ULTRAFILTERS AND HOW TO USE THEM

BURAK KAYA

Abstract. These are the lecture notes of a one-week course I taught at the Nesin Mathematics Village in Şirince, İzmir, Turkey during Summer 2019. The aim of the course was to introduce ultrafilters, prove some fundamental facts and cover various constructions and results in algebra, model theory, topology and social choice theory that use ultrafilters.

Contents

The main purpose of this course is to introduce ultrafilters and their basic applications in various fields of mathematics. This course is intended for third and fourth year undergraduate students and beginning graduate students. Although the necessary background from algebra, topology, set theory and model theory was briefly reviewed in class, it is not included in these lectures notes. We refer the reader to [Hun80], [Mun00], [Jec03], [SS17] and [Hod93] for the basic definitions and facts necessary for these lecture notes.

1. Big sets and small sets

In this section, we shall introduce ultrafilters on sets, prove some basic facts about them and show their existence.

Let $\mathbf{X}$ be a non-empty set. A *filter* on the set $\mathbf{X}$ is a collection of sets $\mathcal{F} \subseteq \mathcal{P}(\mathbf{X})$ such that

- $\emptyset \notin \mathcal{F}$ and $\mathbf{X} \in \mathcal{F}$,
- if $A \in \mathcal{F}$ and $A \subseteq B \subseteq \mathbf{X}$, then $B \in \mathcal{F}$, and
- if $A \in \mathcal{F}$ and $B \in \mathcal{F}$, then $A \cap B \in \mathcal{F}$.

The most basic example of a filter is the *trivial filter* $\mathcal{F} = \{\mathbf{X}\}$. For any non-empty subset $\hat{X} \subseteq \mathbf{X}$, the collection $\mathcal{F} = \{A \subseteq \mathbf{X} : \hat{X} \subseteq A\}$ is also a filter which is called the *principal filter generated by* X.

We shall next make some observations about filters which immediately follow from the definition. The intersection of a collection of filters on $\mathbf{X}$ is also a filter on $\mathbf{X}$ and the union of a collection of filters which are pairwise $\subseteq$-comparable is a filter on $\mathbf{X}$. Filters, being closed under intersection of two sets, have the finite intersection property[1]. It turns out that collections with finite intersection property can be extended to filters as well.

Proposition 1. *Let $\mathbf{X}$ be a non-empty set and $\mathcal{C} \subseteq \mathcal{P}(\mathbf{X})$ be a non-empty set with the finite intersection property. Then there exists a filter $\mathcal{F} \supseteq \mathcal{C}$ on $\mathbf{X}$.*

Proof. It is an exercise to the reader to check that the collection

$$\mathcal{F} = \{A : \exists n \in \mathbb{N}^+ \ \exists C_1, C_2, \ldots, C_n \in \mathcal{C} \ \ C_1 \cap C_2 \cap \ldots C_n \subseteq A\}$$

is a filter on $\mathbf{X}$ containing $\mathcal{C}$ as a subset. □

A filter $\mathcal{U}$ on the set $\mathbf{X}$ is said to be an *ultrafilter*[2] if it satisfies that

- $A \in \mathcal{U}$ or $\mathbf{X} - A \in \mathcal{U}$ for every $A \in \mathcal{P}(\mathbf{X})$.

An ultrafilter on a set can be considered as a tool to split the power set of that set into two: "Big sets" and "small sets". Big sets are those that are in the ultrafilter[3]. Indeed, in order for the reader to understand this motivation, we expect the reader to check that an ultrafilter $\mathcal{U}$ on a set $\mathbf{X}$ canonically induces a finitely additive measure $\mu_{\mathcal{U}} : \mathcal{P}(\mathbf{X}) \to \{0, 1\}$ and vice versa.

Let $\mathbf{X}$ be a non-empty set and $x \in \mathbf{X}$. Then, since $x \in A$ or $x \notin A$ for every $A \subseteq \mathbf{X}$, the principal filter generated by $\{x\}$ is an ultrafilter and is called the *principal ultrafilter generated by x*. It is also clear that any ultrafilter which is also principal has a generating set that is a singleton and has to be of this form. Before we proceed, let us prove a basic but important fact.

Proposition 2. *Any ultrafilter on a non-empty finite set is principal.*

Proof. Let $\mathbf{X}$ be a non-empty finite set and $\mathcal{U}$ be an ultrafilter. Consider the set $A = \bigcap_{U \in \mathcal{U}} U$. Since $\mathbf{X}$ is finite, so is the ultrafilter $\mathcal{U}$. It then follows from the properties of a filter that $A \in \mathcal{U}$ and hence $A \neq \emptyset$. (Up to now, we have only used that $\mathcal{U}$ is a filter and indeed shown that any filter on a non-empty set is principal whose generating set is the intersection of all sets in the filter.)

Let $a \in A$. Being an ultrafilter, we know that either $\{a\} \in \mathcal{U}$ or $\mathbf{X} - \{a\} \in \mathcal{U}$. If it were the case that $\mathbf{X} - \{a\} \in \mathcal{U}$, then we would have $a \in A \subseteq \mathbf{X} - \{a\}$ which is a contradiction. Thus, $\{a\} \in \mathcal{U}$. We claim that $\mathcal{U}$ is generated by a. If it were not, then there would be $B \in \mathcal{U}$ such that $\{a\} \nsubseteq B$ in which case $\{a\} \cap B = \emptyset \in \mathcal{U}$ which is a contradiction. Therefore, $\mathcal{U}$ is generated by a and is principal. □

2. Fantastic beasts and where to find them

What about ultrafilters on infinite sets? Are there any non-principal ultrafilters on infinite sets? As we shall see soon, the answer to this question is affirmative. However, we will not be able to "explicitly see" any of these non-principal ultrafilters. In order to prove the existence of non-principal ultrafilters, we need to introduce a very special filter.

[1] A collection $\mathcal{C}$ of sets is said to have the *finite intersection property* if for every $n \in \mathbb{N}^+$ and $C_1, C_2, \ldots, C_n$ we have that $C_1 \cap C_2 \cap \cdots \cap C_n \neq \emptyset$.

[2] Our definition corresponds to the special case of that definition where one considers the Boolean algebra $(\mathcal{P}(\mathbf{X}), \cap, \cup, \emptyset, \mathbf{X})$

[3] The notion of a filter has a dual notion, namely, the notion of an ideal. Ideals contain the empty set, are closed $\subseteq$-downwards and with respect to finite unions. Small sets are those that are in the corresponding ideal.

Let $\mathbf{X}$ be an infinite set. The collection $\mathcal{F} = \{A \subseteq \mathbf{X} : \mathbf{X} - A \text{ is finite}\}$ consisting of cofinite subsets of $\mathbf{X}$ is a filter called the *Fréchet filter* on $\mathbf{X}$. The Fréchet filter is clearly non-principal. As will be shown in the next proposition, whether an ultrafilter is non-principal is determined by whether it contains the Fréchet filter as a subset.

Proposition 3. *Let $\mathcal{U}$ be an ultrafilter on an infinite set $\mathbf{X}$. Then $\mathcal{U}$ is non-principal if and only if it contains the Fréchet filter.*

Proof. Assume that $\mathcal{U}$ is principal, say, $\mathcal{U} = \{A \subseteq \mathbf{X} : a \in A\}$. Then, since $\{a\} \in \mathcal{U}$, we have that $\mathbf{X} - \{a\} \notin \mathcal{U}$. On the other hand, $\mathbf{X} - \{a\}$ is cofinite. Thus $\mathcal{U}$ does not contain the Fréchet filter.

Assume that $\mathcal{U}$ does not contain the Fréchet filter. Then there exists a cofinite $A \subseteq \mathbf{X}$ such that $A \notin \mathcal{U}$ and hence $\mathbf{X} - A \in \mathcal{U}$. Enumerate the set $\mathbf{X} - A$, say, $\mathbf{X} - A = \{x_1, x_2, \ldots, x_n\}$. If it were that $\mathbf{X} - \{x_i\} \in \mathcal{U}$ for every $1 \leq i \leq n$, then we would have

$$\bigcap_{i=1}^{n} \mathbf{X} - \{x_i\} = \mathbf{X} - \{x_1, x_2, \ldots, x_n\} \in \mathcal{U}$$

which is a contradiction, as the intersection of this set with $\mathbf{X} - A$ is empty. Therefore, there exists $1 \leq i \leq n$ such that $X - \{x_i\} \notin U$ and hence $\{x_i\} \in \mathcal{U}$. Then, as in the proof of Proposition 2, we necessarily have $\{A \subseteq \mathbf{X} : x_i \in A\} = \mathcal{U}$. $\square$

Thus, in order to show the existence of non-principal ultrafilters, we just need to find one that contains the Fréchet filter. How are we supposed to do that? Let us take the Fréchet filter on an infinite set $\mathbf{X}$. The Fréchet filter is clearly not an ultrafilter. We can fix this by extending the Fréchet filter as transfinitely going through all elements of $\mathcal{P}(\mathbf{X})$ and appropriately adding each element or its complement. Although such processes are usually done by transfinite recursion, we are not expecting every person taking this course to be familiar with transfinite recursion and hence we will handle this procedure using Zorn's lemma.

Theorem 1 (Tarski)**.** *Let $\mathbf{X}$ be a non-empty set and $\hat{\mathcal{F}} \subseteq \mathcal{P}(\mathbf{X})$ be a filter. Then there exists an ultrafilter $\mathcal{U}$ on $\mathbf{X}$ such that $\hat{\mathcal{F}} \subseteq \mathcal{U}$.*

Proof. Let $\mathbb{P} = \{\mathcal{F} : \mathcal{F} \text{ is a filter and } \hat{\mathcal{F}} \subseteq \mathcal{F}\}$ and consider the partially ordered set $(\mathbb{P}, \subseteq)$. Given a chain $C \subseteq \mathbb{P}$, as pointed out before, the set $\bigcup C$ is a filter on $\mathbf{X}$ containing $\hat{\mathcal{F}}$ and is an upper bound for C. Therefore, $(\mathbb{P}, \subseteq)$ satisfies the hypotheses of Zorn's lemma and has a maximal element, say, $\mathcal{U}$ is a maximal element of $(\mathbb{P}, \subseteq)$.

We claim that $\mathcal{U}$ is indeed an ultrafilter. Assume towards a contradiction that it is not. Then there exists $A \subseteq \mathbf{X}$ such that $A \notin \mathcal{U}$ and $\mathbf{X} - A \notin \mathcal{U}$. Consider the collection $\mathcal{C} = \mathcal{U} \cup \{A\}$. We claim that $\mathcal{C}$ has the finite intersection property. Let $X_1, X_2, \ldots, X_n \in \mathcal{C}$. We split into cases.

- **Case 1:** Assume that $X_i \in \mathcal{U}$ for every $1 \leq i \leq n$. Then $X_1 \cap X_2 \cap \cdots \cap X_n \in \mathcal{U} \subseteq \mathcal{C}$ since $\mathcal{U}$ has the finite intersection property.
- **Case 2:** Assume that $X_i \notin \mathcal{U}$ for some $1 \leq i \leq n$. By replacing these sets without changing their intersection, we can assume without loss of generality that $X_1 = A$ and $X_2, \ldots, X_n \in \mathcal{U}$. As $\mathcal{U}$ has the finite intersection property, we have that $X_2 \cap \cdots \cap X_n \in U$. It follows that any superset of $X_2 \cap \cdots \cap X_n$ is in $\mathcal{U}$. On the other hand, $\mathbf{X} - A \notin \mathcal{U}$ and hence $X_2 \cap \cdots \cap X_n \nsubseteq \mathbf{X} - A$, that is, $A \cap X_2 \cap \cdots \cap X_n \neq \emptyset$.

Therefore $\mathcal{C}$ has the finite intersection property and can be extended to a filter $\mathcal{F} \supseteq \mathcal{C} \supseteq \mathcal{U}$ which contradicts the maximality of $\mathcal{U}$.[4] □

We are now ready to construct a non-principal ultrafilter. Let $\mathbf{X}$ be a non-empty infinite set. By Theorem 1, there exists an ultrafilter $\mathcal{U}$ on $\mathbf{X}$ containing the Fréchet filter on $\mathbf{X}$. It follows from Proposition 3 that $\mathcal{U}$ is non-principal.

Objects constructed via Zorn's lemma usually tend to be "pathological". One can ask whether or not it is possible to construct a non-principal ultrafilter without the use of the axiom of choice or its equivalents. The answer to this question is negative. While this fact is too advanced to be covered in this course, we can briefly describe the reasoning behind it:

We can identify $\mathcal{P}(\mathbb{N})$ with the Cantor space $2^{\mathbb{N}}$ through characteristic functions. There is a Borel probability measure on $2^{\mathbb{N}}$ which is induced from the product measure where each component $2 = \{0, 1\}$ has the coin-flipping probability measure. An application of the Kolmogorov 0-1 law to the probability space $2^{\mathbb{N}}$ shows that any *measurable* subset of $2^{\mathbb{N}}$ that is invariant under finite changes[5] has to be of measure 0 or 1. Any non-principal ultrafilter on $\mathbb{N}$ contains the Fréchet filter and closed under finite intersections and consequently, is invariant under finite changes. If a non-principal ultrafilter on $\mathbb{N}$ were measurable as a subset of $2^{\mathbb{N}}$, it would have to have measure 0 or 1. However, the bit-flipping transformation, which takes an ultrafilter to its complement, is a measure-preserving transformation of $2^{\mathbb{N}}$ and hence an ultrafilter has to have measure $1/2$ as a subset of $2^{\mathbb{N}}$. Therefore, any non-principal ultrafilter is necessarily non-measurable as a subset of $2^{\mathbb{N}}$. Assuming the relative consistency of an inaccessible cardinal with ZFC, Robert Solovay proved in his celebrated work [Sol70] that the existence of a non-measurable subset of $2^{\mathbb{N}}$ cannot be proven using ZF+DC. Hence, in order to show the existence of a non-principal ultrafilter on $\mathbb{N}$, one should use "more choice" than DC.

Having shown that non-principal ultrafilters on an infinite set exist, the next obvious question is: *How many* such ultrafilters are there? For any set $\mathbf{X}$ with cardinality κ, the cardinality of the set of ultrafilters on $\mathbf{X}$ is trivially less than or equal to the cardinality of $\mathcal{P}(\mathcal{P}(\mathbf{X}))$ which is $2^{2^{\kappa}}$. It turns out that, for an infinite set $\mathbf{X}$, this upper bound is always achieved and there are exactly $2^{2^{\kappa}}$ non-principal ultrafilters on $\mathbf{X}$.

In order to prove this result, we shall need a result of Hausdorff on independent families which is important on its own. Let us recall the notion of an independent family. Given a set $\mathbf{X}$, a collection $\mathcal{C} \subseteq \mathcal{P}(\mathbf{X})$ is called an *independent family on* $\mathbf{X}$ if

$$X_1 \cap X_2 \cap \cdots \cap X_n \cap (\mathbf{X} - Y_1) \cap (\mathbf{X} - Y_2) \cap \cdots \cap (\mathbf{X} - Y_m) \neq \emptyset$$

for every distinct $X_1, X_2, \ldots, X_n, Y_1, Y_2, \ldots, Y_m \in \mathcal{C}$. The next lemma shows that independent families can get as big as possible.

Lemma 1 (Hausdorff, Fichtenholz-Kantorovich)**.** *Let* $\mathbf{X}$ *be an infinite set with cardinality* κ*. Then there exists an independent family* $\mathcal{C}$ *on* $\mathbf{X}$ *of size* 2^{κ}*.*

[4]The reader should be aware of that the argument in the second part of this proof in fact shows that $\subseteq$-maximal filters are ultrafilters. The converse proposition also holds. Thus, a filter is an ultrafilter if and only if it is $\subseteq$-maximal.

[5]A subset $\mathcal{C} \subseteq \mathcal{P}(\mathbb{N})$ is said to be invariant under finite changes if $C \in \mathcal{C}$, $A \subseteq \mathbb{N}$ and $|C \Delta A| < \infty$ imply that $A \in \mathcal{C}$.

Proof. Consider the set

$$\hat{\mathbf{X}} = \{(F, \mathcal{F}) : F \in \mathcal{P}_{fin}(\mathbf{X}) \text{ and } \mathcal{F} \in \mathcal{P}_{fin}(\mathcal{P}_{fin}(\mathbf{X}))\}$$

It is easily seen that $|\hat{\mathbf{X}}| = |\mathbf{X}| = \kappa$ and that, using a fixed bijection between these sets, one can obtain an independent family on $\mathbf{X}$ from an independent family on $\mathbf{X}$ of the same size. Thus it suffices to find an independent family of size 2^κ on $\hat{\mathbf{X}}$. For each subset $A \subseteq \mathbf{X}$, set

$$X_A = \{(F, \mathcal{F}) \in \hat{\mathbf{X}} : F \cap A \in \mathcal{F}\}$$

Note that if $A, B \subseteq \mathbf{X}$ and $x \in A - B$, then we have that $(\{x\}, \{\{x\}\}) \in X_A - X_B$. It follows that $X_A \neq X_B$ whenever $A \neq B$. Thus the collection $\hat{\mathcal{C}} = \{X_A : A \in \mathcal{P}(\hat{\mathbf{X}})\}$ has size 2^κ. We claim that $\hat{\mathcal{C}}$ is an independent family on $\hat{\mathbf{X}}$.

Let $X_{A_1}, X_{A_2}, \ldots, X_{A_n}, X_{B_1}, X_{B_2}, \ldots, X_{B_m} \in \hat{\mathcal{C}}$ be distinct. For each $1 \leq i \leq n$ and $1 \leq j \leq m$, choose an element $x_{ij} \in A_i \Delta B_j$. Consider the finite set

$$K = \{x_{ij} : 1 \leq i \leq n \text{ and } 1 \leq j \leq m\}$$

Set $\mathcal{F} = \{K \cap A_i : 1 \leq i \leq n\}$. Then, by the choice of x_{ij}'s, we have that $(K, \mathcal{F}) \in X_{A_i} - X_{B_j}$ for all $1 \leq i \leq n$ and $1 \leq j \leq m$. It follows that

$$X_{A_1} \cap X_{A_2} \cap \cdots \cap X_{A_n} \cap (\hat{\mathbf{X}} - X_{B_1}) \cap (\hat{\mathbf{X}} - X_{B_2}) \cap \cdots \cap (\hat{\mathbf{X}} - X_{B_m}) \neq \emptyset$$

Thus $\hat{\mathcal{C}}$ is an independent family on $\hat{\mathbf{X}}$. □

We are now ready to prove the final result of this section.

Theorem 2 (Pospíšil)**.** *Let $\mathbf{X}$ be an infinite set with cardinality κ. Then there exist 2^{2^κ} ultrafilters on $\mathbf{X}$.*

Proof. By Lemma 1, there exists an independent family $\mathcal{C} \subseteq \mathcal{P}(\mathbf{X})$ of size 2^κ. Let $f : \mathcal{C} \to \{0, 1\}$ be a function and consider the collection

$$\mathcal{B}_f = \{A \in \mathcal{C} : f(A) = 1\} \cup \{\mathbf{X} - A \in \mathcal{C} : f(A) = 0\}$$

Since $\mathcal{C}$ is an independent family, $\mathcal{B}_f$ has the finite intersection property and can be extended to an ultrafilter $\mathcal{U}_f$ by Proposition 1 and Theorem 1. It is easily seen that, if $f, g \in {}^{\mathcal{C}}\{0, 1\}$ are distinct, then there exists $A \in \mathcal{C}$ such that $f(A) \neq g(A)$ in which case A belongs to exactly one of $\mathcal{B}_f$ and $\mathcal{B}_g$ and $\mathbf{X} - A$ belongs to the other one. It follows that $\mathcal{U}_f \neq \mathcal{U}_g$ whenever $f \neq g$ and hence the collection $\{\mathcal{U}_f : f \in {}^{\mathcal{C}}\{0, 1\}\}$ of ultrafilters on $\mathbf{X}$ has cardinality $|{}^{\mathcal{C}}\{0, 1\}| = 2^{2^\kappa}$ □

3. Ultraproducts and Łos's theorem

In this section, we shall learn a very important model-theoretic construction that is frequently used in algebra and model theory, namely, the ultraproduct of a collection of structures. For the rest of this section, fix a language $\mathcal{L}$ whose sets of functions symbols, relation symbols and constant symbols will be shown by $\mathcal{L}_F$, $\mathcal{L}_R$ and $\mathcal{L}_C$ respectively.

Let I be an infinite set, $\mathcal{U}$ be an ultrafilter on I and $\{\mathcal{M}_i : i \in I\}$ be a set of $\mathcal{L}$-structures, say, we have $\mathcal{M}_i = (M_i, \{f_s^{\mathcal{M}_i}\}_{s \in \mathcal{L}_F}, \{R_s^{\mathcal{M}_i}\}_{s \in \mathcal{L}_R}, \{c_s^{\mathcal{M}_i}\}_{s \in \mathcal{L}_C})$. Consider the relation $\sim$ defined on $\prod_{i \in I} M_i$ given by

$$(x_i)_{i \in I} \sim (y_i)_{i \in I} \text{ if and only if } \{i \in I : x_i = y_i\} \in \mathcal{U}$$

In other words, two sequences in $\prod_{i \in I} M_i$ are related under $\sim$ if and only if they agree on a "large" set. A moment's thought reveals that $\sim$ is an equivalence relation.

We shall now create a structure whose underlying set is the quotient set

$$\mathcal{M} = \prod_{i\in I} M_i / \sim$$

In order to do this, we need to appropriately interpret each symbol in $\mathcal{L}$ on $\mathcal{M}$. For notational simplicity, the equivalence class $[(x_i)_{i\in I}]_\sim$ will be denoted by $\overline{(x_i)_{i\in I}}$. For every function symbol $f \in \mathcal{L}_F$, we define

$$f^{\mathcal{M}}\big(\overline{(x_i)_{i\in I}}, \overline{(y_i)_{i\in I}}, \dots, \overline{(w_i)_{i\in I}}\big) = \overline{\big(f^{\mathcal{M}_i}(x_i, y_i, \dots, w_i)\big)_{i\in I}}$$

We claim that the relation $f^{\mathcal{M}}$ defined above is well-defined and hence is indeed a function.

Lemma 2. *$f^{\mathcal{M}}$ is well-defined.*

Proof. We shall assume throughout the proof that f is a binary function symbol for, otherwise, there would be too many indices around creating ugliness and confusion.

Let $(x_i)_{i\in I}, (x'_i)_{i\in I}, (y_i)_{i\in I}, (y_i)_{i\in I} \in \prod_{i\in I} M_i$ be such that $\overline{(x_i)_{i\in I}} = \overline{(x'_i)_{i\in I}}$ and $\overline{(y_i)_{i\in I}} = \overline{(y'_i)_{i\in I}}$. By definition, this means that $\{i \in I : x_i = x'_i\} \in \mathcal{U}$ and $\{i \in I : y_i = y'_i\} \in \mathcal{U}$. As $\mathcal{U}$ is closed under finite intersections, we have that $\{i \in I : (x_i, y_i) = (x'_i, y'_i)\} \in \mathcal{U}$. As $\mathcal{U}$ is closed upwards, we have that

$$\{i \in I : (x_i, y_i) = (x'_i, y'_i)\} \subseteq \{i \in I : f^{\mathcal{M}_i}(x_i, y_i) = f^{\mathcal{M}_i}(x'_i, y'_i)\} \in \mathcal{U}$$

and hence $f^{\mathcal{M}}(\overline{(x_i)_{i\in I}}) = \overline{\big(f^{\mathcal{M}_i}(x_i)\big)_{i\in I}} = \overline{\big(f^{\mathcal{M}_i}(x'_i)\big)_{i\in I}} = f^{\mathcal{M}}(\overline{(x'_i)_{i\in I}})$. Therefore $f^{\mathcal{M}}$ is well-defined. □

For every relation $R \in \mathcal{L}_F$, we define

$$\big(\overline{(x_i)_{i\in I}}, \overline{(y_i)_{i\in I}}, \dots, \overline{(w_i)_{i\in I}}\big) \in R^{\mathcal{M}} \Leftrightarrow \{i \in I : (x_i, y_i, \dots, w_i) \in R^{\mathcal{M}_i}\} \in \mathcal{U}$$

That is, if the relations $R^{\mathcal{M}_i}$ componentwise hold on a "large" set, then $R^{\mathcal{M}}$ holds between the elements of $\mathcal{M}$ defined by the corresponding sequences. As before, whether or not $R^{\mathcal{M}}$ holds between elements of $\mathcal{M}$ is independent of the choice of representative sequences of these elements. The reader is expected to check this fact. For every constant symbol $c \in \mathcal{L}_C$, we define

$$c^{\mathcal{M}} = \overline{(c^{\mathcal{M}_i})_{i\in I}}$$

The structure

$$\prod_{i\in I} \mathcal{M}_i/\mathcal{U} = \big(\mathcal{M}, \{f_s^{\mathcal{M}}\}_{s\in\mathcal{L}_F}, \{R_s^{\mathcal{M}}\}_{s\in\mathcal{L}_R}, \{c_s^{\mathcal{M}}\}_{s\in\mathcal{L}_C}\big)$$

is called the *ultraproduct* of $\{\mathcal{M}_i\}_{i\in I}$ with respect to $\mathcal{U}$. In the case that $\mathcal{M}_i = \mathcal{N}$ for all $i \in I$ for a fixed $\mathcal{L}$-structure $\mathcal{N}$, the ultraproduct $\prod_{i\in I} \mathcal{N}/\mathcal{U}$ is called the *ultrapower* of $\mathcal{N}$ with respect to $\mathcal{U}$.

Next shall be proven a remarkable theorem of Łos, which states that a first-order $\mathcal{L}$-sentence holds in an ultraproduct if and only if it componentwise holds on a "large" set.

Theorem 3 (Łos's theorem)**.** *Let $\varphi(\alpha_1, \dots, \alpha_n)$ be a first-order $\mathcal{L}$-formula with n free variables and let $\overline{(x_i^1)_{i\in I}}, \dots, \overline{(x_i^n)_{i\in I}} \in \mathcal{M}$. Then*

$$\prod_{i\in I} \mathcal{M}_i/\mathcal{U} \models \varphi(\overline{(x_i^1)_{i\in I}}, \dots, \overline{(x_i^n)_{i\in I}}) \ \Leftrightarrow\ \{i \in I : \mathcal{M}_i \models \varphi(x_i^1, \dots, x_i^n)\} \in \mathcal{U}$$

Proof. We shall prove this by induction on the complexity of the formulas. As the base case, one should first show that the statement of the theorem holds for all atomic formulas and for all $n \in \mathbb{N}$. It is an (pedagogically important) exercise to the reader to prove this fact by induction. We continue the proof assuming that the statement has been proven for all atomic formulas and for all $n \in \mathbb{N}$.

Let $\overline{(x_i^1)_{i\in I}}, \ldots, \overline{(x_i^n)_{i\in I}} \in \mathcal{M}$ and let $\psi(\alpha_1, \ldots, \alpha_n)$ be a formula which is of the form $\neg\varphi(\alpha_1, \ldots, \alpha_n)$. As inductive hypothesis, assume that the statement holds for $\varphi(\alpha_1, \ldots, \alpha_n)$. Then, by definition, we have that

$$\begin{aligned}
\prod_{i\in I} \mathcal{M}_i/\mathcal{U} \models \psi(\overline{(x_i^1)_{i\in I}}, \ldots, \overline{(x_i^n)_{i\in I}}) &\Leftrightarrow \prod_{i\in I} \mathcal{M}_i/\mathcal{U} \models \neg\varphi(\overline{(x_i^1)_{i\in I}}, \ldots, \overline{(x_i^n)_{i\in I}}) \\
&\Leftrightarrow \prod_{i\in I} \mathcal{M}_i/\mathcal{U} \not\models \varphi(\overline{(x_i^1)_{i\in I}}, \ldots, \overline{(x_i^n)_{i\in I}}) \\
&\Leftrightarrow \{i \in I : \mathcal{M}_i \models \varphi(x_i^1, \ldots, x_i^n)\} \notin \mathcal{U} \\
&\Leftrightarrow \{i \in I : \mathcal{M}_i \not\models \varphi(x_i^1, \ldots, x_i^n)\} \in \mathcal{U} \\
&\Leftrightarrow \{i \in I : \mathcal{M}_i \models \neg\varphi(x_i^1, \ldots, x_i^n)\} \in \mathcal{U} \\
&\Leftrightarrow \{i \in I : \mathcal{M}_i \models \psi(x_i^1, \ldots, x_i^n)\} \in \mathcal{U}
\end{aligned}$$

Now, let $\overline{(x_i^1)_{i\in I}}, \ldots, \overline{(x_i^n)_{i\in I}} \in \mathcal{M}$ and let $\psi(\alpha_1, \ldots, \alpha_n)$ be a formula which is of the form $\varphi \wedge \chi$.[6] As inductive hypothesis, assume that the statement holds for both φ and χ. Then, we have that

$$\begin{aligned}
\prod_{i\in I} \mathcal{M}_i/\mathcal{U} \models \psi(\overline{(x_i^1)_{i\in I}}, \ldots, \overline{(x_i^n)_{i\in I}}) &\Leftrightarrow \prod_{i\in I} \mathcal{M}_i/\mathcal{U} \models \varphi \wedge \chi \\
&\Leftrightarrow \prod_{i\in I} \mathcal{M}_i/\mathcal{U} \models \varphi \text{ and } \prod_{i\in I} \mathcal{M}_i/\mathcal{U} \models \chi \\
&\Leftrightarrow \{i \in I : \mathcal{M}_i \models \varphi\} \in \mathcal{U} \\
&\qquad\qquad \text{and} \\
&\qquad \{i \in I : \mathcal{M}_i \models \chi\} \in \mathcal{U} \\
&\Leftrightarrow \{i \in I : \mathcal{M}_i \models \varphi \text{ and } \mathcal{M}_i \models \chi\} \in \mathcal{U} \\
&\Leftrightarrow \{i \in I : \mathcal{M}_i \models \varphi \wedge \chi\} \in \mathcal{U} \\
&\Leftrightarrow \{i \in I : \mathcal{M}_i \models \psi(x_i^1, \ldots, x_i^n)\} \in \mathcal{U}
\end{aligned}$$

Finally, let $\overline{(x_i^1)_{i\in I}}, \ldots, \overline{(x_i^n)_{i\in I}} \in \mathcal{M}$ and let $\psi(\alpha_1, \ldots, \alpha_n)$ be a formula which is of the form $\exists\alpha\ \varphi(\alpha, \alpha_1, \ldots, \alpha_n)$. As inductive hypothesis, assume that the statement holds for $\varphi(\overline{(y_i^0)_{i\in I}}, \ldots, \overline{(y_i^n)_{i\in I}})$ and for all $\overline{(y_i^0)_{i\in I}}, \ldots, \overline{(y_i^n)_{i\in I}} \in \mathcal{M}$. We shall show that the statement holds for $\psi(\overline{(x_i^1)_{i\in I}}, \ldots, \overline{(x_i^n)_{i\in I}})$.

Assume that $\prod_{i\in I} \mathcal{M}_i/\mathcal{U} \models \psi(\overline{(x_i^1)_{i\in I}}, \ldots, \overline{(x_i^n)_{i\in I}})$. Then, by definition, there exists $\overline{(x_i^0)_{i\in I}} \in \mathcal{M}$ such that $\prod_{i\in I} \mathcal{M}_i/\mathcal{U} \models \varphi(\overline{(x_i^0)_{i\in I}}, \overline{(x_i^1)_{i\in I}}, \ldots, \overline{(x_i^n)_{i\in I}})$. It follows from the inductive hypothesis that $\{i \in I : \mathcal{M}_i \models \varphi(x_i^0, x_i^1, \ldots, x_i^n)\} \in \mathcal{U}$. This implies that $\{i \in I : \mathcal{M}_i \models \exists\alpha\ \varphi(\alpha, x_i^1, \ldots, x_i^n)\} \in \mathcal{U}$ and hence we have $\{i \in I : \mathcal{M}_i \models \psi(x_i^1, \ldots, x_i^n)\} \in \mathcal{U}$.

[6] We shall supress the free variables in these subformulas to save space.

For the converse direction, assume that $\{i \in I : \mathcal{M}_i \models \psi(x_i^1, \dots, x_i^n)\} \in \mathcal{U}$. Then, $A = \{i \in I : \mathcal{M}_i \models \exists\alpha\ \varphi(\alpha, x_i^1, \dots, x_i^n)\} \in \mathcal{U}$. Using the axiom of choice,

- For $i \in A$, choose an element $x_i^0 \in M_i$ such that $\mathcal{M}_i \models \varphi(x_i^0, x_i^1, \dots, x_i^n)$.
- For $i \notin A$, choose x_i^0 to be an arbitrary element of M_i.

Then, $\overline{(x_i^0)_{i\in I}} \in \mathcal{M}$ and moreover, by construction, we have that

$$\{i \in I : \mathcal{M}_i \models \varphi(x_i^0, x_i^1, \dots, x_i^n)\} \in \mathcal{U}$$

It follows from the inductive hypothesis that

$$\prod_{i\in I} \mathcal{M}_i/\mathcal{U} \models \varphi(\overline{(x_i^0)_{i\in I}}, \overline{(x_i^1)_{i\in I}}, \dots, \overline{(x_i^n)_{i\in I}})$$

and hence

$$\prod_{i\in I} \mathcal{M}_i/\mathcal{U} \models \exists\alpha\ \varphi(\alpha, \overline{(x_i^1)_{i\in I}}, \dots, \overline{(x_i^n)_{i\in I}})$$

This completes the proof that the statement holds for $\psi(\overline{(x_i^1)_{i\in I}}, \dots, \overline{(x_i^n)_{i\in I}})$. Since the statement holds for atomic formulas and it holds for a formula whenever it holds for its subformulas, it follows by induction on the complexity of formulas that the statement holds for all formulas. □

As a consequence of Łos's theorem, a structure and its ultrapowers are elementarily equivalent. However, while they satisfy the same first-order sentences, structures and their ultrapowers may behave in vastly different ways. To illustrate this, we shall next construct a field of *hyperreal* numbers.

Let $\mathcal{L} = \{+, *, 0, 1, <\}$ be the language of ordered fields and let $\mathcal{U}$ be a non-principal ultrafilter on $\mathbb{N}$. Consider the ultrapower

$$^*\mathbb{R} = \prod_{i\in\mathbb{N}} \mathbb{R}/\mathcal{U}$$

On the one hand, since the cartesian product $\prod_{i\in\mathbb{N}} \mathbb{R}$ has cardinality $\mathfrak{c}$, we have that $|^*\mathbb{R}| \le \mathfrak{c}$. On the other hand, the map given by $x \mapsto \overline{(x, x, \dots)}$ is injective[7] and hence $\mathfrak{c} \le |^*\mathbb{R}|$. It follows that $|^*\mathbb{R}| = \mathfrak{c}$. By Łos's theorem, $^*\mathbb{R}$ is elementarily equivalent to the ordered field of real numbers and hence is an ordered field itself. So far, we have constructed an ordered field of cardinality $\mathfrak{c}$ which, from the point of view of first-order statements, behaves exactly like the real numbers.

Set $\epsilon = \overline{(\frac{1}{1}, \frac{1}{2}, \frac{1}{3}, \dots)} \in {}^*\mathbb{R}$. Let $r \in \mathbb{R}^+$ and consider $\overline{(r, r, r \dots)} \in {}^*\mathbb{R}$. As $\mathbb{R}$ is a Archimedean field, there exists $K \in \mathbb{N}$ such that $\frac{1}{K} < r$. Since $\mathcal{U}$ is non-principal, it contains the Fréchet filter and hence

$$\{i \in \mathbb{N} : K \le i\} \subseteq \left\{i \in \mathbb{N} : \mathbb{R} \models \frac{1}{i+1} < r\right\} \in \mathcal{U}$$

Thus, by Łos's theorem, we have that $\epsilon < \overline{(r, r, r \dots)}$ for every $r \in \mathbb{R}^+$. In other words, the ordered field $^*\mathbb{R}$ contains *infinitesimal* elements and is non-Archimedean. One can similarly show that the multiplicative inverse of ϵ which is $\epsilon^{-1} = \overline{(1, 2, 3, \dots)}$ is an "infinite" number in the sense that $\mathbf{1} < \epsilon^{-1}$, $\mathbf{1} + \mathbf{1} < \epsilon^{-1}$, $\mathbf{1} + \mathbf{1} + \mathbf{1} < \epsilon^{-1}$ and so on where $\mathbf{1} = \overline{(1, 1, 1, \dots)}$. The reader may find it intriguing that one can

[7]In general, by Łos's theorem, the diagonal map $x \mapsto \overline{(x, x, x, \dots)}$ is an elementary embedding from a structure into any of its ultrapowers.

carry out all of Calculus using the hyperreal field via the help of infinitesimal just as Newton and Leibniz originally imagined.[8]

4. Topologizing ultrafilters and Stone-Čech compactification

In general topology, there are various ways to "compactify" a topological space, that is, to densely embed a topological space into a compact topological space. In this section, we shall learn one of these compactification methods, namely, the Stone-Čech compactification. While there are several ways to carry out the construction of this compactification for arbitrary spaces, we shall focus on the construction of the Stone-Čech compactification of a discrete space using ultrafilters.

Fix $\mathbf{X}$ be a discrete topological space and let

$$\beta\mathbf{X} = \{\mathcal{U} : \ \mathcal{U} \text{ is an ultrafilter on } \mathbf{X}\}$$

For any subset $A \subseteq \mathbf{X}$, consider the set

$$O_A = \{\mathcal{U} \in \beta\mathbf{X} : \ A \in \mathcal{U}\}$$

It is straightforward to check that $O_A \cap O_B = O_{A\cap B}$ and hence the collection

$$\mathcal{B} = \{O_A : A \in \mathcal{P}(\mathbf{X})\} \subseteq \mathcal{P}(\beta\mathbf{X})$$

is closed under finite intersection. Consequently, it is a base for a topology on $\beta\mathbf{X}$, namely, the topology τ whose open sets are unions of sets of the form O_A. We shall next explore some basic properties of the topological space $(\beta\mathbf{X}, \tau)$.

Theorem 4. *$\beta\mathbf{X}$ is a compact Hausdorff space.*

Proof. Let $\mathcal{U}, \mathcal{V}$ be distinct points in $\beta\mathbf{X}$. Then there exists $A \subseteq \mathbf{X}$ such that $A \in \mathcal{U} - \mathcal{V}$ and $\mathbf{X} - A \in \mathcal{V} - \mathcal{U}$. It follows that $\mathcal{U} \in O_A$ and $V \in O_{\mathbf{X}-A}$. Therefore $\beta\mathbf{X}$ is Hausdorff.

To show that $\beta\mathbf{X}$ is compact, we shall use the following characterization of compactness: A Hausdorff space is compact if and only if every collection of closed sets satisfying the finite intersection property has non-empty intersection.

We claim that any collection $\mathcal{A}$ with the finite intersection property which consists of sets of the form O_A has non-empty intersection. Let $\mathcal{A}$ be such a family, say, $\mathcal{A} = \{O_A : A \in D\}$ for some set $D \subseteq \mathcal{P}(\mathbf{X})$. Now, for any finite subset $F \subseteq D$, by assumption, there exists $\mathcal{U} \in \bigcap_{A\in F} O_A$ and hence $\bigcap_{A\in F} A \in \mathcal{U}$, which implies that $\bigcap_{A\in F} A \neq \emptyset$. Therefore, D has the finite intersection property. By Proposition 1 and Theorem 1, there exists an ultrafilter $\mathcal{V} \supseteq D$ on $\mathbf{X}$. Then, by definition, we have that $\mathcal{V} \in \bigcap_{A\in D} O_A = \bigcap \mathcal{A}$. This completes the proof of the claim.

Now, observe that $\beta\mathbf{X} - O_A = O_{\mathbf{X}-A}$ and hence the collection of sets of the form O_A are also a base for closed sets, that is, any closed set is the intersection of such sets. Thus, if we take a collection $\hat{\mathcal{A}}$ of closed sets with the finite intersection property, we can find a collection $\mathcal{A}$ with the finite intersection property which consists of sets of the form O_A such that $\bigcap \hat{\mathcal{A}} = \bigcap \mathcal{A}$. However, we have just proven that $\bigcap \mathcal{A} =\neq$ for such families. Thus, every collection of closed sets satisfying the finite intersection property has non-empty intersection and hence $\beta\mathbf{X}$ is compact. □

Theorem 5. *The collection of clopen sets of $\beta\mathbf{X}$ is $\{O_A : A \in \mathcal{P}(\mathbf{X})\}$.*

[8] The curious reader may Google the term "non-standard analysis".

Proof. As observed in the previous proof, we have that $\beta\mathbf{X} - O_A = O_{\mathbf{X}-A}$ for all $A \subseteq \mathbf{X}$ and hence O_A is clopen for every $A \subseteq \mathbf{X}$. Now, let $C \subseteq \beta\mathbf{X}$ be a clopen set. Then, $\mathcal{A} = \{O_A : A \subseteq \mathbf{X} \text{ and } O_A \subseteq C\}$ is an open cover of C as O_A's form a base for the topology of $\beta\mathbf{X}$. As C is closed and $\beta\mathbf{X}$ is compact by Theorem 4, C is compact as well and hence there exists finitely many $A_1, A_2, \ldots, A_n \subseteq \mathbf{X}$ such that $C = \bigcup_{i=1}^n O_{A_i} = O_{A_1 \cup A_2 \cup \cdots \cup A_n}$. □

For the rest of this section, given $x \in \mathbf{X}$, the principal ultrafilter $\{A \subseteq \mathbf{X} : x \in A\}$ will be denoted by $\mathcal{U}_x$. It is easily seen that $x \neq y$ implies that $\mathcal{U}_x \neq \mathcal{U}_y$.

Theorem 6. *The collection $\{\mathcal{U}_x \in \beta\mathbf{X} : x \in X\}$ of principal ultrafilters on $\mathbf{X}$ is precisely the set of isolated points of $\beta\mathbf{X}$ and is dense in $\beta\mathbf{X}$.*

Proof. Let $x \in X$. Then $O_{\{x\}} = \mathcal{U}_x$ and hence $\mathcal{U}_x$ is an isolated point of $\beta\mathbf{X}$. Now, let $\mathcal{U}$ be an isolated point of $\beta\mathbf{X}$. Then, by definition, there exists a non-empty $A \subseteq \mathbf{X}$ such that $O_A = \{\mathcal{U}\}$. Let $x \in A$. Then $\{x\} \subseteq A \in \mathcal{U}_x$ and so $\mathcal{U}_x \in O_A$ which implies that $\mathcal{U}_x = \mathcal{U}$.

To see that $\{\mathcal{U}_x \in \beta\mathbf{X} : x \in X\}$ is dense in $\beta\mathbf{X}$, we need to show that this set intersects every non-empty open set. Let $O \subseteq \beta\mathbf{X}$ be a non-empty open set. Then, by definition, there exists a non-empty $A \subseteq \mathbf{X}$ such that $O_A \subseteq O$. Let $x \in A$. Then $\{x\} \subseteq A \in \mathcal{U}_x$ and hence $\mathcal{U}_x \in O_A \subseteq O$, which completes the proof. □

Before we prove the main result of this section, let us recall some definitions. Given a completely regular topological space X, a *Stone-Čech compactification* of X is a compact Hausdorff space Z together with an embedding $\varphi : X \to Z$ such that $\varphi[X]$ is dense in Z and for every compact Hausdorff space Y and for every continuous map $f : X \to Y$ there exists a continuous map $g : Z \to Y$ with $f = g \circ \varphi$. It is not difficult to check that *the* Stone-Čech compactification is unique up to homeomorphism leaving the image $\varphi[X]$ is fixed.

Theorem 7. *$\beta\mathbf{X}$ is the Stone-Čech compactification of the discrete space $\mathbf{X}$, where $\mathbf{X}$ is embedded into $\beta\mathbf{X}$ via the map $\varphi : \mathbf{X} \to \beta\mathbf{X}$ given by $x \mapsto \mathcal{U}_x$.*

Proof. In Theorem 4 and Theorem 6, we have already proven that $\beta\mathbf{X}$ is a compact Hausdorff space and that $\varphi[\mathbf{X}]$ is dense in $\beta\mathbf{X}$.

It remains to check that $\beta\mathbf{X}$ has the aforementioned lifting property. Let Y be a compact Hausdorff space and $f : \mathbf{X} \to Y$ be a continuous map. Define a map $g : \beta\mathbf{X} \to Y$ as follows: For every $\mathcal{U} \in \beta\mathbf{X}$, choose some element

$$g(\mathcal{U}) \in \bigcap_{A \in \mathcal{U}} \overline{f[A]}$$

where $\overline{f[A]}$ denotes the closure of the set $f[A]$. We claim that g is continuous and satisfies $f = g \circ \varphi$. To see the latter, let $x \in \mathbf{X}$. Then $\varphi(x) = \{x\} \in \mathcal{U}_x$ and so

$$g(\varphi(x)) = g(\mathcal{U}_x) \in \bigcap_{A \in \mathcal{U}_x} \overline{f[A]} \subseteq \overline{f[\{x\}]} = \overline{\{f(x)\}} = \{f(x)\}$$

which means that $g(\varphi(x)) = f(x)$.

We now show that g is continuous. Let $O \subseteq Y$ be an open set. We wish to show that $g^{-1}[O]$ is open. If $O \cap im(g) = \emptyset$, then $g^{-1}[O] = \emptyset$ which is clearly open. Assume that $O \cap im(g) \neq \emptyset$, say, $\mathcal{U} \in g^{-1}[O]$ for some $\mathcal{U} \in \beta\mathbf{X}$. Since Y is regular, we can find some open set $V \subseteq O$ such that $g(\mathcal{U}) \in V \subseteq \overline{V} \subseteq O$. Set $A = f^{-1}[V]$.

We claim that $A \in \mathcal{U}$. Assume towards a contradiction that $A \notin \mathcal{U}$. Then $\mathbf{X} - A \in \mathcal{U}$ and, by the definition of g, we have that

$$g(\mathcal{U}) \in \bigcap_{S \in \mathcal{U}} \overline{f[S]} \subseteq \overline{f[\mathbf{X} - A]}$$

But then, the definition of closure and V being a neighborhood of $g(\mathcal{U})$ together that $V \cap f[\mathbf{X} - A] \neq \emptyset$, which is a contradiction as $V = f^{-1}[A]$. Therefore $A \in \mathcal{U}$.

Since $A \in \mathcal{U}$, we have that $\mathcal{U} \in O_A$. We claim that $O_A \subseteq g^{-1}[O]$ which would complete the proof that O is open. Assume towards a contradiction that there exists $\mathcal{V} \in O_A$ such that $g(\mathcal{V}) \notin O$. Then, $Y - \overline{V}$ is an open neighborhood of $g(\mathcal{V})$. On the other hand, by the definition of g, we have that

$$g(\mathcal{V}) \in \bigcap_{S \in \mathcal{V}} \overline{f[S]} \subseteq \overline{f[A]}$$

and hence $Y - \overline{V} \cap f[A] \neq \emptyset$, contradicting that $A = f^{-1}[V]$. This completes the proof that g is continuous. Therefore, $\beta\mathbf{X}$ is the Stone-Čech compactification of $\mathbf{X}$. $\square$

The study of the Stone-Čech compactification of discrete spaces, such as $\beta\mathbb{N}$, has been an active area of research which has important applications. The curious reader may check [HS12] which is just devoted to this topic.

5. Ultralimits and amenability of $\mathbb{Z}$

(To be added later.)

6. Principal ultrafilters and dictators

(To be added later.)

References

[Hod93] Wilfrid Hodges, *Model theory*, Encyclopedia of Mathematics and its Applications, vol. 42, Cambridge University Press, Cambridge, 1993. MR 1221741

[HS12] Neil Hindman and Dona Strauss, *Algebra in the Stone-Čech compactification*, De Gruyter Textbook, Walter de Gruyter & Co., Berlin, 2012, Theory and applications, Second revised and extended edition [of MR1642231]. MR 2893605

[Hun80] Thomas W. Hungerford, *Algebra*, Graduate Texts in Mathematics, vol. 73, Springer-Verlag, New York-Berlin, 1980, Reprint of the 1974 original. MR 600654

[Jec03] Thomas Jech, *Set theory*, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2003, The third millennium edition, revised and expanded. MR 1940513

[Mun00] James R. Munkres, *Topology*, Prentice Hall, Inc., Upper Saddle River, NJ, 2000, Second edition of [MR0464128]. MR 3728284

[Sol70] Robert M. Solovay, *A model of set-theory in which every set of reals is Lebesgue measurable*, Ann. of Math. (2) **92** (1970), 1–56. MR 265151

[SS17] Haimanti Sarbadhikari and Shashi Mohan Srivastava, *A course on basic model theory*, Springer, Singapore, 2017. MR 3700974

Middle East Technical University, Ankara, Turkey,
E-mail address: `burakk@metu.edu.tr`