2 Ultraproducts

The next tool we look at is ultraproducts. These are a tool for building new structures from old ones allowing to approximate the desired properties of the end structure in the structures along the way. We start by looking at the general notion of a structure.

2.1 Structures

From the viewpoint of logic almost any construct you come across in mathematics is a structure. Examples include graphs, groups, fields, vector spaces, Banach spaces, operator algebras, etc. A structure is just a set equipped with various constants, relations and functions. These are determined by the vocabulary, that determines which aspects of a familiar structure we are studying: e.g., when looking at the reals $\mathbb{R}$, do we mean the dense linear order, the field, or something else?

Definition 2.1. A *vocabulary* is a set of constant symbols c, relation symbols R and function symbols f. We always assume that the relation and function symbols have associated to them a unique natural number n_R (or n_f), the *arity* of the relation (function).

Example 2.2.

- The smallest vocabulary is the empty vocabulary, that is used when studying pure sets.
- The vocabulary of rings is $\{+, -, \cdot, 0, 1\}$, where $+$ and $\cdot$ are binary functions, $-$ is a unary function and $0, 1$ are constants.
- The vocabulary of graphs is $\{E\}$, where E is the binary edge relation.
- The vocabulary of linear orders is $\{<\}$.

Definition 2.3. Given a vocabulary L, an *L-structure* $\mathcal{M}$ consists of

- the *domain* of $\mathcal{M}$, $\mathrm{dom}(\mathcal{M}) = M$, which is a nonempty set,
- for each constant symbol $c \in L$, an element $c^{\mathcal{M}} \in M$,
- for each relation symbol $R \in L$, a n_R-ary relation $R^{\mathcal{M}} \subseteq M^{n_R}$,
- for each function symbol $f \in L$, a n_f-ary function $f^{\mathcal{M}} : M^{n_f} \to M$.

Often one uses the same symbol both for the symbols of the vocabulary and their interpretation in a given model.

Example 2.4. A graph $\mathcal{G} = (V, E)$ consists of a nonempty set of vertices V and the (irreflexive and symmetric) binary edge relation $E \subset V^2$ (see Figure 2).

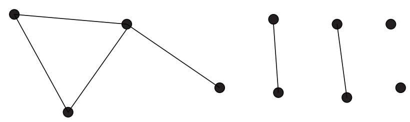

Figure 2: A graph.

Definition 2.5. Suppose $\mathcal{M}$ and $\mathcal{N}$ are L-structures with universes M and N respectively. An *L-embedding* $F : \mathcal{M} \to \mathcal{N}$ is an injection $F : M \to N$ that preserves the interpretation of all of the symbols of L, i.e.,

- $F(c^{\mathcal{M}}) = c^{\mathcal{N}}$ for all constant symbols $c \in L$,
- $(a_0, \ldots, a_{n-1}) \in R^{\mathcal{M}}$ if and only if $(F(a_0), \ldots, F(a_{n-1})) \in R^{\mathcal{N}}$ for all relation symbols $R \in L$ and $a_0, \ldots, a_{n-1} \in M$,
- $F(f^{\mathcal{M}}(a_0, \ldots, a_{n-1})) = f^{\mathcal{N}}(F(a_0), \ldots, F(a_{n-1}))$ for all function symbols $f \in L$ and $a_0, \ldots, a_{n-1} \in M$.

A bijective L-embedding is an *L-isomorphism.*

If $M \subseteq N$ and the inclusion map is an L-embedding, we say that $\mathcal{M}$ is a *substructure* of $\mathcal{N}$ and that $\mathcal{N}$ is an *extension* of $\mathcal{M}$.

Note that substructures need to have interpretations for all constants and they need to be closed under all functions. However, for a purely relational vocabulary, any subset of a structure is a substructure (with relations restricted to the appropriate powers of the new domain).

2.2 Filters

To define ultraproducts we need the notion of ultrafilter. These are used also in topology, e.g., to define a more general notion of convergence than the sequential one.

Definition 2.6. Let I be a set. A *filter* D on I is a nonempty collection of subsets of I satisfying

- $\emptyset \notin D$,
- if $X, Y \in D$, then $X \cap Y \in D$,

- if $X \in D$ and $X \subseteq Y \subseteq I$, then $Y \in D$.

A filter D on I is an *ultrafilter* if in addition

- for all $X \subseteq I$, either $X \in D$ or $I \backslash X \in D$.

A filter D is called *principal*, if it is of the form

$$D = \{X \subseteq I : i \in X\}$$

for some $i \in I$.

Principal filters are ultrafilters, but we are usually interested in the *non-principal* ones. A particularly useful filter is the *Frechet filter*:

Example 2.7. The *Frechet filter* or cofinite filter on $\mathbb{N}$ is defined by:

$$F = \{X \subseteq \mathbb{N} : N \backslash X \text{ is finite}\}.$$

It is clearly nonempty ($\mathbb{N} \in F$), $\emptyset \notin F$ and upwards closed. If $X, Y \in F$, then the complement of $X \cap Y$ is the union of the complements of X and Y, so F is a filter.

The Frechet filter is not an ultrafilter, but that can be mended (assuming AC). For this we need the notion of a generated filter.

Definition 2.8. A collection F of sets is said to have the *finite intersection property*, if any finite collection of sets from F has finite intersection.

A collection F with the finite intersection property *generates* the filter

$$\langle F \rangle = \{X \subseteq \bigcup F : \bigcap_{j \in J} X_j \subseteq X, X_j \in F, F \text{ finite}\}.$$

Note that the finite intersection property is exactly what is needed of F for $\langle F \rangle$ to be a filter. Now we can expand any filter to an ultrafilter (to build an ultrafilter 'out of nothing', let $D = I$ in the theorem).

Theorem 2.9. *Any filter can be extended to an ultrafilter.*

Proof. We prove this by transfinite induction. So let D be a given filter on I. Enumerate the powerset of I,

$$\mathcal{P}(I) = \{X_\alpha : \alpha < 2^{|I|}\}.$$

Now inductively define filters D_α as follows: $D_0 = D$. When D_α has been defined, look at X_α. If $X \cap X_\alpha = \emptyset$ for some $X \in D_\alpha$, let $D_{\alpha+1}$ be the filter generated by $D_\alpha \cup \{I \backslash X_\alpha\}$, either let $D_{\alpha+1}$ be generated by $D_\alpha \cup \{X_\alpha\}$. We need to check that $D_{\alpha+1}$ is a filter. If $X \cap X_\alpha \neq \emptyset$ for any $X \in D_\alpha$, then clearly $D_\alpha \cup \{X_\alpha\}$ has the finite intersection property. If $X \cap X_\alpha = \emptyset$ for some X, we need to show that $D_\alpha \cup \{I \backslash X_\alpha\}$ has the finite intersection property. If not, there are $X_1, \ldots, X_n \in D_\alpha$ such that $X_1 \cap \cdots \cap X_n \cap (I \backslash X_\alpha) = \emptyset$, i.e., $X_1 \cap \cdots \cap X_n \subseteq X_\alpha$, so $X_1 \cap \cdots \cap X_n \cap X \subseteq X_\alpha \cap X = \emptyset$, contradicting the fact that D_α was a filter. So $D_\alpha \cup \{I \backslash X_\alpha\}$ must have the finite intersection property, and $D_{\alpha+1}$ can be defined.

For limit δ, let $D_\delta = \bigcup_{\alpha<\delta} D_\alpha$. It is easy to see that D_δ is a filter.

In the end, $D_{2^{|I|}}$ will be an ultrafilter extending D. □

2.3 Ultraproducts

Definition 2.10. Let I be a set and D a filter over I. For each $i \in I$, let X_i be a set. Then we can define an equivalence relation $\sim_D$ on the cartesian product $\prod_{i \in I} X_i$ by

$$(a_i)_{i \in I} \sim_D (b_i)_{i \in I} \text{ iff } \{i \in I : a_i = b_i\} \in D.$$

This is an equivalence relation: reflexivity and symmetry are trivial, and transitivity follows from the assumption that D is a filter and noticing that

$$\{i \in I : a_i = c_i\} \supseteq \{i \in I : a_i = b_i\} \cap \{i \in I : b_i = c_i\}.$$

The set $\prod_{i \in I} X_i/D$ denotes the set of equivalence classes of $\sim_D$. It is called the *reduced product of X_i modulo D.*

If D is an ultrafilter, the reduced product is called an *ultraproduct*. If all the sets X_i are the same set X, the ultraproduct is called the *ultrapower of X modulo D.*

This construction is the basis for defining ultraproducts of models.

Definition 2.11. Let L be a vocabulary. Let I be a nonempty set, D an ultrafilter over I and for each $i \in I$, let $\mathcal{M}_i$ be an L-structure. The ultraproduct $\prod_{i \in I} \mathcal{M}_i/D$ is the L-structure $\mathcal{M}$ defined by the following:

- The domain of $\prod_{i \in I} \mathcal{M}_i/D$ is $\prod_{i \in I} M_i/D$ (where M_i is the domain of $\mathcal{M}_i$.
- If $c \in L$ is a constant symbol and c_i denotes the interpretation of c in $\mathcal{M}_i$, $c^{\mathcal{M}} = (c_i)_{i \in I}/D$.

- If $R \in L$ is an n–ary relation symbol, then $R^{\mathcal{M}}$ is defined by

  $(a^1_D, \dots a^n_D) \in R^{\mathcal{M}}$ if and only if $\{i \in I : (a^1(i), \dots, a^n(i)) \in R^{\mathcal{M}_i}\} \in D$.

  for any elements $a^k_D = (a^k(i))_{i \in I}/D \in \prod_{i \in I} M_i/D$.

- If $F \in L$ is an n–ary function symbol, then

$$F^{\mathcal{M}}(a^1_D, \dots a^n_D) = (F^{\mathcal{M}_i}(a^1(i), \dots, a^n(i)))_{i \in I}/D.$$

A crucial part of the definition is the following lemma showing that the structure on $\mathcal{M}$ is well-defined. Note that the interpretation of constant symbols is fine, since it does not depend on any chosen representatives but only gives an element of $\prod_{i \in I} M_i/D$.

Lemma 2.12. *Let I, D, L and $\mathcal{M}_i$ be as in Definition 2.11. Suppose that $a^1_d = b^1_D, \dots, a^n_D = b^n_D$. Then*

$$\{i \in I : (a^1(i), \dots, a^n(i)) \in R^{\mathcal{M}_i}\} \in D \text{ if and only if } \{i \in I : (b^1(i), \dots, b^n(i)) \in R^{\mathcal{M}_i}\} \in D$$

and

$$(F^{\mathcal{M}_i}(a^1(i), \dots, a^n(i)))_{i \in I} \sim_D (F^{\mathcal{M}_i}(b^1(i), \dots, b^n(i)))_{i \in I}.$$

Proof. Denote $X_k = \{i \in I : a^k(i) = b^k(i)\}$. By assumption $X_k \in D$ for all $k \in \{1, \dots, n\}$.

For the first claim assume $X := \{i \in I : (a^1(i), \dots, a^n(i)) \in R^{\mathcal{M}_i}\} \in D$. But now

$$X \cap X_1 \cap \dots \cap X_n \subseteq \{i \in I : (b^1(i), \dots, b^n(i)) \in R^{\mathcal{M}_i}\}$$

so as D is a filter, $\{i \in I : (b^1(i), \dots, b^n(i)) \in R^{\mathcal{M}_i} \in D$. The other direction is symmetrical.

For the second claim, note that $F^{\mathcal{M}_i}(b^1(i), \dots, b^n(i)) = F^{\mathcal{M}_i}(a^1(i), \dots, a^n(i))$, whenever $b^k(i) = a^k(i)$ for each $k \in \{1, \dots, n\}$ (as $F^{\mathcal{M}_i}$ is a function. Thus

$$X_1 \cap \dots X_n \subseteq \{i \in I : F^{\mathcal{M}_i}(b^1(i), \dots, b^n(i)) = F^{\mathcal{M}_i}(a^1(i), \dots, a^n(i))\}$$

so $\{i \in I : F^{\mathcal{M}_i}(b^1(i), \dots, b^n(i)) = F^{\mathcal{M}_i}(a^1(i), \dots, a^n(i))\} \in D$ which is the claim. □

The above lemma makes sure that Definition 2.11 gives a well-defined L-structure. We will look closer at this structure and its properties once we have a logic to express these properties.

2.4 Predicate logic

Predicate logic (or first order logic) is a formal language used as a tool in studying structures. It is built from basic statements, *atomic formulas*, using *connectives* ('not' $\neg$, 'and' $\wedge$, 'or' $\vee$) and *quantifiers* ('there exists' $\exists$, 'for all' $\forall$).

The objects formulas 'talk about' are called *terms*. They are built from *variables* and named elements of the structure.

Definition 2.13. Let L be a vocabulary. The set of L–terms is defied by:

- If v_j is a variable, then v_j is an L-term.
- If $c \in L$ is a constant symbol, then c is an L-term.
- If $t_1, \ldots, t_n$ are L-terms and $F \in L$ is an n-ary function symbol, then $F(t_1, \ldots, t_n)$ is an L–term.

The most simple formulas one can form are the atomic formulas:

Definition 2.14. Let L be a vocabulary. The set of atomic L-formulas is defined by:

- If t_1, t_2 are L–terms, then $t_1 = t_2$ is an atomic L-formula.
- If $R \in L$ is an n-ary relation symbol, and $t_1, \ldots, t_n$ are L-terms, then $R(t_1, \ldots, t_n)$ is an atomic L–formula.

Combining atomic formulas by connectives and quantifiers gives the full set of L-formulas. Note that we do not include all connectives, as the omitted ones can be expressed using $\neg$ and $\wedge$.

Definition 2.15. Let L be a vocabulary. The set of L-formulas is defined by:

- Atomic L-formulas are L-formulas.
- If φ, ψ are L-formulas, then so are $\neg\varphi$, $(\varphi \wedge \psi)$ and $(\varphi \vee \psi)$.
- If φ is an L-formula and v_j is a variable, then $\exists v_j \varphi$ and $\forall v_j \varphi$ are L–formulas.

Before we can define how to interpret the meaning of formulas in structures, we need to note the difference between free and bound variables.

Definition 2.16. The *free* variables of a formula φ, $Free(\varphi)$ can be defied inductively as follows:

- The free variables of an atomic formula are all the variables occurring in the formula.
- $Free(\neg\varphi) = Free(\varphi)$ and $Free((\varphi \wedge \psi) = Free(\varphi \vee \psi) = Free(\varphi) \cup Free(\psi)$.
- $Free(\exists v_j \varphi) = Free(\forall v_j \varphi) = Free(\varphi) \backslash \{v_j\}$.

The variables occurring in a formula that are not free are *bound*. A formula with no free variables is a *sentence*.

The role of sentences and formulas with free variables differ a bit: formulas express properties of tuples of elements from a model while sentences express properties of the models as a whole. To see exactly what a formula expresses we need a definition of when it is satisfied and when not.

In the following writing a formula φ in the form $\varphi(v_1, \ldots, v_n)$ means that the free variables occurring in φ are among the variables $v_1, \ldots, v_n$. We then define what it means for a formula $\varphi(v_1, \ldots, v_n)$ to hold of a tuple $(a_1, \ldots, a_n) \in M^n$ in a model $\mathcal{M}$. For that we need to know how to interpret terms:

Definition 2.17. Let L be a vocabulary, $\mathcal{M}$ an L-structure, and t an L–term built using variables from $\bar{v} = (v_1, \ldots, v_n)$. The *value of t at $\bar{a}$*, $t[\bar{a}]$, where $\bar{a} = (a_1, \ldots, a_n) \in M^n$ is defined inductively by:

1. If $t = v_i$, then $t[\bar{a}] = a_i$.
2. If $t = c$ a constant symbol in L, then $t[\bar{a}] = c^{\mathcal{M}}$.
3. If $t = F(t_1, \ldots, t_n)$, where $F \in L$ is an n–ary function symbol and $t_1, \ldots, t_n$ are L-terms, then
$$t[\bar{a}] = F^{\mathcal{M}}(t_1[\bar{a}], \ldots, t_n[\bar{a}]).$$

We can then define satisfaction of a formula.

Definition 2.18. Let L be a vocabulary, $\mathcal{M}$ an L-structure, $\varphi(v_1, \ldots, v_n)$ an L–formula and $\bar{a} = (a_1, \ldots, a_n) \in M^n$. The concept of $\mathcal{M}$ *satisfying* $\varphi[\bar{a}]$ (or $\bar{a}$ *satisfying* φ *in* $\mathcal{M}$), denoted $\mathcal{M} \models \varphi[\bar{a}]$, is defined inductively by:

1. If φ is $t_1 = t_2$, then $\mathcal{M} \models \varphi[\bar{a}]$ iff $t_1[\bar{a}] = t_2[\bar{a}]$.
2. If φ is $R(t_1, \ldots, t_n)$, then $\mathcal{M} \models \varphi[\bar{a}]$ iff $(t_1[\bar{a}], \ldots, t_n[\bar{a}]) \in R^{\mathcal{M}}$.
3. If φ is $\neg\psi$ then $\mathcal{M} \models \varphi[\bar{a}]$ iff $\mathcal{M} \not\models \psi[\bar{a}]$.

4. If φ is $(\psi \wedge \theta)$ then $\mathcal{M} \models \varphi[\bar{a}]$ iff both $\mathcal{M} \models \psi[\bar{a}]$ and $\mathcal{M} \models \theta[\bar{a}]$ hold.
5. If φ is $(\psi \vee \theta)$ then $\mathcal{M} \models \varphi[\bar{a}]$ iff at least one of $\mathcal{M} \models \psi[\bar{a}]$ and $\mathcal{M} \models \theta[\bar{a}]$ holds.
6. If $\varphi(v_1, \ldots, v_n)$ is $\exists v_j \psi(v_1, \ldots, v_n, v_j)$ then $\mathcal{M} \models \varphi[a_1, \ldots, a_n]$ iff there is some $b \in M$ such that $\mathcal{M} \models \psi[a_1, \ldots, a_n, b]$.
7. If $\varphi(v_1, \ldots, v_n)$ is $\forall v_j \psi(v_1, \ldots, v_n, v_j)$ then $\mathcal{M} \models \varphi[a_1, \ldots, a_n]$ iff for any $b \in M$, $\mathcal{M} \models \psi[a_1, \ldots, a_n, b]$ holds.

2.5 Łos's theorem

The following theorem by Łos is central in using ultraproducts, and it is also known under the name *The fundamental theorem of ultraproducts.*

Theorem 2.19 (Łos's theorem)**.** *Let I be an infinite set, D an ultrafilter over I, L a vocabulary, and for each $i \in I$, let $\mathcal{M}_i$ an L-structure. Denote the ultraproduct $\mathcal{M} := \prod_{i \in I} \mathcal{M}_i / D$.*

1. *For any L-term $t(v_1, \ldots, v_n)$ and elements $a^1_D, \ldots, a^n_D \in M$,*

$$t[a^1_D, \ldots, a^n_D] = (t^{\mathcal{M}_i}[a^1(i), \ldots, a^n(i)])_{i \in I}/D,$$

where $t^{\mathcal{M}_i}$ refers to the value of t being calculated in $\mathcal{M}_i$.

2. *For any L-formula $\varphi(v_1, \ldots, v_n)$ and any $a^1_D, \ldots, a^n_D \in \mathcal{M}$,*

$$\mathcal{M} \models \varphi[a^1_D, \ldots, a^n_D] \text{ if and only if } \{i \in I : \mathcal{M}_i \models \varphi[a^1(i), \ldots, a^n(i)]\} \in D.$$

Proof. 1. The claim is proved by induction on the structure of terms. If t is a constant or a variable, this holds by definition. If t is of the form $F(t_1, \ldots, t_m)$ and the claim holds for the terms $t_1, \ldots, t_m$, then , denoting $\bar{a}_D = (a^1_D, \ldots, a^n_D)$ and $\bar{a}(i) = (a^1(i), \ldots, a^n(i))$,

$$\begin{aligned} t^{\mathcal{M}}[\bar{a}_D] &= F^{\mathcal{M}}(t_1[\bar{a}_D], \ldots, t_m[\bar{a}_D]) \\ &= F^{\mathcal{M}}((t_1^{\mathcal{M}_i}[\bar{a}(i)])_{i \in I}/D, \ldots, (t_m^{\mathcal{M}_i}[\bar{a}(i)])_{i \in I}/D) \\ &= (F^{\mathcal{M}_i}(t_1^{\mathcal{M}_i}[\bar{a}(i)], t_m^{\mathcal{M}_i}[\bar{a}(i)]))_{i \in I}/D \\ &= (t^{\mathcal{M}_i}[\bar{a}(i)])_{i \in I}/D, \end{aligned}$$

proving the claim.

2. The claim is proved by induction on the structure of formulas.

(a) If φ is of the form $t_1 = t_2$, then by the above

$$t_1[a_D^1, \ldots, a_D^n] = t_2[a_D^1, \ldots, a_D^n] \text{ if and only if}$$
$$(t_1^{\mathcal{M}_i}[a^1(i), \ldots, a^n(i)])_{i \in I} \sim_D (t_2^{\mathcal{M}_i}[a^1(i), \ldots, a^n(i)])_{i \in I},$$

i.e., if $\{i \in I : \mathcal{M}_i \models (t_1 = t_2)[a^1(i), \ldots, a^n(i)]\} \in D$.

(b) The case where $\varphi(v_1, \ldots, v_n)$ is of the form $R(t_1, \ldots, t_m)$ is left as an exercise.

(c) The cases where φ is of the form $\neg\psi$, $(\psi \wedge \theta)$ or $(\psi \vee \theta)$ are left as an exercise.

(d) Let $\varphi(v_1, \ldots, v_n) = \exists v_j \psi(v_1, \ldots, v_n, v_j)$ and assume the claim holds for ψ. Now if

$$\mathcal{M} \models \varphi[a_D^1, \ldots, a_D^n],$$

then there is some $a_D^j \in M$ such that $\mathcal{M} \models \psi[a_D^1, \ldots, a_D^n, a_D^j]$. By induction hypothesis

$$X := \{i \in I : \mathcal{M}_i \models \psi[a^1(i), \ldots, a^n(i), a^j(i)]\} \in D$$

and $X \subseteq \{i \in I : \mathcal{M}_i \models \exists v_j \psi[a^1(i), \ldots, a^n(i)]\} \in D$. For the other direction, if

$$X' := \{i \in I : \mathcal{M}_i \models \exists v_j \psi[a^1(i), \ldots, a^n(i)]\} \in D,$$

then for each $i \in X'$ let $a^j(i)$ be s.t. $\mathcal{M}_i \models \psi[a^1(i), \ldots, a^n(i), a^j(i)]$. For $i \in I \backslash X'$, let $a^j(i)$ be arbitrary. Let $a_D^j = (a^j(i))_{i \in I}/D$. Then $\{i \in I : \mathcal{M}_i \models \psi[a^1(i), \ldots, a^n(i), a^j(i)]\} \supseteq X'$ so the set is in D. By induction hypothesis $\mathcal{M} \models \psi[a_D^1, \ldots, a_D^n, a_D^j]$, so $\mathcal{M} \models \exists v_j \psi[a_D^1, \ldots, a_D^n]$.

(e) The case where $\varphi(v_1, \ldots, v_n) = \forall v_j \psi(v_1, \ldots, v_n, v_j)$ is left as an exercise.

□

2.6 Applications

One of the most important consequences of Łos's theorem is the compactness theorem of first order logic (this can also be proved by other methods, as is done, e.g., on the course 'Matemaattinen logiikka').

Definition 2.20. Let L be a vocabulary. A set T of L-sentences is *satisfiable* if there exist an L-structure $\mathcal{M}$ such that $\mathcal{M} \models \varphi$ for all $\varphi \in T$. This is often written $\mathcal{M} \models T$.

Theorem 2.21 (Compactness theorem)**.** *Let T be a set of sentences of first order logic. Then T is satisfiable if and only if every finite subset of T is.*

Proof. The direction from left to right is trivial. So we prove the direction from right to left. Assume every finite subset of T is satisfiable. Let I be the set of finite subsets of T. By assumption, for each $i \in I$ we can find a model $\mathcal{M}_i$ such that $\mathcal{M}_i \models i$. Let $X_i = \{j \in I : i \subseteq j\}$. Then the collection $\{X_i : i \in I\}$ has the finite intersection property and thus generates a filter, which can be extended to an ultrafilter D (by Theorem 2.9). Now by Łos's theorem $\prod_{i \in I} \mathcal{M}_i / D \models T$. $\square$

Ultraproducts can also be used to transfer properties, e.g., between fields of positive and of zero characteristic. Recall that a field is a structure in the vocabulary $\{+, -, 0, 1\}$, where $+$ and $\cdot$ are binary function symbols and 0 and 1 are constant symbols. The axioms state $(M, +)$ is an Abelian group with neutral element 0

$$\forall x \forall y \forall z\, x + (y + z) = (x + y) + z,\ \forall x (x + 0 = x \wedge 0 + x = x),$$
$$\forall x \exists y (x + y = 0 \wedge y + x = 0),\ \forall x \forall y\, x + y = y + x,$$

that $(M \backslash 0, \cdot)$ is an Abelian group with neutral element 1

$$\forall x \forall y \forall z\, x \cdot (y \cdot z) = (x \cdot y) \cdot z,\ \forall x (x \cdot 1 = x \wedge 1 \cdot x = x),$$
$$\forall x \exists y (x = 0 \vee (x \cdot y = 1 \wedge y \cdot x = 1)),\ \forall x \forall y\, x \cdot y = y \cdot x,$$

as well as distributivity

$$\forall x \forall y \forall z\, x \cdot (y + z) = (x \cdot y) + (x \cdot z)$$

and non-triviality

$$\neg 0 = 1.$$

So we see that the property of being a field can be expressed with first order logic (above we have used ordinary shorthand such as x, y, z for v_0, v_1, v_2 and $x + y$ for $+(x, y)$). We could also have included functions for the inverses in the vocabulary (and then defined, e.g., $0^{-}1 = 0$ just in order to have a total function).

Now recall that if there is $n \in \mathbb{N}$ such that $1 + 1 + \cdots + 1$ (n repeated terms) equals 0, then the smallest such n is the *characteristic* of the field. The smallest such number must be a prime number. If no such number exists the characteristic of the field is said to be zero.

Theorem 2.22. *Let P be an infinite set of primes and for each $p \in P$ let F_p be a field of characteristic p. If D is a non-principal ultrafilter over P, then $\prod_{p\in P} F_p/D$ is a field of characteristic zero.*

Proof. Since the axioms expressed above capture the notion of a field and are true in all F_p, the ultraproduct $\prod_{p\in P} F_p/D$ is a field. Now let $\Sigma = \{\varphi_n : n \geq 1\}$ where φ_n states that $n \cdot 1 \neq 0$,i.e.

$$\varphi_n = \neg 1 + 1 + \cdots + 1 = 0 \text{ (n repetitions of 1).}$$

Now for each $n \geq 1$, the set $\{p \in P : F_p \models \varphi_n\}$ is cofinite ($F_p \models \varphi$ iff $p \nmid n$). By Exercise 5.6 a non-principal ultrafilter always contains the cofinite filter, so $\{p \in P : F_p \models \varphi_n\} \in D$. Thus $\prod_{p\in P} F_p/D \models \Sigma$ and thus cannot have positive characteristic. So it has characteristic 0. $\square$

As a corollary one sees, that if ψ expresses a property which is true in all fields of positive characteristic, then ψ is true in some field of characteristic zero. And conversely, if ψ is true in fields of zero characteristic, it is true in 'almost all' fields of positive characteristics. This phenomenon, and the related phenomenon of axiomatizability vs. non-axiomatizability is explored in the exercises.