & & \\ {}^*f(a+M) & = & (f(a_0), & f(a_1), & f(a_2), & \dots)+M \end{array}$$

Figure 2. Extension of a function

Theorem 14. $\mathbb{N}$ *consists of the finite elements of* ${}^*\mathbb{N}$.

Proof. By definition, the elements of $\mathbb{N}$ are finite. Suppose n is a finite element of ${}^*\mathbb{N}$. Then the set $\{x \in \mathbb{N} \colon n < x + 1\}$ is nonempty, so it has a least element, m. Then

$$m \leqslant n < m + 1,$$

unless $n < 1$; but this cannot happen since

$$\mathbb{R} \vDash \forall x\ (x \in \mathbb{N} \implies 1 \leqslant x),$$

so

$${}^*\mathbb{R} \vDash \forall x\ (x \in {}^*\mathbb{N} \implies 1 \leqslant x).$$

Likewise, since

$$\mathbb{R} \vDash \forall x\ (x \in \mathbb{N}\ \&\ m \leqslant x < m + 1 \implies m = x),$$

we have also

$${}^*\mathbb{R} \vDash \forall x\ (x \in {}^*\mathbb{N}\ \&\ m \leqslant x < m + 1 \implies m = x),$$

and therefore $m = n$, so $n \in \mathbb{N}$. □

See Figure 3. In proving the theorem, we use that $\mathbb{N}$ is **well ordered,** that is, every nonempty subset has a least element. Is ${}^*\mathbb{N}$ well ordered? No: ${}^*\mathbb{N} \smallsetminus \mathbb{N}$ has no least element, since if n is infinite, so is $n - 1$.

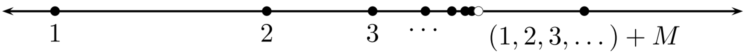

Figure 3. Non-standard natural numbers

By Łoś's Theorem, we have (and have used)

$$\mathbb{R} \preccurlyeq {}^*\mathbb{R}; \tag{15}$$

this means all **first-order** sentences that are true in $\mathbb{R}$ are true in ${}^*\mathbb{R}$. first-order sentences are the kinds of sentences defined on page 26. In a first-order sentence, variables refer to individuals, not subsets.

If $a\colon \mathbb{N} \to \mathbb{R}$, that is, $a \in \mathbb{R}^{\mathbb{N}}$, or

$$a = (a_k \colon k \in \mathbb{N}),$$

then we obtain ${}^*a$ from ${}^*\mathbb{N}$ to ${}^*\mathbb{R}$, that is,

$${}^*a = (a_k \colon k \in {}^*\mathbb{N}).$$

This is the extension of a, as ${}^*f$ is the extension of f, and ${}^*\mathbb{N}$ of $\mathbb{N}$. If $k = (k(i) \colon i \in \omega) + M$, then

$$a_k = (a_{k(i)} \colon i \in \omega) + M.$$

The sequence a is **bounded** if

$$\exists R\, \forall n\, (n \in \mathbb{N} \implies |a_n| \leqslant R); \tag{16}$$

this applies to ${}^*a$ if we replace $\mathbb{N}$ with ${}^*\mathbb{N}$. Then what condition on ${}^*a$ is equivalent to boundedness of a?

Theorem 15. *A sequence a in $\mathbb{R}^{\mathbb{N}}$ is bounded if and only if every element of ${}^*a$ is finite.*

Proof. If a is bounded, then for some S in $\mathbb{R}$,

$$\mathbb{R} \vDash \forall n\,(n \in \mathbb{N} \implies |a_n| \leqslant S),$$
$$^*\mathbb{R} \vDash \forall n\,(n \in {}^*\mathbb{N} \implies |a_n| \leqslant S),$$

so $^*a$ has no infinite elements (since S is finite).

If a is not bounded, then

$$\mathbb{R} \vDash \forall S\,\exists n\,(n \in \mathbb{N} \;\&\; |a_n| > S),$$
$$^*\mathbb{R} \vDash \forall S\,\exists n\,(n \in {}^*\mathbb{N} \;\&\; |a_n| > S).$$

Letting S be positive and infinite, we obtain that a_n is infinite for some n in $^*\mathbb{N}$. □

Note how we had to manipulate the quantifiers. The *standard* definition of boundedness of a sequence is first order; it is the condition in (16). The theorem shows that this condition is equivalent to a *non-standard* definition that, as such, is not expressed in a first-order way.

The same is true of Theorem 3 (page 14). By analogy with limits of functions, we have the following; the proof is an EXERCISE.

Theorem 16. $\lim_{n\to\infty} a_n = L$ *if and only if, for all infinite* n, $a_n \simeq L$.

A *filter* of a Boolean ring $\mathscr{P}(\Omega)$ is "dual" to an ideal of the ring. By "dualizing" the conditions in Theorem 6 (page 19), we say that a subset F of $\mathscr{P}(\Omega)$ is a **filter of** $\mathscr{P}(\Omega)$, or a **filter on** Ω, if

$$X, Y \in F \implies X \cap Y \in F,$$
$$X \in F \;\&\; X \subseteq Y \subseteq \Omega \implies Y \in F,$$
$$\Omega \in F.$$

Then $\{\Omega \smallsetminus X : X \in F\}$ is an ideal, namely the **dual ideal** of the filter. A maximal proper filter is called an **ultrafilter;** its dual ideal is also its complement and is a maximal ideal.

In Łoś's Theorem, if $\mathscr{U} = \mathscr{P}(\Omega) \smallsetminus \mathfrak{m}$, then the structure $\mathfrak{B}$ is denoted by

$$\prod_{i\in\Omega} \mathfrak{A}_i \Big/ \mathscr{U}$$

(or $\prod_{\mathscr{U}} \mathfrak{A}_i$ or some such) and is called the **ultraproduct** of the structures $\mathfrak{A}_i$.

2 Second week

2.1 Monday

Recall that **infinite** means greater than all n in $\mathbb{N}$, where

$$\mathbb{N} = \{1, 2, 3, \dots\};$$

infinitesimal means less in absolute value than $1/n$ for all n in $\mathbb{N}$; $a_n \simeq L$ means $a_n - L$ is infinitesimal. $\mathscr{U}$ is an **ultrafilter** on ω (which is $\{0, 1, 2, \dots\}$); this means $\mathscr{U} \subseteq \mathscr{P}(\omega)$ and, for all X and Y in $\mathscr{P}(\omega)$,

$$X, Y \in \mathscr{U} \iff X \cap Y \in \mathscr{U},$$
$$X \in \mathscr{U} \iff \omega \smallsetminus X \notin \mathscr{U}.$$

Indeed, these conditions imply

$$X \in \mathscr{U} \ \& \ X \subseteq Y \subseteq \omega \implies Y \in \mathscr{U},$$

since if $X \subseteq Y$, then $X \cap Y = X$. We define an equivalence relation on $\mathbb{R}^\omega$ so that, if $a/\mathscr{U}$ is the equivalence class of an element a, that is, $(a_k : k \in \omega)$, of $\mathbb{R}^\omega$, then

$$a/\mathscr{U} = b/\mathscr{U} \iff \{k \in \omega : a_k = b_k\} \in \mathscr{U}.$$

The elements of $\mathscr{U}$ are **large.** Similarly we obtain ${}^*\mathbb{N}$. If $a \in \mathbb{R}^\mathbb{N}$, we obtain ${}^*a$, namely $(a_n : n \in {}^*\mathbb{N})$, where

$$a_n = \begin{cases} a_n, & \text{if } n \in \mathbb{N}, \\ (a_{n_k} : k \in \omega)/\mathscr{U}, & \text{if } n = (n_k : k \in \omega)/\mathscr{U}. \end{cases}$$

If $n \in \mathbb{N}$, we identify n with $(n, n, n, \dots)/\mathscr{U}$; thus, in the definition of a_n for n in ${}^*\mathbb{N}$, it suffices to give only the second case.

Theorem 16 is that $\lim_{n\to\infty} a_n = L$ if and only if, for all infinite n, $a_n \simeq L$. Suppose we try to prove the foreward direction as follows. Assuming $\lim_{n\to\infty} a_n = L$, we have that for some positive ε in $\mathbb{R}$,

$$\mathbb{R} \vDash \exists M\ \forall n\ (n \in \mathbb{N}\ \&\ n \geqslant M \implies |a_n - L| < \varepsilon),$$
$${}^*\mathbb{R} \vDash \exists M\ \forall n\ (n \in {}^*\mathbb{N}\ \&\ n \geqslant M \implies |a_n - L| < \varepsilon).$$

This is true, but is not what we want. Rather, we want to observe that, for some positive ε in $\mathbb{R}$, for some M in $\mathbb{R}$,

$$\mathbb{R} \vDash \forall n\ (n \in \mathbb{N}\ \&\ n \geqslant M \implies |a_n - L| < \varepsilon),$$
$${}^*\mathbb{R} \vDash \forall n\ (n \in {}^*\mathbb{N}\ \&\ n \geqslant M \implies |a_n - L| < \varepsilon).$$

Thus, if n is infinite, then $|a_n - L| < \varepsilon$. This being true for all positive real ε, $a_n \simeq L$.

Suppose conversely $\lim_{n\to\infty} a_n \neq L$. Then for some positive ε in $\mathbb{R}$,

$$\mathbb{R} \vDash \forall M\ \exists n\ (n \in \mathbb{N}\ \&\ n \geqslant M\ \&\ |a_n - L| \geqslant \varepsilon),$$
$${}^*\mathbb{R} \vDash \forall M\ \exists n\ (n \in {}^*\mathbb{N}\ \&\ n \geqslant M\ \&\ |a_n - L| \geqslant \varepsilon).$$

Letting M be positive and infinite, we find infinite n so that $|a_n - L| \geqslant \varepsilon$ and therefore $a_n \not\simeq L$.

If you grew up using δ-ε proofs, but now want to do infinitesimal calculus, then you have to prove things like the foregoing. Alternatively, you can just define $\lim_{n\to\infty} a_n = L$ to mean $a_n \simeq L$ for all infinite n. Then other proofs become easier.

Lemma 1. *Suppose a in $\mathbb{R}^{\mathbb{N}}$ is convergent. Then every term of ${}^*a$ is finite.*

Theorem 17. *Suppose $a, b \in \mathbb{R}^{\mathbb{N}}$ and $\lim(a) = L$, $\lim(b) = M$, and $r \in \mathbb{R}$. Then*

$$\lim(a+b) = L + M,$$
$$\lim(ra) = rL,$$
$$\lim(ab) = LM,$$

and if $L \neq 0$, then

$$\lim a^{-1} = L^{-1}. \tag{17}$$

Proof. Follow the method of the non-standard proofs of Theorems 1 and 2, using that the infinitesimals of ${}^*\mathbb{R}$ compose an ideal of the ring of finite elements. For (17), we need the foregoing lemma—an EXERCISE—that a_n is finite for all n. If $L \neq 0$, possibly some a_n are 0, but we prove that $a_n \neq 0$ if n is infinite. □

Theorem 18. *A sequence a is convergent if and only if, for all infinite m and n, $a_m \simeq a_n$.*

The proof, an EXERCISE, will want *standard parts.* If a is a finite element of ${}^*\mathbb{R}$, its **standard part** is an element $\operatorname{st}(a)$ of $\mathbb{R}$ such that $\operatorname{st}(a) \simeq a$. We achieve this by defining

$$\operatorname{st}(a) = \sup\{x \in \mathbb{R} \colon x < a\}.$$

(compare (4) on page 12). To prove that this works, consider the alternative whereby $0 < \delta < |a - \operatorname{st}(a)|$ for some δ in $\mathbb{R}$.

We showed yesterday that, since ${}^*\mathbb{N}$ is not well ordered, the property of being well ordered is not first order. Similarly being complete is not first order, since the set of finite elements of ${}^*\mathbb{R}$ has no supremum.

However, in the two-sorted structure $\mathbb{N} \sqcup \mathscr{P}(\mathbb{N})$, the first-order sentence

$$\forall Y \ \bigl(1 \in Y \ \ \& \ \ \forall x \ (x \in Y \implies x + 1 \in Y) \implies \forall x \ x \in Y\bigr),$$

so it is true in ${}^*(\mathbb{N} \sqcup \mathscr{P}(\mathbb{N}))$, which can be understood as ${}^*\mathbb{N} \sqcup {}^*\mathscr{P}(\mathbb{N})$, where ${}^*\mathscr{P}(\mathbb{N})$ consists of the subsets

$$\left\{ x/\mathscr{U} : x \in \prod_{k \in \omega} Y_k \right\}$$

of ${}^*\mathbb{N}$, where $(Y_k : k \in \omega) \in \mathscr{P}(\mathbb{N})^\omega$. Thus ${}^*\mathscr{P}(\mathbb{N}) \subseteq \mathscr{P}({}^*\mathbb{N})$, but the inclusion is proper, since for example

$$\mathbb{N} \in \mathscr{P}({}^*\mathbb{N}) \smallsetminus {}^*\mathscr{P}(\mathbb{N}).$$

2.2 Tuesday

If $a \simeq b$, and b is finite, why must a be finite? If $a \simeq b$, and b is finite, then

$$|a| - |b| \leqslant |a - b| \leqslant 1,$$
$$|a| \leqslant 1 + |b|,$$

so a is finite.

Since all terms of convergent sequences are finite, convergent sequences are bounded.

What is the standard proof of this? If $\lim(a) = L$, then for some N, if $n > N$, then $|a_n - L| < 1$, so $|a_n| < |L| + 1$. Now let $M = \min\{|a_1|, \ldots, |a_N|, |L| + 1\}$. Then $|a_n| \leqslant M$ for all n in $\mathbb{N}$.

An arbitrary subset A of $\mathbb{R}$ is **bounded** if

$$\exists x\, \forall y\, (y \in A \implies |y| \leqslant x).$$

Theorem 19. *A subset A of $\mathbb{R}$ is bounded if and only if every element of ${}^*A$ is finite.*

A **limit point** of a subset E of $\mathbb{R}$ is an element p of $\mathbb{R}$ such that, for all positive ε,

$$(p - \varepsilon, p + \varepsilon) \cap (E \setminus \{p\}) \neq \varnothing,$$

that is,

$$\exists x\, (x \in E \;\&\; 0 < |x - p| < \varepsilon). \tag{18}$$

What is the non-standard formulation of this definition?

Theorem 20. *p is a limit point of E if and only if, for some q in ${}^*E$,*

$$q \simeq p \;\&\; q \neq p. \tag{19}$$

Proof. For the forward direction, let q be x in (18) when ε is a positive infinitesimal. More precisely, we have

$$\mathbb{R} \vDash \forall \varepsilon\, (\varepsilon > 0 \implies \exists q\, (q \in E \;\&\; 0 < |q - p| < \varepsilon)),$$
$${}^*\mathbb{R} \vDash \forall \varepsilon\, (\varepsilon > 0 \implies \exists q\, (q \in {}^*E \;\&\; 0 < |q - p| < \varepsilon)),$$

so if we let ε be a positive infinitesimal, we have q as desired.

Conversely, suppose (19) holds. Then for all positive ε in $\mathbb{R}$, (18) holds in ${}^*\mathbb{R}$, so it holds in $\mathbb{R}$: that is, we have $0 < |q - p| < \varepsilon$, and so,

$${}^*\mathbb{R} \vDash \exists x\, (x \in {}^*E \;\&\; 0 < |x - p| < \varepsilon),$$
$$\mathbb{R} \vDash \exists x\, (x \in E \;\&\; 0 < |x - p| < \varepsilon).$$

This being true for all positive ε in $\mathbb{R}$, p is a limit point of E. □

An alternative proof of the reverse direction uses the construction of ${}^*\mathbb{R}$ as follows. We can write out q as $(q_k\colon k \in \omega)/\mathscr{U}$, while p is just $(p, p, p, \dots)/\mathscr{U}$. Then for all n in $\mathbb{N}$, since $|q - p| < 1/n$, we have

$$\{k \in \omega\colon |q_k - p| < 1/n\} \in \mathscr{U}.$$

Also, since $q \neq p$,

$$\{k \in \omega\colon q_k \neq p\} \in \mathscr{U}.$$

Since $\mathscr{U}$ is closed under $\cap$, we conclude

$$\{k \in \omega\colon |q_k - p| < 1/n \;\;\&\;\; q_k \neq p\} \in \mathscr{U}.$$

Finally, since $q \in {}^*E$, we have

$$\{k \in \omega\colon q_k \in E\} \in \mathscr{U}.$$

Indeed, we can consider ${}^*E$ as $E^\omega/\mathscr{U}$, but more precisely

$${}^*E = \{x/\mathscr{U} \in {}^*\mathbb{R}\colon x \in E^\omega\};$$

this allows for the possibility that $x/\mathscr{U} = y/\mathscr{U}$, but $y \notin E^\omega$. Summing up,

$$\{k \in \omega\colon |q_k - p| < 1/n \;\;\&\;\; q_k \neq p \;\;\&\;\; q_k \in E\} \in \mathscr{U}.$$

Thus for all n in $\mathbb{N}$, there is q in E such that $q \neq p$, but $|q - p| < 1/n$. Therefore p is a limit point of E.

In this alternative argument, it does not follow that $\lim_{n\to\infty} q_n = p$. For example, possibly $\{\text{odds}\} \in \mathscr{U}$, and $q_{2k} = 2k$, so $(q_n\colon n \in \omega)$ diverges. But there is still a subsequence that converges to p.

The definitions of bounded sets and limit points, as well as the following theorem, can be adapted to apply to $\mathbb{R}^n$ for any n in $\mathbb{N}$. We stick with $\mathbb{R}$ for simplicity.

Theorem 21 (Bolzano–Weierstrass). *A bounded infinite subset of* $\mathbb{R}$ *has a limit point.*

The standard proof is to "divide and conquer": If the infinite subset E of $\mathbb{R}$ is bounded, it is included in a closed bounded interval. If we divide the interval in half, then at least one of the halves contains infinitely many points of E. Then we divide that interval in two, and continue. The sequence of left endpoints of the intervals that we find has a limit, which is a limit point of E.

Non-standard proof. Since E is infinite, there is a nonrepeating sequence $(a_n : n \in \omega)$ such that each term is in E. Let n be infinite. Since a_n is finite, it has a standard part b. Since $a_n \simeq b$, if $a_n \neq b$ we are done. Suppose $a_n = b$. Then for some finite k, $a_k = b$, so $a_k = a_n$. Thus $(a_n : n \in \omega)$ repeats. □

A neater non-standard proof uses the following.

Theorem 22. *A subset E of $\mathbb{R}$ is finite if and only if ${}^*E = E$.*

The proof is an EXERCISE; or see page 45. Meanwhile, if E is infinite and bounded, we can find a in ${}^*E \smallsetminus E$ and then let $b = \operatorname{st}(a)$; then b is a limit point of E.

2.3 Wednesday

Recall that p is a *limit point* of A if there is q in ${}^*A$ such that $q \neq p$, but $q \simeq p$. (This makes sense in $\mathbb{R}^n$ for all n in $\mathbb{N}$, not just in $\mathbb{R}$; but again, we shall officially stay with $\mathbb{R}$.) A set is **closed** if it contains all of its limit points. The complement of a closed set is **open.** Then a subset U of $\mathbb{R}$ is open if and only if, for all p in U, if $q \in {}^*\mathbb{R}$ and $p \simeq q$, then $q \in {}^*U$. Thus uses that ${}^*\mathbb{R}$ is the disjoint union of ${}^*U$ and ${}^*(\mathbb{R} \smallsetminus U)$:

$$ {}^*\mathbb{R} \smallsetminus {}^*(\mathbb{R} \smallsetminus U) = {}^*U. $$

If $p \in \mathbb{R}$, let

$$\mu(p) = \{q \in {}^*\mathbb{R} \colon q \simeq p\},$$

the **monad** of p. This is Robinson's term [3, pp. 57, 90], but whether Robinson is alluding to Leibniz's philosophical use of the term is not clear. In passing from $\mathbb{R}$ to ${}^*\mathbb{R}$, each point is replaced by a "cloud"—its monad, monads of distinct points being disjoint.

In any case, we now have that U is open if and only if, for all p in U,

$$\mu(p) \subseteq U.$$

Theorem 23. *The open subsets of* $\mathbb{R}$ *are just the unions of sets of open intervals.*

Proof. Suppose U is open. Then for every p in U, $\mu(p) \subseteq U$, and so $(p - \delta, p + \delta)$ for infinitesimal positive δ. Thus

$$\begin{gathered} {}^*\mathbb{R} \vDash \exists x \; {}^*(p - x, p + x) \subseteq {}^*U, \\ \mathbb{R} \vDash \exists x \; (p - x, p + x) \subseteq U, \end{gathered}$$

so for some positive δ_p in $\mathbb{R}$, $(p - \delta_p, p + \delta_p) \subseteq U$. Thus

$$U = \bigcup\{(p - \delta_p, p + \delta_p) \colon p \in U\}.$$

The converse is easy. □

Porism 1. *The intersection of every set of closed sets is a closed set. Every closed interval is a closed set.*

Not every closed set is a union of closed intervals. Examples are singleton sets and $\{1/n \colon n \in \mathbb{N}\} \cup \{0\}$. An uncountable example is the **Cantor set,** which results from starting with the closed interval $[0, 1]$, removing the open interval $(1/3, 2/3)$, then removing the middle thirds of the remaining intervals, and so on.

Theorem 24. *A subset A of $\mathbb{R}$ is closed and bounded if and only if for all q in ${}^*A$ there is p in A such that $p \simeq q$.*

Proof. Suppose A is closed and bounded, and $q \in {}^*A$. Since A is bounded, q is finite, by Theorem 19. Now let p be the standard part of q. Either $p = q$, or p is a limit point of A. In either case, $p \in A$, since this is closed.

Suppose conversely for all q in ${}^*A$ there is p in A such that $p \simeq q$. Since p is finite, q must be finite. Thus A is bounded. Suppose p is a limit point of A. For some q in ${}^*A$, $p \simeq q$, but $p \neq q$. But for some p' in A, $p' \simeq q$. Therefore $p \simeq p'$, so $p = p'$, and $p \in A$. □

We are going to talk about *compactness,* both in the present context and in logic. The **Compactness Theorem** of first-order logic is that if Γ is set of first-order sentences, and every finite subset of Γ has a model, then Γ has a model. (See page 59).

A collection $\mathscr{O}$ of open subsets of $\mathbb{R}$ is a **cover** (or **open cover**) of A if

$$A \subseteq \bigcup \mathscr{O}.$$

Then A is **compact** if for all open covers $\mathscr{O}$ of A, there is a finite subset of $\mathscr{O}$ that covers A.

Theorem 25 (Heine–Borel)**.** *A subset A of $\mathbb{R}$ (or $\mathbb{R}^n$) is compact if and only if closed and bounded (that is, for all q in ${}^*A$ there is p in A such that $q \in \mu(p)$).*

Proof. ($\Rightarrow$) (Standard) $A \subseteq \bigcup\{(-n, n) \colon n \in \mathbb{N}\}$. If $p \notin A$, then

$$A \subseteq \bigcup\{\mathbb{R} \smallsetminus [p - \delta, p + \delta] \colon \delta > 0\}.$$

($\Rightarrow$) (Non-standard) Suppose for some q in ${}^*A$, for all p in A, $q \notin \mu(p)$. But (EXERCISE)

$$\mu(p) = \bigcap\{{}^*U : U \text{ open and } p \in U\}.$$

So for some open set U_p containing p, $q \notin {}^*U_p$. Then $\{U_p : p \in A\}$ covers A, but $\{{}^*U_p : p \in A\}$ does not cover ${}^*A$. However, if $\{p(0), \dots, p(n)\}$ is a finite subset of A, then

$${}^*(U_{p(0)} \cup \dots \cup U_{p(n)}) = {}^*U_{p(0)} \cup \dots \cup {}^*U_{p(n)},$$

so $U_{p(0)} \cup \dots \cup U_{p(n)}$ cannot include A. That is, if

$$\mathbb{R} \vDash \forall x\ (x \in A \Rightarrow x \in U_{p(0)} \vee \dots \vee x \in U_{p(n)}),$$

then

$${}^*\mathbb{R} \vDash \forall x\ (x \in {}^*A \Rightarrow x \in {}^*U_{p(0)} \vee \dots \vee x \in {}^*U_{p(n)}),$$

which cannot be.

($\Leftarrow$) (Non-standard) Suppose $\mathscr{O}$ is an open cover of A with no finite subcover. We may assume $\mathscr{O}$ is countable by the **Lindelöf Covering Theorem:** For every p in A, there is U in $\mathscr{O}$ such that $p \in U$; but then for some a_p and b_p in $\mathbb{Q}$,

$$p \in (a_p, b_p)\ \ \&\ \ (a_p, b_p) \subseteq U.$$

Thus we can replace $\mathscr{O}$ with $\{(a_p, b_p) : p \in A\}$, which is countable. If this has a finite subcover of A, so does $\mathscr{O}$.

Say then $\mathscr{O} = \{U_k : k \in \omega\}$. Let

$$q_k \in A \smallsetminus \bigcup_{i<k} U_i = A \smallsetminus (U_0 \cup \dots \cup U_{k-1}),$$

$$q = (q_k : k \in \omega)/\mathscr{U}.$$

If $p \in A$, then $p \in U_k$ for some k; but then

$$\{\ell \in \omega : q_\ell \in U_k\} \subseteq \{0, 1, \dots, k\}$$

since

$$k < \ell \implies q_\ell \notin U_0 \cup \dots \cup U_k \cup \dots \cup U_{\ell-1}.$$

Since finite sets are small, $q \notin U_k$, so $q \notin \mu(p)$. □

In the non-standard proof of the sufficiency of the "$q \in \mu(p)$" condition, we have used more than the **Transfer Principle,** which is what we have written as $\mathbb{R} \preccurlyeq {}^*\mathbb{R}$ in (15) on page 33. Likewise we need more than the Transfer Principle for:

Proof of Theorem 22. Suppose $E = \{a_1, \dots, a_n\}$, but q is an element $(q_1, q_2, \dots)/\mathscr{U}$ of ${}^*E \smallsetminus E$. If

$$A_i = \{k \in \omega : q_k = a_i\},$$

then A_i is small; but

$$\omega = A_1 \cup \dots \cup A_n,$$

which is therefore small, which is absurd.

Conversely (or inversely), suppose $\{q_n : n \in \omega\} \subseteq A$, all q_n different from one another. Let $q = (q_n : n \in \omega)/\mathscr{U}$. If $a \in A$, then $q \neq a$, since the set

$$\{n \in \omega : q_n = a\}$$

has at most one element. □

Alternatively, for the first part, by the Transfer Principle we have

$$\mathbb{R} \vDash \forall x\ (x \in A \implies x = a_1 \lor \dots \lor x = a_n),$$
$${}^*\mathbb{R} \vDash \forall x\ (x \in {}^*A \implies x = a_1 \lor \dots \lor x = a_n).$$

For an example of the logical Compactness Theorem, given a structure $\mathfrak{A}$, let

$$\mathrm{Th}(\mathfrak{A}) = \{\sigma\colon \mathfrak{A} \vDash \sigma\}, \tag{20}$$

the **theory of** $\mathfrak{A}$. Here σ is a first-order sentence. Let

$$\sigma_n = \exists x_0, \dots x_{n-1} \bigwedge_{i<j<n} x_i \neq x_j.$$

By Compactness, $\bigcap\{\mathrm{Th}(\mathbb{F}_p)\colon p \text{ prime}\} \cup \{\sigma_n\colon n \in \mathbb{N}\}$ has a model. In fact $\prod_{p \text{ prime}} \mathbb{F}_p/\mathscr{U}$ is a model, $\mathscr{U}$ being a nonprincipal ultrafilter on $\{\text{primes}\}$.

By contrast, $\{\text{Peano axioms}\} \cup \{c \neq \overbrace{1 + \dots + 1}^{n}\colon n \in \mathbb{N}\}$ has no model; so the Peano axioms for $\mathbb{N}$ have no first-order formulation.

2.4 Friday

Again, the **open** subsets of $\mathbb{R}$ are just the unions of sets of open intervals. Let $\mathscr{O}$ be the set of open subsets of $\mathbb{R}$. Then

1. $\mathscr{X} \subseteq \mathscr{O} \implies \bigcup \mathscr{X} \in \mathscr{O}$.
2. In particular, $\varnothing \in \mathscr{O}$ (since $\bigcup \varnothing = \varnothing$).
3. If $X, Y \in \mathscr{O}$, then $X \cap Y \in \mathscr{O}$, since

$$\bigcup \mathscr{X} \cap \bigcup \mathscr{Y} = \bigcup \{Z \cap W\colon Z \in \mathscr{X} \ \ \& \ \ W \in \mathscr{Y}\}.$$

4. $\mathbb{R} \in \mathscr{O}$.

If A is an arbitrary set, and $\mathscr{O}$ is a subset of $\mathscr{P}(A)$ with the four properties above (with $A \in \mathscr{O}$ as the fourth property), then $\mathscr{O}$ is the set of **open subsets** of A in a **topology on** A, and $(A, \mathscr{O})$ is a **topological space.**

We are going to be interested in the case where A is the set of prime ideals of some ring R; this set is called

$$\mathrm{Spec}(R),$$

the **spectrum** of R. Recall then that $P \in \mathrm{Spec}(R)$ if and only if

$$x \in P \ \ \& \ \ y \in P \implies x + y \in P,$$
$$x \in P \ \text{ OR } \ y \in P \iff xy \in P,$$
$$0 \in P \ \ \& \ \ 1 \notin P.$$

For example:

- Writing $(n) = n\mathbb{Z} = \{nx \colon x \in \mathbb{Z}\}$, we have

  $$\mathrm{Spec}(\mathbb{Z}) = \{(p) \colon p \text{ prime}\} \cup \{(0)\}.$$

- If K is an algebraically closed field like $\mathbb{C}$, then

  $$\mathrm{Spec}(K[X]) = \{(X - a) \colon a \in K\} \cup \{(0)\},$$

  while $\mathrm{Spec}(K[X, Y])$ consists of the ideals
  - $(X - a, Y - b)$, where $a, b \in K$;
  - (f), where f is an irreducible element of $K[X, Y]$;
  - (0).

If R is a ring, and $a \in R$, let

$$[a] = \{P \in \mathrm{Spec}(R) \colon a \notin P\}.$$

If also $b \in R$, then

$$[a] \cap [b] = [ab].$$

Therefore unions of sets of sets $[a]$ are the open sets in a topology on $\mathrm{Spec}(R)$. The sets $[a]$ themselves are like open intervals.

As with closed bounded subsets of $\mathbb{R}$, a topological space $(A, \mathscr{O})$ is **compact** if for all subsets $\mathscr{X}$ of $\mathscr{O}$, if $\bigcup \mathscr{X} = A$, then for some finite subset $\mathscr{X}_0$ of $\mathscr{X}$,

$$\bigcup \mathscr{X}_0 = A.$$

The open subsets of $\operatorname{Spec}(\mathbb{Z})$ are just the complements of finite sets of ideals (p). In particular, (0) belongs to every nonempty open set. Then easily $\operatorname{Spec}(\mathbb{Z})$ is compact.

Theorem 26. *For all rings R,* $\operatorname{Spec}(R)$ *is compact.*

Standard proof. Suppose $A \subseteq R$ and $\bigcup\{[x] \colon x \in A\} = \operatorname{Spec}(R)$. Then

$$\bigcap\{[x]^{\mathrm{c}} \colon x \in A\} = \varnothing;$$

but in general

$$\bigcap\{[x]^{\mathrm{c}} \colon x \in A\} = \{P \in \operatorname{Spec}(R) \colon A \subseteq P\}.$$

By the Prime Ideal Theorem (page 28), A is included in no proper ideal. In general, the smallest ideal including A is

$$\{a_0x_0 + \cdots + a_{n-1}x_{n-1} \colon n \in \omega \ \ \& \ \ \boldsymbol{a} \in A^n \ \ \& \ \ \boldsymbol{x} \in R^n\};$$

this is denoted by

$$(A).$$

In the present case, $1 \in (A)$, so

$$1 = a_0x_0 + \dots a_{n-1}x_{n-1}$$

for some a_i in A. Then

$$[a_0]^{\mathrm{c}} \cap \cdots \cap [a_{n-1}]^{\mathrm{c}} = \varnothing,$$
$$[a_0] \cup \cdots \cup [a_{n-1}] = \operatorname{Spec}(R). \qquad \square$$

For a non-standard proof, if $(A, \mathscr{O})$ is a topological space, and Ω is some infinite set, and $\mathscr{U}$ is a nonprincipal ultrafilter on Ω, we define

$$^*A = A^\Omega/\mathscr{U}$$

as before. If $P \in A$, we now *define*

$$\mu(P) = \bigcap\{^*U \in \mathscr{O} : P \in U\}.$$

Then the non-standard proof of the Heine–Borel Theorem (page 43) gives us:

Theorem 27. *A topological space A is compact if and only if, for all choices of Ω and $\mathscr{U}$, for all Q in $^*A$ there is P in A such that*

$$Q \in \mu(P).$$

We must adjust the proof of sufficiency given for the Heine–Borel Theorem, since now we cannot assume $\mathscr{O}$ countable. We let Ω be the set of finite subsets of $\mathscr{O}$, and we let $\mathscr{U}$ be an ultrafilter on Ω that contains, for each E in $\mathscr{O}$, the set

$$\{X \in \Omega : E \in X\}.$$

(Such sets do generate a proper filter on Ω, since $\mathscr{O}$ is closed under taking intersections.) If $X \in \Omega$, we can let

$$q_X \in A \smallsetminus \bigcup X$$

The resulting point $(q_X : X \in \Omega)/\mathscr{U}$ of $^*A$ is not in $\mu(p)$. Now we can proceed with:

Non-standard proof of Theorem 26. Given Q in $^*\mathrm{Spec}(R)$, we are going to find P in $\mathrm{Spec}(R)$ such that

$$Q \in \mu(P).$$

Here

$$\mu(P) = \bigcap\{{}^*[x] \colon P \in [x]\} = \bigcap\{{}^*[x] \colon a \notin P\},$$

and for some $(Q_i \colon i \in \Omega)$ in $(\mathrm{Spec}(R))^\Omega$,

$$Q = (Q_i \colon i \in \Omega)/\mathscr{U},$$
$$Q \in {}^*[x] \iff \{i \in \Omega \colon x \notin Q_i\} \in \mathscr{U}.$$

Now define

$$P = \bigl\{x \in R \colon \{i \in \Omega \colon x \in Q_i\} \in \mathscr{U}\bigr\}.$$

Then $P \in \mathrm{Spec}(R)$ and $Q \in \mu(P)$. Indeed, if $x, y \in P$, then

$$\{i \in \Omega \colon x \in Q_i\}, \{i \in \Omega \colon y \in Q_i\} \in \mathscr{U},$$
$$\{i \in \Omega \colon x, y \in Q_i\} \in \mathscr{U},$$
$$\{i \in \Omega \colon x + y \in Q_i\} \in \mathscr{U}$$

(using that $\mathscr{U}$ is closed under intersections and taking supersets), so $x + y \in P$. Likewise

$$x \in P \ \text{OR} \ y \in P \iff xy \in P.$$

Finally, $0 \in P \ \& \ 1 \notin P$. □

The non-standard proof seems to work by magic. One need not recognize that if $1 \in (A)$, then $1 \in (A_0)$ for some finite subset A_0 of A.

In the remainder of the course, we want to establish the following:

- The Prime Ideal Theorem (page 28) is equivalent to the Compactness Theorem of first-order logic (pages 30, 43, and 59).
- The Maximal Ideal Theorem (page 22) is equivalent to:
  - the Prime Ideal Theorem with Łoś's Theorem (page 30);

- the Axiom of Choice;
- Zorn's Lemma (page 22).

In the background is Zermelo–Fraenkel set theory.

Theorem 28. *The Compactness Theorem implies the Prime Ideal Theorem.*

Proof. Supposing R is a ring, let $\operatorname{diag}(R)$ be the set of quantifier-free sentences in $\{0, 1, -, +, \times\}$, with parameters from R, that are true in R. Let P be a new singulary relation symbol, and let

$$\begin{aligned}\Gamma = \operatorname{diag}(R) \cup \{Pa \land Pb \Rightarrow P(a+b) \colon a, b \in R\}\\ \cup \{Pa \lor Pb \Leftrightarrow P(ab) \colon a, b \in R\} \cup \{P0, \neg P1\}.\end{aligned}$$

Then Γ has a model, because every finite subset Γ_0 of Γ has a model. Indeed, Γ_0 involves only finitely many parameters, so they generate a *countable* sub-ring R_0 of R.

The Prime Ideal Theorem holds for nontrivial countable rings. For, say $R_0 = \{b_n \colon n \in \omega\}$. Let $B_0 = \varnothing$ and

$$B_{n+1} = \begin{cases} B_n, & \text{if } (B_n \cup (b_n)) = R_0,\\ B_n \cup \{b_n\}, & \text{otherwise.}\end{cases}$$

Then $\bigcup_{n\in\omega} B_n$ is a maximal ideal of R_0, hence a prime ideal of R_0.

Now the symbol P can be interpreted as this prime ideal of R_0, so that R_0 is a model of Γ_0.

By Compactness, Γ has a model (S, Q). We may assume $R \subseteq S$, and then $R \cap Q$ is a prime ideal of R.

So far, we have shown only that R has a prime ideal. In particular, if I is a proper ideal of R, then R/I has a prime ideal; but this will be of the form P/I for some prime ideal P of R that includes I. □

If one knows any **Galois Theory,** then what we are doing can be seen in this context. A **relation from** a set (or class) A **to** a set B is a subset of $A \times B$. Suppose G is such a relation. If $Y \subseteq B$, then the set
$$\bigcap_{y \in Y} \{x \in A \colon x \mathrel{G} y\}$$
is a **closed** subset of A with respect to G. The closed subsets of B are defined similarly. It is not too hard to establish that there is a one-to-one, inclusion-reversing correspondence between the closed subsets of A and the closed subsets of B. It may be more difficult to give an interesting characterization of the closed subsets. In some cases, the closed subsets of A are closed with respect to a topology on A; but this is not always true.

In the original case, B is a field K, and A is its group $\mathrm{Aut}(K)$ of automorphisms, and
$$G = \{(\sigma, x) \in \mathrm{Aut}(K) \times K \colon x^\sigma = x\}.$$
We are interested in three other cases:

- B is a ring R, and A is its spectrum $\mathrm{Spec}(R)$, and
$$G = \{(\mathfrak{p}, x) \in \mathrm{Spec}(R) \times R \colon x \notin \mathfrak{p}\}.$$
- B is a *Boolean algebra* (defined below), A is the set $\mathrm{S}(B)$ of its ultrafilters, and
$$G = \{(\mathscr{U}, x) \in \mathrm{S}(B) \times B \colon x \in \mathscr{U}\}.$$
- For some logical signature $\mathscr{S}$, B is the set $\mathrm{Sn}(\mathscr{S})$ of sentences of $\mathscr{S}$, and A is the *class* $\mathrm{Mod}(\mathscr{S})$ of structures having signature $\mathscr{S}$, and G is the relation of truth:
$$G = \{(\mathfrak{A}, \sigma) \in \mathrm{Mod}(\mathscr{S}) \times \mathrm{Sn}(\mathscr{S}) \colon \mathfrak{A} \models \sigma\}.$$
(See tomorrow's lecture.)

For any set Ω, the structure

$$(\mathscr{P}(\Omega), \varnothing, \Omega, \triangle, \cap\}$$

is a Boolean ring; but the structure

$$(\mathscr{P}(\Omega), \varnothing, \Omega, {}^{\mathrm{c}}, \cup, \cap\}$$

is a **Boolean algebra.** If $(R, 0, 1, +, \times)$ is an arbitrary Boolean ring, then $(R, 0, 1, \neg, \vee, \wedge)$ is a Boolean algebra, where

$$\neg x = 1 + x,$$
$$x \vee y = x + y + xy,$$
$$x \wedge y = xy,$$

and also

$$x + y = (x \wedge \neg y) \vee (y \wedge \neg x), xy = x \wedge y.$$

We are going to prove (on page 57):

Theorem 29 (Stone Representation Theorem)**.** *Every Boolean algebra* $\mathfrak{A}$ *embeds in an algebra* $\mathscr{P}(\Omega)$*. Here* Ω *can be taken as the set* $\mathrm{S}(\mathfrak{A})$ *of ultrafilters of* $\mathfrak{A}$*.*

This is analogous to the simpler:

Theorem 30 (Cayley's Theorem)**.** *Every group* G *embeds in a symmetry group*

$$(\mathrm{Sym}(\Omega), \mathrm{id}_\Omega, {}^{-1}, \circ),$$

where $\mathrm{Sym}(\Omega)$ *is the set of bijections from* Ω *to* Ω*. Here* Ω *can be taken as* G *itself, and the embedding of* G *in* $\mathrm{Sym}(G)$ *is*

$$g \mapsto \lambda_g,$$

where $\lambda_g(x) = gx$*.*

The original groups are a symmetry groups. Such groups are seen to satisfy certain axioms, and it turns out, as in Cayley's Theorem, that all structures that satisfy these axioms are isomorphic to symmetry groups. The same will be seen to be true for Boolean algebras.

2.5 Saturday

Given a first-order signature $\mathscr{S}$, we let

$$\begin{aligned}\mathrm{Mod}(\mathscr{S}) &= \{\text{structures of } \mathscr{S}\},\\ \mathrm{Sn}(\mathscr{S}) &= \{\text{sentences of } \mathscr{S}\}.\end{aligned}$$

The relation $\vDash$ of **truth** is from $\mathrm{Mod}(\mathscr{S})$ to $\mathrm{Sn}(\mathscr{S})$. If $\sigma, \tau \in \mathrm{Sn}(\mathscr{S})$, and for all $\mathfrak{A}$ in $\mathrm{Mod}(\mathscr{S})$,

$$\mathfrak{A} \vDash \sigma \iff \mathfrak{A} \vDash \tau,$$

then we say σ and τ are **logically equivalent,** and we write

$$\sigma \sim \tau.$$

Theorem 31. $\mathrm{Sn}(\mathscr{S})/{\sim}$ *is a Boolean algebra.*

The Boolean algebra of the theorem is called a **Lindenbaum algebra** after a student of Tarski murdered by the Nazis; it can be denoted by

$$\mathrm{Lin}_0(\mathscr{S}).$$

The subscript 0 indicates that there are no free variables in the formulas whose logical equivalence classes compose the algebra. (Later, on page 59, we shall have reason to consider the Lindenbaum algebra of formulas in one free variable.)

If $\mathfrak{A}, \mathfrak{B} \in \mathrm{Mod}(\mathscr{S})$, and for all σ in $\mathrm{Sn}(\mathscr{S})$,

$$\mathfrak{A} \vDash \sigma \iff \mathfrak{B} \vDash \sigma,$$

then we say $\mathfrak{A}$ and $\mathfrak{B}$ are **elementarily equivalent,** and we write

$$\mathfrak{A} \equiv \mathfrak{B}.$$

By defining

$$\mathrm{Th}(\mathfrak{A}) = \{\sigma \in \mathrm{Sn}(\mathscr{S}) \colon \mathfrak{A} \vDash \sigma\}$$

as in (20) on page 46, we have

$$\mathfrak{A} \equiv \mathfrak{B} \iff \mathrm{Th}(\mathfrak{A}) = \mathrm{Th}(\mathfrak{B}).$$

Theorem 32. *For all* $\mathfrak{A}$ *in* $\mathrm{Mod}(\mathscr{S})$, *the quotient*

$$\mathrm{Th}(\mathfrak{A})/{\sim}$$

is an ultrafilter of $\mathrm{Lin}_0(\mathscr{S})$.

Strictly, $\mathrm{Th}(\mathfrak{A})/{\sim}$ here should be understood as

$$\{\sigma/{\sim} \in \mathrm{Lin}_0(\mathscr{S}) \colon \mathfrak{A} \vDash \sigma\}.$$

If now $\mathfrak{A}$ is an arbitrary Boolean algebra, we let

$$\mathrm{S}(\mathfrak{A}) = \{\text{ultrafilters of } \mathfrak{A}\};$$

this will be the **Stone space** of $\mathfrak{A}$. In case $\mathfrak{A} = \mathrm{Lin}_0(\mathscr{S})$, the situation is as in Figure 4. In general, if $U \in \mathrm{S}(\mathfrak{A})$, this means

$$x \in U \;\;\&\;\; y \in U \iff x \wedge y \in U,$$
$$\neg x \in U \iff x \notin U.$$

Hence

$$x \in U \;\;\text{OR}\;\; y \in U \iff x \vee y \in U,$$
$$1 \in U, \qquad 0 \notin U.$$

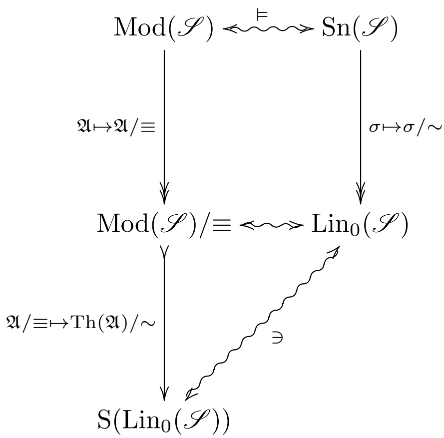

Figure 4. Stone space of Lindenbaum algebra

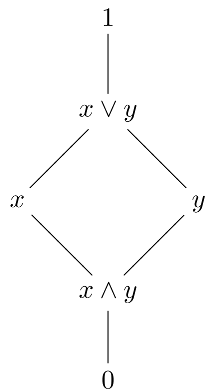

Figure 5. Ordering of a Boolean algebra

The Boolean algebra $\mathfrak{A}$ is (partially) ordered by

$$\begin{aligned} x \leqslant y &\iff x \wedge y = x \\ &\iff x \vee y = y; \end{aligned}$$

see Figure 5. If $a \in A$, let

$$[a] = \{U \in \mathrm{S}(\mathfrak{A})\colon a \in U\}.$$

Proof of the Stone Representation Theorem. In the Boolean algebra $\mathfrak{A}$, we now have

$$\begin{gathered} [x] \cap [y] = [x \wedge y], \\ [\neg x] = [x]^{\mathrm{c}}, \\ [x] \cup [y] = [x \vee y], \\ [1] = \mathrm{S}(\mathfrak{A}), \qquad [0] = \varnothing. \end{gathered}$$

Thus the map $x \mapsto [x]$ from A to $\mathscr{P}(\mathrm{S}(\mathfrak{A}))$ is a homomorphism of Boolean algebras. Moreover, it is an embedding, since if $a \neq b$, then we may assume $a \wedge b \neq a$, so

$$a \wedge \neg b = a \wedge (1 + b) = a + (a \wedge b) \neq 0,$$

since $a + a = 0$. Therefore $a \wedge \neg b$ generates a proper filter, namely

$$\{x \vee (a \wedge \neg b)\colon x \in A\},$$

which by the Prime Ideal Theorem (page 28) is included in an element U of $\mathrm{S}(\mathfrak{A})$. Then

$$U \in [a] \smallsetminus [b],$$

so $[a] \neq [b]$. □

Suppose $(A, \mathscr{O})$ is a topological space as on page 46. If $\mathscr{B} \subseteq \mathscr{O}$, and every element of $\mathscr{O}$ is $\bigcup \mathscr{X}$ for some subset $\mathscr{X}$ of $\mathscr{B}$, then $\mathscr{B}$ is a **base** or **basis** of $\mathscr{O}$. For example,

- $\{(a,b)\colon a,b\in\mathbb{R}\}$ is a base for the topology of $\mathbb{R}$.
- $\{[a]\colon a\in R\}$ is a base for the topology of $\operatorname{Spec}(R)$ for any ring R. Hence:
- $\{[a]\colon a\in A\}$ is a base for the topology of $\mathrm{S}(\mathfrak{A})$ for any Boolean algebra $\mathfrak{A}$, since $\mathrm{S}(\mathfrak{A})=\{P^{\mathrm{c}}\colon P\in\operatorname{Spec}(\mathfrak{A})\}$. The subsets $[a]$ of $\mathrm{S}(\mathfrak{A})$ are both open and closed: they are **clopen.**

A topological space $(A,\mathscr{O})$ with base $\mathscr{B}$ is compact if and only if, for all subsets $\mathscr{X}$ of $\mathscr{B}$ such that $\bigcup\mathscr{X}=A$, there is a finite subset $\mathscr{X}_0$ of $\mathscr{X}$ such that $\bigcup\mathscr{X}_0=A$. Taking the contrapositive and taking complements, we have:

Theorem 33. *A topological space $(A,\mathscr{O})$ with base $\mathscr{B}$ is compact if and only if, for all subsets $\mathscr{X}$ of $\mathscr{B}$ such that, for all finite subsets $\mathscr{X}_0$ of $\mathscr{X}$,*

$$\bigcap\{X^{\mathrm{c}}\colon X\in\mathscr{X}_0\}\neq\varnothing,$$

it follows that

$$\bigcap\{X^{\mathrm{c}}\colon X\in\mathscr{X}\}\neq\varnothing.$$

The hypothesis of this condition is that $\{X^{\mathrm{c}}\colon X\in\mathscr{X}\}$ has the **finite intersection property** or **FIP.**

For a Boolean algebra $\mathfrak{A}$, the theorem is that $\mathrm{S}(\mathfrak{A})$ is compact if and only if, for all subsets X of A such that every finite subset of A is included in an ultrafilter, X itself is included in an ultrafilter. This condition is immediately satisfied.

Let C be the subspace

$$\{\operatorname{Th}(\mathfrak{A})/{\sim}\colon\mathfrak{A}\in\operatorname{Mod}(\mathscr{S})\}$$

of $\mathrm{S}(\operatorname{Lin}_0(\mathscr{S}))$. Then C is compact if and only if, for every subset Γ of $\operatorname{Sn}(\mathscr{S})$, if every finite subset of Γ has a model, then Γ itself has a model. Thus the compactness of C is equivalent to the Compactness Theorem (page 43).

The subspace C is **dense** in $\mathrm{S}(\mathrm{Lin}_0(\mathscr{S}))$, because if $[\sigma/{\sim}] \neq \varnothing$, then

$$\sigma \not\sim \exists x\ x \neq x,$$

so σ has a model $\mathfrak{A}$, and then $\mathrm{Th}(\mathfrak{A})/{\sim} \in [\sigma/{\sim}]$.

Thus every element of $\mathrm{S}(\mathrm{Lin}_0(\mathscr{S})) \smallsetminus C$ is a limit point of C. Therefore C is closed if and only if $C = \mathrm{S}(\mathrm{Lin}_0(\mathscr{S}))$. In particular, if C is closed, then it is compact. Conversely, since $\mathrm{S}(\mathrm{Lin}_0(\mathscr{S}))$ is a *Hausdorff space,* its every compact subset is closed.

We shall not pursue this further, but a topological space is a **Hausdorff space** if for any two distinct points of the space, there are disjoint open sets containing the points respectively. Easily a Stone space $\mathrm{S}(\mathfrak{A})$ is Hausforff: if U_0 and U_1 are distinct elements, then we may assume there is a in $U_0 \smallsetminus U_1$, and so $U_0 \in [a]$ and $U_1 \in [\neg a]$.

Now we prove the theorem stated originally on page 43; we implicitly proved a version of it in proving Łoś's Theorem (page 30).

Theorem 34 (Compactness)**.** *If* Γ *is a subset of* $\mathrm{Sn}(\mathscr{S})$ *whose every finite subset has a model, then* Γ *itself has a model.*

Proof. We can assume $\mathscr{S}$ has been expanded so that, for every singulary formula $\varphi(x)$, there is a new constant symbol

$$c_\varphi.$$

Now let

$$\Gamma^* = \Gamma \cup \{\exists x\ \varphi(x) \Rightarrow \varphi(c_\varphi) \colon \varphi/{\sim} \in \mathrm{Lin}_1(\mathscr{S})\}.$$

Then every finite subset of Γ^* has a model. By the Prime Ideal Theorem (page 28), there is a subset T of $\mathrm{Sn}(\mathscr{S})$ such that $\Gamma \subseteq T$ and $T/{\sim}$ is an ultrafilter of $\mathrm{Lin}_0(\mathscr{S})$. Then T has a model $\mathfrak{A}$ such that

$$A = \{c_\varphi \colon \varphi\}/{\approx},$$

where

$$c_\varphi \approx c_\psi \iff (c_\varphi = c_\psi) \in T.$$

For example, if F is a singulary operation symbol in $\mathscr{S}$, then

$$F^{\mathfrak{A}}(c_\varphi) = c_\psi \iff (Fc_\varphi = c_\psi) \in T.$$

There are many details to check, but this is the idea. □

We obtained ${}^*\mathbb{R}$ as an ultrapower $\mathbb{R}^\omega/\mathscr{U}$; but Robinson [3, p. 55] can be understood to obtain it by the Compactness Theorem as a model of

$$\mathrm{Th}(\mathbb{R}) \cup \{c > n \colon n \in \mathbb{N}\}.$$

(Here the signature of $\mathbb{R}$ has everything we might want.) However, Robinson [3, p. 13] proves the Compactness Theorem by means of ultraproducts and Łoś's Theorem (page 30).

2.6 Sunday

We have shown the equivalence in Zermelo–Fraenkel set theory of:

- the Compactness Theorem (page 59),
- The Prime Ideal Theorem (page 28), and
- The Boolean $\left\{\begin{array}{c}\text{Prime}\\ \text{Maximal}\end{array}\right\}$ Ideal Theorem.

We are going to show the equivalence of:

- The Prime Ideal Theorem, together with Łoś's Theorem, and
- The Axiom of Choice.*

As we noted (page 30), Compactness can be understood as a weak form of Łoś's Theorem (with the Prime Ideal Theorem). Indeed, suppose $\Gamma \subseteq \mathrm{Sn}(\mathscr{S})$, and every Δ in $\mathscr{P}_\omega(\Gamma)$ has a model $\mathfrak{A}_\Delta$. If $\sigma \in \Gamma$, let

$$[\sigma] = \{\Delta \in \mathscr{P}_\omega(\Gamma) \colon \sigma \in \Delta\}.$$

Then for all Δ in $\mathscr{P}_\omega(\Gamma)$,

$$\Delta \in \bigcap_{\sigma \in \Delta} [\sigma].$$

Thus $\{[\sigma] \colon \sigma \in \Gamma\}$ generates a proper filter on $\mathscr{P}_\omega(\Gamma)$. By the Prime Ideal Theorem, this filter is included in an ultrafilter $\mathscr{U}$. Let

$$A = \prod_{\Delta \in \mathscr{P}_\omega(\Gamma)} A_\Delta,$$

and if a is an element $(a_\Delta \colon \Delta \in \mathscr{P}_\omega(\Gamma))$ of A, let a be interpreted in $\mathfrak{A}_\Delta$ as a_Δ. For all σ in $\mathscr{S}(A)$, let

$$\|\sigma\| = \{\Delta \in \mathscr{P}_\omega(\Gamma) \colon \mathfrak{A}_\Delta \vDash \sigma\}.$$

Finally, let

$$T = \{\sigma \in \mathrm{Sn}(\mathscr{S}(A)) \colon \|\sigma\| \in \mathscr{U}\}.$$

Then $\Gamma \subseteq T$, since if $\sigma \in \Gamma$, then $[\sigma] \subseteq \|\sigma\|$, and $[\sigma] \in \mathscr{U}$, so $\|\sigma\| \in \mathscr{U}$. Also, if $\{\sigma_k \colon k < n\} \subseteq T$, then

$$\left\| \bigwedge_{k<n} \sigma_k \right\| = \bigcap_{k<n} \|\sigma_k\|,$$

*We could add to the list examples like the *Tychonoff Theorem*: the product of compact spaces is compact. Restricted to Hausdorff spaces, the theorem is equivalent to the Prime Ideal Theorem. See Rubin and Rubin [4] and their references.

which is in $\mathscr{U}$, so in particular it is nonempty, which means the set $\{\sigma_k \colon k < n\}$ has a model. By the Compactness Theorem, T has a model $\mathfrak{C}$, and then $T = \mathrm{Th}(\mathfrak{C})$.

We may assume $A/\mathscr{U} \subseteq C$. Then $\mathfrak{C}$ has a substructure $\mathfrak{B}$ whose universe B is precisely $A/\mathscr{U}$; for if F is an n-ary operation symbol in $\mathscr{S}$ for some n in $\mathbb{N}$, then

$$F^{\mathfrak{C}}(a^0/\mathscr{U}, \dots, a^{n-1}/\mathscr{U}) = (F^{\mathfrak{A}_\Delta}(a^0_\Delta, \dots, a^{n-1}_\Delta) \colon \Delta \in \mathscr{P}_\omega(\Gamma))/\mathscr{U},$$

so B is closed under the operations of $\mathfrak{C}$. Thus

$$\mathfrak{B} \subseteq \mathfrak{C}.$$

Łoś's Theorem is then

$$\mathfrak{B} \preccurlyeq \mathfrak{C},$$

that is, for all σ in $\mathrm{Sn}(\mathscr{S}(B))$ or $\mathrm{Sn}(\mathscr{S}(A))$,

$$\mathfrak{B} \vDash \sigma \iff \mathfrak{C} \vDash \sigma.$$

To prove this, by the **Tarski–Vaught Test,** it is enough to show that, for all formulas $\varphi(x)$ of $\mathscr{S}(A)$, if

$$\mathfrak{C} \vDash \exists x\ \varphi(x),$$

then for some b in B, $\mathfrak{C} \vDash \varphi(b)$. So we show this. If $\mathfrak{A}_\Delta \vDash \exists x\ \varphi(x)$, then for some b_Δ in A_Δ, $\mathfrak{A}_\Delta \vDash \varphi(b_\Delta)$. Otherwise let b_Δ be arbitrary. Now let

$$b = (b_\Delta \colon \Delta \in \mathscr{P}_\omega(\Gamma))/\mathscr{U}.$$

So Łoś's Theorem holds. Note however that we used the Axiom of Choice to obtain b.

Conversely, the following is shown by Howard [2]:

Theorem 35. *The Axiom of Choice follows from Łoś's Theorem and the Prime Ideal Theorem.*

Proof. Suppose $\mathscr{A}$ is a set of nonempty sets with no **choice function,** that is, no function f from $\mathscr{A}$ to $\bigcup\mathscr{A}$ such that, for all X in $\mathscr{A}$, $f(X) \in X$. Let

$$\Omega = \bigcup\mathscr{A} \cup \mathscr{A},$$
$$R = \Big\{(x, Y) \in \bigcup\mathscr{A} \times \mathscr{A} : x \in Y\Big\} \cup \Big\{(y, y) \colon y \in \bigcup\mathscr{A}\Big\}.$$

Then

$$(\Omega, R) \vDash \forall y\ \exists x\ x\ R\ y.$$

Now, the subset

$$\{\mathscr{X} \subseteq \mathscr{A} : \mathscr{X} \text{ has no choice function}\}$$

of $\mathscr{P}(\mathscr{A})$ is a proper ideal on $\mathscr{A}$, so by the Prime Ideal Theorem, its complement includes an ultrafilter $\mathscr{U}$ on $\mathscr{A}$. This embeds in an ultrafilter $\mathscr{U}^*$ on Ω. by Łoś's Theorem,

$$(\Omega, R)^{\Omega}/\mathscr{U}^* \vDash \forall y\ \exists x\ x\ R\ y.$$

Replacing y with $(i \colon i \in \Omega)/\mathscr{U}^*$, we get $(a_i \colon i \in \Omega)/\mathscr{U}^*$ such that $\mathscr{U}$ contains $\{i \in \Omega \colon a_i\ R\ i\}$. Let

$$\begin{aligned} C &= \{X \in A \colon a_X\ R\ X\} \\ &= \{X \in A \colon a_X \in X\}. \end{aligned}$$

Then $X \mapsto a_X$ is a choice function on C. But $C \in \mathscr{U}$, so by definition of $\mathscr{U}$, C cannot have a choice function. This contradiction implies that $\mathscr{A}$ must have had a choice function. □

References

[1] Wilfrid Hodges. *Model theory*, volume 42 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1993.

[2] Paul E. Howard. Łoś' theorem and the Boolean prime ideal theorem imply the axiom of choice. *Proc. Amer. Math. Soc.*, 49:426–428, 1975.

[3] Abraham Robinson. *Non-standard analysis.* Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 1996. Reprint of the second (1974) edition, With a foreword by Wilhelmus A. J. Luxemburg.

[4] Herman Rubin and Jean E. Rubin. *Equivalents of the axiom of choice. II*, volume 116 of *Studies in Logic and the Foundations of Mathematics.* North-Holland Publishing Co., Amsterdam, 1985.