2010 SUMMER COURSE ON MODEL THEORY

PETE L. CLARK

Contents

Introduction

0.1. Some theorems in mathematics with snappy model-theoretic proofs.

1) The Nullstellensatz and the $\mathbb{R}$-Nullstellensatz.
2) Chevalley's Theorem: the image of a constructible set is constructible.
3) (Grothendieck, Ax) An injective polynomial map from $\mathbb{C}^n$ to $\mathbb{C}^n$ is surjective.
4) Hilbert's 17th problem: a positive semidefinite rational function $f \in \mathbb{R}(t_1, \ldots, t_n)$ is a sum of squares.
5) Polynomially compact operators have invariant subspaces. (Let V be a complex Hilbert space, L a bounded linear operator on H, and $0 \neq P(t) \in \mathbb{C}[t]$. Suppose that $P(L)$ is a compact operator: the image of the unit ball has compact closure. Then there exists a nontrivial, proper closed subspace W of V which is L-invariant.)
6) (Ax-Kochen) For each $n \in \mathbb{Z}^+$ and all sufficiently large primes p, a homogeneous form with coefficients in $\mathbb{Q}_p$ with at least $n^2 + 1$ variables has a nontrivial zero.
7) (Duesler-Knecht) An analogue of the Ax-Katz theorem for rationally connected varieties over the maximal unramified extension of $\mathbb{Q}_p$.
8) (Faltings, Hrushovski) Mordell-Lang Conjecture.

Remark: So far as I know, these are in increasing order of difficulty. However, I have barely glanced at the proof of 5), so this is just a guess.

1. Languages, structures, sentences and theories

1.1. Languages.

Definition: A **language** $\mathcal{L}$ is the supply of symbols that are deemed admissible when appearing in an expression. There are three types of symbols that make up a language: for each $n \in \mathbb{Z}^+$, a set of **n-ary function symbols** $f = f(x_1, \ldots, x_n)$ – formally f is just a symbol which has the number n associated to it; for each $n \in \mathbb{Z}^+$, a set of **n-ary relation symbols** $R = R(x_1, \ldots, x_n)$ – formally exactly the same as a function – and a set of **constant symbols**.

We assume that all of these sets are disjoint – i.e., that given an element $a \in \mathcal{L}$, we can say unambiguously that it is an n-ary function symbol for a unique $n \in \mathbb{Z}^+$, an n-ary relation symbol (for a unique n) or a constant. Otherwise the sets are quite arbitrary: any or all of them may be empty (and most of them usually will be in any given application), and the sets may be infinite, even uncountably infinite.

An $\mathcal{L}$-structure X is given by the data of
- an underlying set[1], which we (abusively) also denote X,
- for all n and for each n-ary function f a map of sets n-ary function $f_X : X^n \to X$,
- for all n and for each n-ary relation R a subset $R_X \subset X^n$,
- for each constant c an element $c_X \in X$.

That is, to endow a set X with an $\mathcal{L}$-structure is to *assign* to each element "formal n-ary function" $f \in \mathcal{L}$ an actual n-ary function on X, to each "formal n-ary relation" $R \in \mathcal{L}$ an actual n-ary relation on X, and to each "formal constant" $c \in \mathcal{L}$ an actual element ("constant") of X.

Exercise 1.1: Let $\mathcal{L}$ be a language, and let X be a set.
a) If X is nonempty, show that X can always be endowed with an $\mathcal{L}$-structure.
b) For which languages $\mathcal{L}$ can the empty set be endowed with an $\mathcal{L}$-structure?

Exercise 1.2: Explain how a constant can be viewed as a 0-ary function. (Nevertheless, we will just call them constants.)

We also have the notion of morphism of $\mathcal{L}$-structures. In fact, there are two natural such notions.

Let X and Y be $\mathcal{L}$-structures. A **homomorphism** $\varphi : X \to Y$ is a map of the underlying sets such that:
(HS1) For each constant $c \in \mathcal{L}$, $\varphi(c_X) = c_Y$.
(HS2) For each n-ary function $f \in \mathcal{L}$ and all $(x_1, ldots, x_n) \in X^n$, we have

$$f_Y((\varphi(x_1), \ldots, \varphi(x_n)) = \varphi(f_X(x_1, \ldots, x_n)).$$

(HS3) For each n-ary relation $R \in \mathcal{L}$ and all $(x_1, \ldots, x_n) \in R_X$, we have

$$(\varphi(x_1), \ldots, \varphi(x_n)) \in R_Y.$$

On the other hand, an **embedding** $\iota : X \to Y$ of $\mathcal{L}$-structures is an *injective* mapping on the underlying sets satisfying (HS1), (HS2) and the following strengthened version of (HS3):

(IS3) For each n-ary relation $R \in \mathcal{L}$ and all $(x_1, \ldots, x_n) \in X^n$, we have $(x_1, \ldots, x_n) \in R_X \iff (\varphi(x_1), \ldots, \varphi(x_n)) \in R_Y$.

An **isomorphism** of $\mathcal{L}$-structures is, as usual, a homomorphism of $\mathcal{L}$-structures which has a two-sided inverse homomorphism.

Exercise 1.3: For a homomorphism $\varphi : X \to Y$ of $\mathcal{L}$-structures, TFAE:
(i) φ is an isomorphism of $\mathcal{L}$-structures.
(ii) φ is a surjective embedding of $\mathcal{L}$-structures.

[1]Most of the texts I have consulted require X to be nonempty. Although this will certainly be the case of interest to us, I do not see – yet – why it is necessary or helpful to assume this at the outset. Let me know if you spot why!

We demonstrate that the notion of an $\mathcal{L}$-structure is a familiar one by identifying some of the smallest possible $\mathcal{L}$-structures as standard examples of mathematical structures.

Example 1.4: If $\mathcal{L} = \emptyset$, an $\mathcal{L}$-structure is precisely a set; a homomorphism of $\mathcal{L}$-structures is precisely a mapping between sets, and an embedding of $\mathcal{L}$-structures is simply an injective mapping of sets.

Example 1.5: Let $\mathcal{L}$ be the language with a single unary relation R. Then an $\mathcal{L}$-structure is a **pair** of sets (X, S), i.e., a set X together with a distinguished subset S. A homomorphism $\varphi : (X_1, S_1) \to (X_2, S_2)$ is a homomorphism of pairs in the usual sense of (e.g.) algebraic topology: a mapping $X_1 \to X_2$ such that $\varphi(S_1) \subset S_2$. An embedding of pairs is an injective map such that $\varphi^{-1}(S_2) \subset S_1$.

This example can evidently be generalized: if $\mathcal{L}$ consists of a set I of unary relations, then an $\mathcal{L}$-structure is an I-tuple: i.e., a set endowed with an I-indexed family of subsets.

Example 1.6: If $\mathcal{L}$ consists of a single unary function f, an $\mathcal{L}$-structure is a function $f : M \to M$: a "discrete dynamical system."

Example 1.7: If $\mathcal{L}$ consists of a single binary function $\cdot$, an $\mathcal{L}$-structure is a binary composition law $\cdot : M \times M \to M$, i.e., what Bourbaki calls a magma. If $\mathcal{L} = \{\cdot, e\}$, i.e., a binary function together with a constant, we get a pointed magma.

Here is a key point: we are probably more interested in more restricted subcategories of (pointed) magmas: e.g., in which the composition law is associative (semigroup) and for which the constant is an identity element (monoid) and with respect to which there are multiplicative inverses (a group). We have an idea that a semigroup is a magma with additional structure. This is not captured by the idea of an $\mathcal{L}$-structure but rather by the idea of imposing additional axioms on the structure that we already have. This is coming up shortly!

Example 1.8: The underlying $\mathcal{L}$-structure of the category of rings consists of two constant symbols 0 and 1 and two binary operations $+$ and $-$.

Example 1.9: The underlying $\mathcal{L}$-structure of the category of ordered rings consists of the above elements together with a binary relation $<$.

Example 1.10: If $\mathcal{L}$ consists of a single binary relation $\sim$, then an $\mathcal{L}$-structure is a (simple, undirected) graph. This is a good example to use to reflect on the difference between homomorphisms and embeddings. For instance, let $n > 1$ be an integer. Let X_1 be the graph on the vertex set $\{1, \ldots, n\}$ with no edges: i.e., the binary relation is the empty relation. Let X_2 be the complete graph on the vertex set $\{1, \ldots, n\}$, i.e., $i \sim j \iff i \neq j$. Let ι be the identity map on $\{1, \ldots, n\}$. Then ι is a bijective homomorphism from X_1 to X_2 but is not an embedding.

Non-example 1.11: There is no appropriate $\mathcal{L}$-structure for topological spaces. That is because the data of a topological space is a set X together with a unary relation (i.e., subset) $\mathcal{T}$ on 2^X. As it turns out, the notion of $\mathcal{L}$-structure is fundamentally unable to treat subsets on the same footing as sets. This is certainly worth thinking about. For instance, the reader may wonder whether we can capture the structure of a topological space via a set I of unary relations as in Example 1.5 above. The problem here is that the set I would have to depend upon the topological space X: think about it!

The general term for $\mathcal{L}$-structures is **relational structures**, i.e., the sort of structure which is given by n-ary relations (after all an n-ary function is a special kind of $(n+1)$-ary relation and a constant is a 0-ary function). As we have seen, many (but not all!) instances of mathematical structure are special cases of relational structure. There is a branch of mathematics which studies the (homomorphism and embedding) categories of general relational structures: it is called **universal algebra**. This is not what we want to study! As alluded to above, in conventional mathematics it is natural to restrict the categories of $\mathcal{L}$-structures by requiring them to satisfy certain *axioms*. Thus to get to model theory at all we need to add the **syntactic** ingredient: i.e., to define first-order formulas and statments.

1.2. Statements and Formulas.

Thinking of $\mathcal{L}$ as giving us an admissible set of symbols, a *formula* is a syntactically correct expression made out of those symbols. For a formal definition of what a formula is, consult [Marker, pp. 9-10]; we will content ourselves with the following description, which should allow the reader to see in each case whether an expression is or isn't a formula. First we give ourselves a supply of *variables* $\{v_i\}_{i=1}^{\infty}$. Actually this is overly formalistic – in practice we can use any single character which is not already denoting something else in a formula; we often use x, y, z for variables. (The reader will see shortly why we do not need to take more than a countable set of variables.) Then a formula is built out of:

variables: formally speaking, we should take a fixed countably infinite set of variables, say $\{x_i\}_{i=1}^{\infty}$; in practice we will use whatever lower case letter we like;
constants, functions, relations
equality: $=$

logical symbols: $\neg, \vee, \wedge, \implies, \iff, \exists, \forall$
parentheses: (,)

We are not trying to do computer programming or anything of the sort, so we will feel free to use new symbols which are commonly understood to be composed out of our basic symbols. For instance, we all understand $\neq$ to be shorthand for $\neg =$, and we will use this freely in formulas.

For example, the following are all valid formulas in the language $(+, \cdot, 0, 1)$ of rings:

$$0 = 1.$$

$$x^2 + y^2 = z^2.$$

(Here we are again employing shorthand: by x^2, we of course really mean $x \cdot x$. Note that x^y does not have an agreed upon meaning in terms of the given symbols, so that would certainly be out of bounds.

$$\forall b \forall c \; \exists x \; (x^2 + bx + c = 0).$$

(Here we understand bx to be an abbreviation for $b \cdot x$.)

The key point is the following: given any formula φ in a language $\mathcal{L}$, and given any $\mathcal{L}$-structure X, the formula has a **semantic interpretation** in the structure X.

Note that the formulas (1) and (3) are of a different nature than formula (2). Formulas (1) and (3), when interpreted in any ring R, make a mathematical statement about the ring R, which is either true or false (and not both!). For instance (1) is true in a ring R iff R is the ring with a single element. Formula (3) asserts that any monic quadratic equation with coefficients in R has a solution in R: this is true, for instance, if $R = \mathbb{C}$ and false if $R = \mathbb{Z}$ or $\mathbb{Q}$ or $\mathbb{R}$. However, intrepreting (2) in the ring $R = \mathbb{R}$ of real numbers, say, it does not make sense to say that it is true or false. It depends on what values the variables x, y, z take.

The logical formalization of this lies in the distinction between **bound** and **free** variables. A variable v appearing in a formula is said to be **bound** if it comes after a quantifier $\exists v$ or $\forall v$: otherwise it is said to be **free**. A **sentence** is a formula in which all variables are bound. (In particular, a formula in which no variables appear is automatically a sentence.)

Of course the sentences will play a distinguished and important role in what follows. But considering formulas with free variables is also important: it is the key to the geometric interpretation of model theory. Namely, a statement must either be true or false, but a formula φ with unbound variables $(x_1, \ldots, x_n)$ may be viewed as a parameterized family of sentences: for each structure X and each $a = (a_1, \ldots, a_n)$ in X^n, the formula φ **defines** a subset of X^n: namely the set of parameters $\varphi(a)$ is true. Note that this directly generalizes the notion of an affine variety as being defined by a finite set of polynomial equations.[2]

1.3. **Satisfaction.**

Let us come back to the case of sentences. Suppose that $\mathcal{L}$ is a language, X is an $\mathcal{L}$-structure, and φ is an $\mathcal{L}$-sentence. Then we have a fundamental dichotomy: either φ is true in the structure X, or it is false. In the former case we say that X **satisfies** φ. This is sufficiently important that it gets its own symbol:

$$X \models \varphi.$$

The following result is another instance of how easy things become when we do not need to worry about quantifiers.

Proposition 1. *Let $X \subset Y$ be an embedding of $\mathcal{L}$-structures, and let $\phi(\overline{v})$ be a quantifier-free formula with n unbound variables. Then for all $\overline{a} \in X^n$, $X \models \phi(\overline{a})$ iff $Y \models \phi(\overline{a})$.*

[2] From a logical standpoint, there is no difference between a single statement and a finite set of statements, since $\varphi_1, \ldots, \varphi_n$ are all true iff the single statement $\bigwedge_{i=1}^n \varphi_i$ is true.

Proof. Let's try to see what this proposition means. Suppose there are no free variables in the formula. Then since there are no quantifiers either, there are no variables at all, i.e. the formula expresses some fact about the constants. Since $X \hookrightarrow Y$ is an embedding of sets preserving the constants, clearly this sentence has the same truth value in X and Y. Now take an n-tuple $\overline{a} \in X^n$. By evaluating at $\overline{a}$ we have essentially reduced to the previous case, except that the coordinates of $\overline{a}$ may not be distinguished constants in the language in question. But contemplate enlarging the language by adding these new constants – then the above argument goes through to show the equivalence of $X \models \phi(\overline{a})$ and $Y \models (\overline{a})$ in the expanded language. But this is just formal nonsense: the truth of $\phi(\overline{a})$ is the same in both languages! $\square$

1.4. **Elementary equivalence.**
Let X_1 and X_2 be two $\mathcal{L}$-structures. We say that X_1 and X_2 are **elementarily equivalent** – denoted $X_1 \equiv X_2$ – if for all $\mathcal{L}$-sentences φ, $X_1 \models \varphi \iff X_2 \models \varphi$. Thus two structures are elementarily equivalent if they satisfy exactly the same first-order sentences.

This is an immediately appealing notion, because it suggests the possibility of an important proof technique, that of **transfer**: suppose one is trying to prove a certain theorem about a structure X. Suppose that the theorem can be expressed as $X \models \varphi$ where φ is an $\mathcal{L}$-sentence (or a set of $\mathcal{L}$-sentences; more on that shortly). Then, we are free to replace X with any structure X' which is elementarily equivalent to X and which is either simpler in some objective way or somehow tailored so as to make the sentence φ easier to prove. Then we are free to prove the sentence φ in X' (by whatever mathematical means we can), and then the truth is *transferred* to X. In fact this is very close – but not quite – the way we will prove that injective polynomial maps on the complex numbers are surjective. A better example is **nonstandard analysis**: there are ordered fields $\mathcal{R}$ elementarily equivalent to the real numbers but with infinitesimal elements: as long as one can restrict the theorems one wishes to prove to first-order statements, it is perfectly acceptable to do one's analysis in the field $\mathcal{R}$. This proof technique was systematized by A. Robinson, who showed that – after a certain logical scaffolding for the "transfer" is set up once and for all – one can with complete rigor prove theorems in calculus and real analysis by reasoning with infinitesimal elements. In other words, the methods of Newton and Leibniz, which had been justly denigrated for hundreds of years, can be made sound!

A good question is whether the introduction of infinitesimals is actually helpful: do they allow one to prove things that the epsilontically trained mathematician cannot? The answer is in principle no, as will be answered by Gödel's Completeness Theorem. However, one can ask this question of almost any mathematical technique: does one really need complex analysis to prove the prime number theorem? In fact no, as Erdos and Selberg famously showed. But it certainly helps! There are definitely theorems that were first proven using nonstandard methods. An early example is the subspace theorem for polynomially compact operators, by Bernstein and Robinson [BR66].

The principle that elementary equivalence is a much coarser relation than isomorphism is the driving force behind model theory. Unfortunately, at the present time it is difficult to give explicit examples of elementarily equivalent but nonisomorphic structures: much of the theory we will develop is pointing in that direction. But at least we can give a cardinality argument.

Exercise 1.12: Let $\mathcal{L}$ be a language.
a) Show that the number of $\mathcal{L}$-sentences is $\max(\aleph_0, |\mathcal{L}|)$.
b) Deduce that the number of pairwise non-elementarily equivalent $\mathcal{L}$-structures is at most $2^{\max(\aleph_0,|\mathcal{L}|)}$.
c) Show that for any cardinal number κ, there are more than κ pairwise nonisomorphic $\mathcal{L}$-structures.
d) Deduce that for any cardinal number κ there exists a set of κ pairwise nonisomorphic $\mathcal{L}$-structures which are all elementarily equivalent.

Exercise 1.13: Let X be a finite $\mathcal{L}$-structure. Show that if an $\mathcal{L}$-structure X' is elementarily equivalent to X, then $|X'| = |X|$.

Exercise 1.14 (harder): Let X be a finite $\mathcal{L}$-structure and X' an $\mathcal{L}$-structure. Show that $X \equiv X'$ iff $X \cong X'$.

1.5. **Theories.**

A **theory** $\mathcal{T}$ in the language $\mathcal{L}$ is a set of $\mathcal{L}$-sentences. If X is an $\mathcal{L}$-structure, then $X \models \mathcal{T}$ means that X satisfies every sentence in $\mathcal{T}$. In such a situation, we say that X is a **model** of $\mathcal{T}$.

Notice that the class of $\mathcal{L}$-structures which are models of a given theory $\mathcal{T}$ form a category (in two different ways), simply by defining a homomorphism (resp. an embedding) of models of $\mathcal{T}$ to be a homomorphism (resp. embedding) of the underlying $\mathcal{L}$-structures.

Example 1.15: Let $\mathcal{L}$ be the language with one binary operation $\cdot$ and one constant symbol e. Let $\mathcal{T}$ consist of the two sentences:

$$\varphi_A : \forall x \forall y \forall z (x \cdot y) \cdot z = x \cdot (y \cdot z).$$

$$\varphi_e : \forall x \ (x \cdot e = x) \wedge (e \cdot x = x).$$

Then $\mathcal{T}$ is the theory of monoids, and the resulting category is precisely the category of monoids (resp. the category of monoids and monoid embeddings).

In a similar way we can define the theories of groups, rings, integral domains and fields. This is possible because in each case, the familiar axioms satisfied by these structures are captured by sentences in the appropriate language. (The reader should check this!)

We say that a class $\mathcal{C}$ of $\mathcal{L}$-structures is **finitely axiomatizable** if there exists a sentence $\varphi_{\mathcal{C}}$ such that an L-structure X lies in $\mathcal{C}$ iff $X \models \varphi_{\mathcal{C}}$. Informally speaking, we may think of membership in a class $\mathcal{C}$ as being a "property" of a structure, and conversely. For instance, in the language of rings, being commutative, being an integral domain, being a field, being the zero ring, are all finitely axiomatizable. A simple but important observation is that if a property is finitely axiomatizable, then its negation is also finitely axiomatizable.

Exercise 1.16: Let n be a positive integer. Let $\mathcal{L}$ be any language. Show that the properties "X has at least n elements", "X has at most n elements" and "X has exactly n elements" are all finitely axiomatizable.

We say that a property $\mathcal{P}$ of $\mathcal{L}$-structures is **elementary** or **first-order** if there exists an $\mathcal{L}$-theory $\mathcal{T}$ such that an $\mathcal{L}$-structure X has property $\mathcal{P}$ iff $X \models \mathcal{T}$.

Example 1.17: We show that the class of algebraically closed fields is elementary. Indeed, we can certainly write down one sentence that enforces commutativity:

$$\varphi_c : \forall x \forall y \; xy = yx$$

and another sentence that enforces that every nonzero element has a multiplicative inverse:

$$\varphi_e : \forall x \exists y \; xy = yx = 1.$$

To get the property of algebraic closure, the evident way to go is to include, for each $n \in \mathbb{Z}^+$, a sentence φ_n asserting that every monic degree n polynomial has a root. For instance, we can take φ_3 to be:

$$\forall a \forall b \forall c \exists x \; x^3 + ax^2 + bx + c = 0.$$

We define $\{\varphi_c, \varphi_e, \varphi_n\}_{n=1}^{\infty}$ to be the theory of algebraically closed fields. Thus the class of algebraically closed fields is elementary.

Note that this leaves open the question as to whether the class of algebraically closed fields is finitely axiomatizable. The reader might think that we could dispose of this by replacing the infinite collection of sentences φ_n by the single sentence

$$\Phi : \forall n \in \mathbb{Z}^+, \forall c_0, \ldots, c_{n-1}, \; \exists x \; x^n + c_{N-1}x^{n-1} + \ldots + c_1 x + c_0 = 0.$$

However, this is **absolutely not a valid sentence**! The problem is that we have quantified over the positive integers, which is not allowed: quantifiers are not allowed to range over auxiliary sets: every quantifier must pertain to the $\mathcal{L}$-structure itself. Related to this, $\ldots$ is not one of our symbols! Similarly, of course by x^n we mean $x \cdot x \cdot \ldots x$ (n times), which is again illegal. The notion of a "generic integer" is firmly off-limits.[3] In fact, we will see later that the class of algebraically closed fields is *not* finitely axiomatizable (and also why this matters).

A reasonable initial reaction to this state of affairs is that first-order sentences are rather inconveniently limited in their expressive value: why are we working with them? This is a great question, and it has a fantastic answer, coming up next: **Gödel's Completeness Theorem**.

2. Big Theorems: Completeness, Compactness and Löwenheim-Skolem

2.1. **The Completeness Theorem.**

Let $\mathcal{T}$ be a theory in a language $\mathcal{L}$. We say that $\mathcal{T}$ is **satisfiable** if it has at least one model.

[3]In contrast, it is okay to write things like $17x^2 + 2 = 0$, it being understood that this is just a shorthand for $(1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1)(x \cdot x) + (1+1) = 0$.

Exercise 2.1: Let $\mathcal{T}$ be a satisfiable $\mathcal{L}$-theory and φ an $\mathcal{L}$-sentence. Show that the following are equivalent.
(i) $\mathcal{T} \not\models \varphi$.
(ii) $\mathcal{T} \cup \{\neg\varphi\}$ is satisfiable.

A basic and important question is: when is a theory satisfiable?

It is easy to give "silly" examples theories which are not satisfiable. Suppose for instance that for some $\mathcal{L}$-sentence φ, the theory $\mathcal{T}$ contains both φ and its negation $\neg\varphi$. Then evidently $\mathcal{T}$ is not satisfiable!

Evidently there are slightly more examples with the same basic silliness: e.g. suppose that there are sentences $\varphi_1, \varphi_2, \varphi_3$ such that our theory $\mathcal{T}$ contains the following sentences:

$$\varphi_1,\ \varphi_1 \implies \varphi_2,\ \varphi_2 \implies \varphi_3, \varphi_3 \implies \neg\varphi_1.$$

This is clearly no good either: it doesn't matter that $\mathcal{T}$ need not contain the sentence $\neg\varphi_1$, because this is a logical consequence of sentences that it does contain. In fact, let's use more precise terminology – after all, in mathematics, what other kinds of consequences are there than logical consequences? The nonsatisfiability of $\mathcal{T}$ has nothing to do with the *semantics* (i.e., meaning) of the sentences φ_i as interpreted in any $\mathcal{L}$-structure. Rather, their contradictory nature follows from a purely *syntactic* manipulation of these formulas using basic rules of formal logic.

This notion can be formalized by something called **predicate calculus**, i.e., a formal system of deducing various sentences from various other sentences from a fixed set of universal logical rules. As an example, if we are given a list of sentences including φ_1 and $\varphi_1 \implies \varphi_2$, then we should be able to append the sentence φ_2 to the list.

It is most convenient for us to think of the process of **formal proof** as a black box satisfying the following assumptions: with respect to a fixed language $\mathcal{L}$, let $\mathcal{T}$ be a theory and φ a sentence. A formal proof of φ from $\mathcal{T}$ consists of a finite series of sentences $\varphi_1, \ldots, \varphi_n$ generated as follows, where given $\varphi_1, \ldots, \varphi_i$, the sentence φ_{i+1} is either an element of $\mathcal{T}$ or is obtained from the previous sentences by applying one of the finitely many preassigned, valid logical deductions. We require that $\varphi_n = \varphi$. Particular proof systems have been given; a popular one is called **predicate calculus**. We isolate the two most important features of the foregoing informal description.

(Soundness): We write $\mathcal{T} \vdash \varphi$ if there is a formal proof of φ from $\mathcal{T}$: we think of this as **syntactic implication** as it has nothing to do with the meaning of the sentence φ in any model. At the other extreme, we define $\mathcal{T} \models \varphi$ if the sentence φ is true in every model of $\mathcal{T}$. This is **semantic implication** and corresponds to the usual notion of mathematical implication for statements of our restricted "first-order" form. Now we certainly want the following soundness property:

$$(\mathcal{T} \vdash \varphi) \implies (\mathcal{T} \models \varphi).$$

In other words, all of the formal proofs are actually true! As long as our "deduction" rules are true logical facts like the above "modus ponens", this is nothing to worry about.

(Finite character): If $\mathcal{T} \vdash \varphi$, then there exists a finite subset $\mathcal{F} \subset \mathcal{T}$ such that $\mathcal{F} \vdash \varphi$. This also follows from our above description: a proof by definition has finite length, so it can only involve introducing finitely many sentences of $\mathcal{T}$.

We may now state a truly fundamental and spectacular result.

Theorem 2. *(Gödel's Completeness Theorem) Let $\mathcal{T}$ be a theory in a language $\mathcal{L}$ and let φ be an $\mathcal{L}$-sentence. Then*

$$(\mathcal{T} \models \varphi) \iff (\mathcal{T} \vdash \varphi.).$$

Proof. We are not going to prove the completeness theorem here (in a sense that will be made clear soon enough, it is not the business of model theory proper to prove this result). For a proof, see e.g. [BS, Ch. 3]. □

Remark: Perhaps because of the similar name, awareness of Theorem 2 in the larger mathematical community has been largely drowned out by Gödel's Incompleteness Theorems. (The completeness theorem appears in Gödel's 1929 doctoral dissertation; the incompleteness theorems were proved in 1931.) It is not our business here to describe this latter – and yes, even more spectacular – result. Among other sources, wikipedia gives a solid introduction.

Of course the fact that syntactic implication implies semantic implication is just the soundness of our deduction rules referred to above: nothing exciting about that. However, the converse is amazing! It says nothing less than that – provided we restrict our attention to first-order sentences – the notions of mathematical truth and purely syntactic provability coincide.

Corollary 3. *Say that a theory $\mathcal{T}$ is* ***syntactically consistent*** *if for no sentence φ is it the case that $\mathcal{T} \vdash (\varphi \wedge \neg\varphi)$. Then a theory $\mathcal{T}$ is syntactically consistent iff it is satisfiable, i.e., if it has a model X.*

Proof. Again, it is clear that any theory which has a model is syntactically consistent. Inversely, suppose that $\mathcal{T}$ is unsatisfiable, i.e., has no model. Let φ be any sentence whatsoever – e.g. $\forall x\ x = x$. Then it is vacuously true that $\mathcal{T} \models (\varphi \wedge \neg\varphi)$. By the Completeness Theorem, it follows that $\mathcal{T} \vdash (\varphi \wedge \neg\varphi)$, i.e., that $\mathcal{T}$ is synactically inconsistent. □

2.2. **Proof-theoretic consequences of the completeness theorem.**

Gödel's completeness theorem has some metamathematical consequences lying outside of the scope of model theory in general and our course in particular. Nevertheless, some of these consequences are too basic and important to be ignored completely, so we give a brief description now that we hope will enable the reader to make sense of later "decidability" statements thrown off briefly and our course and increasingly ubiquitous elsewhere.

Let $\mathcal{L}$ be a countable language. We say that $\mathcal{L}$ is **recursive** if there exists an algorithm (e.g., Turing machine – we will not need to formalize this concept here) to decide whether a finite sequence of symbols is a syntactically valid formula. Note that a language is certainly recursive if it has only finitly many constants, relations and functions: this already covers most of our applications. Further, we say that an $\mathcal{L}$-theory is recursive if there exists an algorithm which, given an $\mathcal{L}$-sentence φ lies in $\mathcal{T}$.

Simple examples: the theories of rings, integral domains, fields, ordered fields, algebraically closed fields are all recursive. At the other extreme, given an $\mathcal{L}$-structure X, the complete theory $\mathcal{T}(X)$ of all sentences true in X is not necessarily recursive: this is indeed what we would like to be able to *prove* in various examples of interest.

Proposition 4. *If $\mathcal{L}$ is a recurive language and $\mathcal{T}$ is a recursive $\mathcal{L}$-theory, then the set of all sentences φ such that $\mathcal{T} \vdash \varphi$ is recursively enumerable.*

In lieu of a proof we give a definition of recursively enumerable! A theory $\mathcal{T}'$ is recursively enumerable iff there exists a semialgorithm to determine membership in $\mathcal{T}'$: i.e., a Turing machine that when inputted a sentence φ halts iff $\varphi \in \mathcal{T}'$. The result should now be clear: given a recursive set of axioms and a proof system, one can then enumerate all possible proofs. Putting $\mathcal{T}'$ to the set of all syntactic consequences of $\mathcal{T}$, if $\varphi \in \mathcal{T}'$, then eventually we will hit upon a proof of φ.

Corollary 5. *Let $\mathcal{T}$ be a* complete *recursive theory in a recursive language $\mathcal{L}$. Then the set $\mathcal{T}'$ of syntactic consequences of $\mathcal{T}$ is recursive.*

Exercise 2.2: Prove Corollary 5. (Simple but enlightening.)

A theory $\mathcal{T}$ whose set of syntactic consequences is recursive is said to be **decidable**. The point being that if $\mathcal{T}$ is a decidable theory, then by Gödel's completeness theorem there is an algorithm that, given a sentence φ, determines whether $\mathcal{T} \models \varphi$. So, for instance, if we can prove that the complete theory of the real numbers $\mathbb{R}$ is decidable, then it follows that there is an algorithm which decides whether an arbitrary system of polynomial equations has a solution. In fact, more is true. Since in $\mathbb{R}$ the formula $\exists x \ (x^2 = y)$ cuts out the set of non-negaive real numbers, the algorithm can also handle polynomial inequalities: e.g. we can determine whether a system of polynomial equations has a solution in a given open disk. In particular, we can compute roots of univariate polynomial equations to arbitrary accuracy.

That $\mathrm{Th}(\mathbb{R})$ is decidable was proved by Tarski and again, later, by Abraham Robinson. In due time we will give Robinson's proof: this is one of the major results of our course.

Finally, we must mention the theorem of Davis-Matijasevic-Putnam-Robinson[4] that there is no algorithm for deciding whether a system of polynomial equations with $\mathbb{Z}$-coefficients has a $\mathbb{Z}$-solution [Mat]. This implies that the complete theory of $\mathbb{Z}$ in the language of rings is undecidable. In fact it is much stronger: a precise statement is that even the set of *existential sentences* – i.e., sentences in which only existential quantifiers appear – in the language of rings which are true in $\mathbb{Z}$ is not recursive.

[4]Here "Robinson" means Julia Bowman Robinson. Her husband was Raphael Robinson, who was also a leading mathematician working at the intersection of mathematical logic, number theory and geometry. So a theorem of Robinson in model theoretic algebra could be due to any of at least three people.

One says that the existential theory of $\mathbb{Z}$ is undecidable.

On the other hand, Julia Robinson proved the following fantastic theorem [JRob49]: consider the formula

$$\varphi(x, y, z) : \exists a \exists b \exists c \ \ xyz^2 + 2 = a^2 + xy^2 - yc^2$$

and the formula $\psi(x)$ be the formula

$$\psi(x) \forall y \forall z ([\phi(y, z, 0) \wedge (\forall w(\phi(y, z, w) \implies \phi(y, z, w + 1)))] \implies \phi(y, z, x)).$$

Then for $x \in \mathbb{Q}$, $x \in \mathbb{Z}$ iff $\mathbb{Q} \models \psi(x)$. In other words, using $\psi(x)$ allows us to express the condition "x is an integer" in our first-order language.

Exercise 2.3: Deduce from Robinson's theorem and the Davis-Matijasevic-Putnam-Robinson theorem that the complete theory of $\mathbb{Q}$ in the language of rings is undecidable.

In contrast, the question of whether the existential theory of $\mathbb{Q}$ remains undecidable – that is, is there an algorithm to decide whether a system of polynomial equations with $\mathbb{Q}$-coefficients has a $\mathbb{Q}$-solution – remains open, and is surely the leading open problem in the area. Most leading experts I've spoken to belive that the there is no such algorithm. I am most familiar with the work of Bjorn Poonen in this area, who has proved some very nice results in this direction, especially undecidability of the existential theory of rings intermediate between $\mathbb{Z}$ and $\mathbb{Q}$ and are, in a certain precise sense, "much closer to $\mathbb{Q}$ than to $\mathbb{Z}$".

2.3. **The Compactness Theorem.**

Theorem 6. *(Gödel's Compactness Theorem) A theory $\mathcal{T}$ is satisfiable iff every finite subset of $\mathcal{T}$ is satisfiable.*

Proof. The implication $\implies$ is certainly clear. Conversely, assume that every finite subset of $\mathcal{T}$ is satisfiable but $\mathcal{T}$ itself is not satisfiable. As in the proof of Corollary 3 above, let φ be any sentence; then $\mathcal{T} \vdash (\varphi \wedge \neg\varphi)$. But by the finite character of proof, this implies that there exists a finite subset $\mathcal{F}$ of $\mathcal{T}$ that proves $(\varphi \wedge \neg\varphi)$. Thus $\mathcal{F}$ is syntactically inconsistent, so it certainly is not satisfiable, contradiction. $\square$

Remark: Define a theory $\mathcal{T}$ to be **finitely satisfiable** if every finite subset $\mathcal{F} \subset \mathcal{T}$ is satisfiable. Then the compactness theorem can be stated more...compactly?...as follows: a theory $\mathcal{T}$ is satisfiable iff it is finitely satisfiable. One might think that the compactness theorem allows us to dispense with the term "finitely satisfiable" but this turns out not quite to be the case – the term is used later on in these notes.

2.4. **Topological interpretation of the compactness theorem.**

When one first learns of the Compactness Theorem, a question springs immediately to mind: what does this have to do with the concept of compactness that one learns in conventional mathematics, i.e., compactness of topological spaces? I am sorry to say that many model theory texts either do not address this question at all, or brush it away with vague remarks about the "analogy" between finite satisfiability and the finite intersection property which characterizes compact spaces. This is a disservice.

In fact, the compactness theorem for $\mathcal{L}$-theories is equivalent to the compactness of a certain naturally associated topological space.

We need a preliminary definition. Let us say that a theory $\mathcal{T}$ is **complete** if for every sentence φ, exactly one of $\mathcal{T} \models \varphi$, $\mathcal{T} \models \neg\varphi$ holds. On the other hand we define a theory $\mathcal{T}$ to be **maximal** if for evey sentence φ, exactly one of $\varphi \in \mathcal{T}$, $\neg\varphi \in \mathcal{T}$ holds.

Exercise: A theory $\mathcal{T}$ is complete iff its closure $\overline{\mathcal{T}}$ is maximal.

Exercise: a) A maximal theory is precisely a maximal element in the set of consistent $\mathcal{L}$-theories partially ordered by inclusion.
b) Show in two different ways that a theory $\mathcal{T}$ is satisfiable iff it is contained in a maximal theory.
First proof: Let X be a model of $\mathcal{T}$. Then the collection of all sentences which are true in X is a maximal theory containing $\mathcal{T}$.
Second proof: Argue by Zorn's Lemma using the Compactness Theorem.

Now let $\mathbb{X} = \mathbb{X}(\mathcal{L})$ be the set of all maximal $\mathcal{L}$-theories. For each $\mathcal{L}$-sentence φ, let $U(\varphi) = \{\mathcal{T} \in \mathbb{X} \mid \varphi \in \mathcal{T}\}$. We topologize $\mathbb{X}$ by taking the $U(\varphi)$ as a subbase for the open sets, i.e., the nonvoid proper open subsets are precisely the finite intersections of the $U(\varphi)$.[5] In fact $U(\varphi_1) \cap U(\varphi_2) = U(\varphi_1 \wedge \varphi_2)$, so the $U(\varphi)$ form a base for the topology. Moreover, by the definition of a complete theory, $\mathbb{X} \setminus U(\varphi) = U(\neg\varphi)$, so that each $U(\varphi)$ is closed as well as open: it follows that the topology on X is totally disconnected.

Theorem 7. *a) The space $\mathbb{X}(\mathcal{L})$ is Hausdorff and totally disconnected.*
b) The space $\mathbb{X}(\mathcal{L})$ is compact iff every finitely satisfiable $\mathcal{L}$-theory is satisfiable.

Proof. a) Let $\mathcal{T}_1 \neq \mathcal{T}_2$ be distinct maximal $\mathcal{L}$-theories. Then there exists a sentence φ such that $\varphi \in \mathcal{T}_1$ and $\neg\varphi \in \mathcal{T}_2$, so that $\mathcal{T}_1 \in U(\varphi)$, $\mathcal{T}_2 \in U(\neg\varphi)$ and $U(\varphi) \cap U(\neg\varphi) = \emptyset$. Thus $\mathbb{X}(\mathcal{L})$ is totally disconnected in the strong sense that any two distinct points lie in distinct elements of some partition of $\mathbb{X}(\mathcal{L})$ into two disjoint open sets, and this immediately implies that $\mathbb{X}(\mathcal{L})$ is Hausdorff as well.
b) Compactness may be checked on coverings by basis elements, so $\mathbb{X}(\mathcal{L})$ is compact iff every open covering by sets $U(\varphi)$ has a finite subcovering. Taking complements – and using the fact that the set of all $U(\varphi)$'s is stable under complementation – this is equivalent to the assertion that for every family $\{\varphi_i\}_{i\in I}$ of sentences, we have $\bigcap_{i\in I} U(\varphi_i) = \emptyset$ iff there exists a finite subset $J \subset I$ such that $\bigcap_{i\in J} U(\varphi_i) = \emptyset$. But put $\mathcal{T} = \{\varphi_i\}_{i\in I}$. Then $\bigcap_{i\in I} U(\varphi_i)$ is the set of all complete theories containing $\mathcal{T}$; as we saw above, this set is empty iff $\mathcal{T}$ is not satisfiable. Similarly, a finite subset $J \subset I$ such that $\bigcap_{i\in J} U(\varphi_i) = \emptyset$ is a finite subset of $\mathcal{T}$ which is not satisfiable. Thus we have shown that the compactness of $\mathbb{X}$ is equivalent to the assertion that a theory $\mathcal{T}$ is not satisfiable iff there exists a finite subset of $\mathcal{T}$ which is not satisfiable, which is certainly equivalent to the compactness theorem. $\square$

Comment: Let $\widetilde{\mathbb{X}(\mathcal{L})}$ be the set of all *complete* $\mathcal{L}$-theories, topologized via the sets $\tilde{U}(\varphi) = \{\mathcal{T} \mid \mathcal{T} \models \varphi\}$. In a previous draft of these notes, I worked with the space

[5]In this case, we in fact already have $\emptyset = U(\exists x\ (x \neq x)$ and $\mathbb{X} = U(\forall x\ (x = x))$; I am just reminding you of the general definition of the topology generated by an arbitary family of subsets.

$\widetilde{\mathbb{X}(\mathcal{L})}$ instead of $\mathbb{X}(\mathcal{L})$. The above proof works to show that the quasi-compactness of $\widetilde{\mathbb{X}(\mathcal{L})}$ is equivalent to the compactness theorem for $\mathcal{L}$-theories. However, as was pointed out in class by J. Stankewicz, the space $\widetilde{\mathbb{X}(\mathcal{L})}$ is not Hausdorff. Indeed it is not even a T_0, or **Kolmogorov space**.[6] Recall that in a topological space X, we define an equivalence relation, **topological indistinguishability**, as follows:for points x, y, write $x \sim y$ if x and y have exactly the same open neighborhoods. Then any space has a universal T_0-quotient, its **Kolmogorov completion**: as a set, we take $X/\sim$, the set of equivalence classes under topological indistinguishability, and we endow $X/\sim$ with the quotient topology via the map $q : X \to X/\sim$. It is easy to check that on $X/\sim$ any two distinct points are topologically distinguishable. It is interesting to remark that in our context, the topology and the model theory matches up nicely: two theories $\mathcal{T}_1$ and $\mathcal{T}_2$ are topologically indistinguishable iff they have exactly the same models iff $\overline{\mathcal{T}_1} = \overline{\mathcal{T}_2}$. Thus the Kolmogorov quotient of $\widetilde{\mathbb{X}(\mathcal{L})}$ may be identified with the space $\mathbb{X}(\mathcal{L})$ of maximal theories, which is compact (i.e., Hausdorff!).

One may also regard $\mathbb{X}(\mathcal{L})$ as the space of elementary equivalence classes of $\mathcal{L}$-structures.

The space $\mathbb{X}(\mathcal{L})$ is often called the **Stone space**[7] of $\mathcal{L}$: as we have seen, it is a compact, totally disconnected space. Such spaces show up frequently in modern mathematics and are important (for instance) because of the **Stone Duality Theorem**: the category of Stone spaces and continuous maps is anti-equivalent to the category of Boolean rings. For mathematicians with a certain background and inclination, this is an invitation to consider **ultrafilters**.

To many readers the previous paragraph will sound quite mysterious. We will come back to this point later and indeed give a direct proof of the compactness of the Stone space $\mathbb{X}(\mathcal{L})$ and hence of the compactness theorem.

Exercise 2.4: Show that a topological space is compact[8] and totally disconnected iff it is an inverse limit of finite discrete spaces.

2.5. **First applications of compactness.**

We claim that the compactness theorem, and not the (stronger!) completeness theorem is one of the fundamental results of model theory. This is partially true by definition: model theory as a branch of mathematical logic is not concerned with formal provability but only with satisfiability. But there are also deeper reasons: unlike the completeness theorem, which necesarily takes place relative to a formal proof system (although there are multiple systems for which the theorem can be proven) and therefore any proof *must* deal with these formal aspects of proof, there

[6] It follows that it is not totally disconnected, since the closure of a singleton set is connected and not all points are closed.

[7] Warning: there are other, related topological spaces which are also called Stone spaces; c.f. [Mar, §4.1].

[8] Following Bourbaki, for me compact means quasi-compact – every open cover has a finite subcover – and Hausdorff.

are *purely semantic* proofs of the compactness theorem. We will give one especially short and elegant proof later on in the course as an application of **ultraproducts**. Finally and most importantly, the compactness theorem has many, many important model-theoretic consequences. We give some first examples.

Theorem 8. *Let $\mathcal{T}$ be a theory which has arbitrarily large finite models. Then $\mathcal{T}$ has an infinite model.*

Proof. As above, for any $n \in \mathbb{Z}^+$, there exists a sentence φ_n whose interpretation in a structure X is "X has at least n elements. For $n \in \mathbb{Z}^+$, let $\mathcal{T}_n = \mathcal{T} \cup \{\varphi_i\}_{i=1}^n$, and let $\mathcal{T}_\infty = \bigcup_n \mathcal{T}_n$. Now, our hypothesis is precisely that for all n, $\mathcal{T}_n$ has a model. Let $\mathcal{F}$ be any finite subset of $\mathcal{T}_\infty$; then in particular $\mathcal{F}$ contains only finitely many of the sentences φ_n, so $\mathcal{F} \subset \mathcal{T}_n$ for some n. Since $\mathcal{T}_n$ is satisfiable, so is $\mathcal{F}$. Therefore every finite subset of $\mathcal{T}_\infty$ has a model, and so, by compactness, $\mathcal{T}_\infty$ has a model, which is precisely an infinite model of $\mathcal{T}$. $\square$

Corollary 9. *The following classes are* not *elementary classes: the class of all finite sets, the class of all finite groups, the class of all finite abelian groups, the class of all finite rings, the class of all finite fields.*

Remark: Note that for each of the classes of Corollary 9, the *complementary class* – i.e., infinite sets, etc. – is an elementary class. Thus it is possible for the complement of an elementary class to be non-elementary. On the other hand, if a class $\mathcal{C}$ is finitely axiomatizable – i.e., there exists a single sentence φ such that $X \in \mathcal{C}$ iff $X \models \varphi$ – then the complementary class is also finitely axiomatizable: it is the set of models of $\neg\varphi$. This gives a technique for showing that an elementary class is not finitely axiomatizable: show that its complementary class is not elementary. We will return later to address the efficacy of this technique to show that an elementary class is not finitely axiomatizable.

The language of ordered fields is $(+, -, \cdot, <, 0, 1)$. The theory of ordered fields is the theory of fields augmented with sentences expressing the compatibility of the order relation with the field axioms, *viz.*:

$$\forall x \ ((x < 0) \vee (x = 0) \vee (0 < x)),$$
$$\forall a \forall b \forall c \forall d \ ((a < c) \wedge (b < d) \implies (a + b < c + d),$$
$$\forall x \forall y \ (0 < x) \wedge (0 < y) \implies (0 < x \cdot y).$$

In practice, it is convenient to informally introduce additional binary relation symbols $\leq, >, \geq$ with the usual meanings, e.g. we regard $x \geq y$ as an abbreviation for $(y < x) \vee (y = x)$.

Exercise 2.5: a) In ordered field, -1 is not a sum of squares.
b) In particular, an ordered field has characteristic 0 and hence has $\mathbb{Q}$ as a canonical subfield.

A positive element x in an ordered field F is **infinitesimal** if for all $n \in \mathbb{Z}^+$, $x < \frac{1}{n}$. Similarly, a positive element x is **infinitely large** if for all $n \in \mathbb{Z}^+$, $x > n$. It is immediate that x is infinitesimal iff $\frac{1}{x}$ is infinitely large. An ordered field is said to be **non-Archimidean** if it contains infinitesimal elements and **Archimedean** otherwise. For instance, the standard ordering on the real numbers is Archimedean, as are all ordered subfields of $\mathbb{R}$.

Exercise 2.6: Show that, conversely, every Archimedean ordered is isomorphic to a subfield of $\mathbb{R}$. In particular, any ordered field of cardinality greater than $c = |\mathbb{R}|$ is necessarily non-Archimedean.

Theorem 10. *There exists a non-Archimedean ordered field $\mathcal{R}$ which is elementarily equivalent to $\mathbb{R}$.*

Proof. Let $\mathcal{T}$ be the complete theory of the real numbers, i.e., the collection of all sentences in the language $\mathcal{L}$ of ordered fields which are true in $\mathbb{R}$. We extend $\mathcal{L}$ by adding a single constant element c, and we extend $\mathcal{T}$ to an $\mathcal{L}^*$ theory $\mathcal{T}^*$ by adding the following infinite family of sentences: for each $n \in \mathbb{Z}^+$,

$$\varphi_n : c > n.$$

Every finite subset of the theory $\mathcal{T}^*$ is satisfiable, so by compactness $\mathcal{T}^*$ is satisfiable: let $\mathcal{R}$ be a model. Evidently the element $c_{\mathcal{R}}$ is infinitely large, so $\mathcal{R}$ is non-Archimedean. However, viewing $\mathcal{R}$ merely as an ordered field (and not as an ordered field with a distinguished infinitely large element), it is certainly a model of $\mathcal{T}$. Moreover, since $\mathcal{T}$ is the set of all sentences which are true in $\mathbb{R}$, it is a maximal satisfiable theory: adding any other sentence would give a contradiction. Therefore $\mathcal{T}$ is precisely the set of sentences in $\mathcal{L}$ which are true for $\mathcal{R}$, whence $\mathcal{R} \equiv \mathbb{R}$. $\square$

Theorem 11. *Let $\kappa \geq \max(\aleph_0, |\mathcal{L}|)$. Let $\mathcal{T}$ be a theory which admits infinite models. Then $\mathcal{T}$ admits a model of cardinality at least κ.*

Proof. Let $\mathcal{L}^*$ be the scalar extension of $\mathcal{L}$ obtained by adding constant symbols c_i for each $i \in \kappa$. For all distinct indices $i \neq j$ in κ, define the $\mathcal{L}^*$-sentence:

$$\varphi_{i,j} : c_i \neq c_j.$$

Arguing similarly as above, since $\mathcal{T}$ admits infinite models, every finite subset of $\mathcal{T}^*$ is satisfiable. By compactness, $\mathcal{T}^*$ is satisfiable, and a model X of $\mathcal{T}^*$ is precisely a model of $\mathcal{T}$ together with an injection $\kappa \hookrightarrow X$, so X has cardinality at least κ. $\square$

Exercise 2.7: Combine Theorem 11 and Exercise 2.6 to prove (again) Theorem 10.

2.6. **The Löwenheim-Skolem Theorems.**

Theorem 11 shows that the class of structures elementarily equivalent to a given infinite structure contains members of arbitrarily large cardinality. This motivates a closer study of the sizes of models. The results here are strikingly strong: for instance, a theory in a countable language which admits infinite models admits models of every infinite cardinality!

Even this last result is not the full extent of what is true. For that, we need another fundamental definition: let $\iota : X \hookrightarrow Y$ be an embedding of $\mathcal{L}$-structures. We say that ι is an **elementary embedding** if for all formulas $\varphi(x_1, \ldots, x_n)$ – here our convention is that $x_1, \ldots, x_n$ are the unbound variables – and all $(a_1, \ldots, a_n) \in X^n$,

$$X \models \varphi(a_1, \ldots, a_n) \iff Y \models \varphi(\iota(a_1), \ldots, \iota(a_n)).$$

Note that this immediately implies that X and Y are elementarily equivalent: this is the $n = 0$ case, i.e., sentences.[9] But it is much stronger than that: e.g. it is possible to have an embedding of elementarily equivalent (or even isomorphic!) structures that is not an elementary embedding.

The notion of an elementary embedding then is one which fundamentally exploits the existence of formulas containing unbound variables. In fact, one of the major themes of model theory[10] is the interplay between bound and unbound variables. One instance of this is the following simple but important result, which reduces the notion of elementary embedding to that of an "ordinary embedding" in an enriched language.

For an $\mathcal{L}$-structure X, we define a new language $\mathcal{L}_X$ as follows: $\mathcal{L}_X$ consists of the language $\mathcal{L}$ together with a new constant symbol c_x for each element $x \in X$. We may canonically extend X to an $\mathcal{L}_X$ structure by interpreting each constant c_x as the element $x \in X$. Similarly, if $\iota : X \hookrightarrow Y$ is an embedding of $\mathcal{L}$-structures, then we may canonically extend Y to an $\mathcal{L}_X$ structure by interpreting each constant c_x as the element $\iota(x) \in Y$.

Proposition 12. *For an embedding $\iota : X \hookrightarrow Y$ of $\mathcal{L}$-structures, TFAE:*
(i) The map ι is an embedding of the induced $\mathcal{L}_X$-structures.
(ii) The map ι is an elementary embedding of $\mathcal{L}$-structures.

The proof amounts to unwinding the definitions and is left as a good exercise.

Exercise 2.8: Let $\iota : K \hookrightarrow L$ be an elementary embedding of fields.
a) Show that K is algebraically closed in L: i.e. an element of L which is the root of a polynomial equation with coefficients in K is already an element of K.
b) Show L/K is regular: K is algebraically closed in L and $L \otimes_K \overline{K}$ is a domain.

Theorem 13. *(Upward Löwenheim-Skolem) Let X be an infinite $\mathcal{L}$-structure and κ a cardinal with $\kappa \geq |X| + |\mathcal{L}|$. Then there exists an $\mathcal{L}$-structure Y with $|Y| = \kappa$ and an elementary embedding $\iota : X \hookrightarrow Y$.*

Exercise 2.9: Show that each of the following structures admits a model of every infinite cardinality: groups, abelian groups, divisible abelian groups, fields, ordered fields, algebraically closed fields.

Theorem 14. *(Downward Löwenheim-Skolem) Let X be an $\mathcal{L}$-structure and A a subset of X. Let κ be a cardinal such that $|\mathcal{L}| + |A| \leq \kappa \leq |X|$. Then there exists a substructure Y of X satisfying all of the following:*
- *$A \subset Y$;*
- *$|Y| = \kappa$;*
- *the inclusion map $Y \hookrightarrow X$ is an elementary embedding.*

The proofs of Theorems 13 and 14 take some time. Since the emphasis in our course is on the applications of model-theoretic results, we omit the proofs for now.

[9] Even if our notation suggests otherwise, by a formula we always allow the possibility of 0 unbound variables. For one thing, it doesn't matter either way: we could always make a sentence φ into a formula by introducing a new variable x and putting $\varphi'(x) = \varphi \wedge (x = x)$.

[10] Please count and tell me at the end of the course how many major themes there are in model theory!

Depending upon interest, we may return to give the proofs later in the course. (If not, see any introductory text on the subject, e.g. [BS], [Hod], [Mar], [Poi]...)

I was very much struck by the Löwenheim-Skolem theorems when I first learned of them: many times had I wondered "Is there an example of a certain mathematical structure X of cardinality κ?" I knew many *ad hoc* constructions, but it was revelatory to learn that there were general results along these lines!

Here is a very pretty example of Löwenheim-Skolem telling us unfamiliar facts about familiar structures.

Theorem 15. *Let G be an infinite simple group. Then for every cardinal κ with $\aleph_0 \leq \kappa \leq |G|$, G has a simple subgroup of cardinality κ.*

Proof. [Hod, p. 72] The language of groups is countable, so by downward Löwenheim-Skolem there exists a subgroup H of G such that $|H| = \kappa$ and the inclusion of H into G is elementary. Note that since being simple refers to subgroups, it is not clear that it is preserved by elementary equivalence (indeed, we will see later that it is *not*). So we have to argue more cleverly, using the stronger property of an elementary embedding.

First observe that for H to be simple, it suffices to show: if $a, b \in H$ are such that $b \neq e$, then a lies in the least normal subgroup of H generated by b. Certainly this holds true in G, since G is simple. Suppose for example that there are elements y and z of G such that $a = y^{-1}by \cdot z^{-1}b^{-1}z$. Then

$$G \models \exists y \exists z \ (a = y^{-1}by \cdot z^{-1}b^{-1}z).$$

Since H is an elementary substructure of G, this is also true in H! Done. $\square$

Remark: I suppose this theorem is more interesting the more infinite simple groups one knows. Perhaps the easiest construction is as follows: let $n \geq 2$ and F be any field. Then the projective special linear group $PSL_n(F)$ is simple unless $(n, |F|) = (2, 2)$ or $(2, 3)$. In particular, taking F to be an infinite field of cardinality κ (e.g. by Löwenheim-Skolem), one gets infinite simple groups of all cardinalities. However, in this case the conclusion is easy to see by hand: we can get the desired simple group by taking a subfield of the desired cardinality. Other examples: a simple Lie group with trivial center, the group of permutations on $\mathbb{Z}$ modulo permutations which move only finitely many objects.

3. Complete and model complete theories

3.1. Maximal and complete theories.

Let us say that a theory $\mathcal{T}$ is **maximal** if it is maximal among all satisfiable theories: equivalently, for any sentence φ, exactly one of $\varphi, \neg\varphi$ lies in $\mathcal{T}$.

For an $\mathcal{L}$-structure X, we define the **complete theory of X**, $\text{Th}(X)$, to be the set of all $\mathcal{L}$-sentences which are true in X.

Proposition 16. *For a theory $\mathcal{T}$, TFAE:*
(i) $\mathcal{T}$ is maximal.
(ii) There exists an $\mathcal{L}$-structure X such that $\mathcal{T} = \text{Th}(X)$.

Proof. (i) $\implies$ (ii): Let $\mathcal{T}$ be a maximal theory, and let X be a model of $\mathcal{T}$. This means $\mathcal{T} \subset \mathrm{Th}(X)$, and by definition of maximality, we must have equality.
(ii) $\implies$ (i): The set of all sentences which are true in a given structure is clearly a maximal theory. $\square$

It is not in the spirit of model theory to distinguish between theories which have the same models. (Moreover, for applications to decidability results, a maximal theory is often unpleasantly large – e.g. it is not clearly recursive.) This motivates the following weaker – but much more useful – notion.

Recall that a theory $\mathcal{T}$ is **complete** if for every sentence φ, either $\mathcal{T} \models \varphi$ or $\mathcal{T} \models \neg\varphi$, but not both.[11]

Proposition 17. *For a theory $\mathcal{T}$, TFAE:*
(i) $\mathcal{T}$ is complete.
(ii) $\mathcal{T}$ is satisfiable, and if X_1 and X_2 are two models of $\mathcal{T}$, then $X_1 \equiv X_2$.

Exercise 3.1: Prove Proposition 17.

Exercise 3.2 Define the **deductive closure** $\overline{\mathcal{T}}$ of a theory to be the set of all sentences such that $\mathcal{T} \vdash \varphi$.[12]
a) Show that $\mathcal{T} \mapsto \mathcal{T}^{\vdash}$ is a closure operator (this involves checking three simple properties; see e.g. wikipedia for the definition).
b) Show that $\mathcal{T}$ is satisfiable iff $\mathcal{T}^{\vdash}$ is a proper subset of the set of all sentences.
c) Show that for any $\mathcal{L}$-structure X, $\mathcal{T} \models X \iff \mathcal{T}^{\vdash} \models X$. d) Show that $\mathcal{T}$ is complete iff $\mathcal{T}^{\vdash}$ is maximal.

Thus one way to resolve the issue of distinct theories with the same set of models is to always pass to the deductive closure. However, this turns out not to be so desirable. (E.g., from a proof-theoretic perspective, if $\mathcal{T}$ is recursive, then $\mathcal{T}^{\models}$ need only be recursively enumerable.) In fact, a major goal in model theory is somehow the opposite: given a maximal theory $\mathcal{T}$, find a subtheory $\mathcal{T}'$ which is as small as possible such that $\mathcal{T}'^{\models} = \mathcal{T}$.

3.2. **Model complete theories.** A theory $\mathcal{T}$ is **model complete** if every embedding $\iota : X \to Y$ of models of $\mathcal{T}$ is an elementary embedding.

Although neither completeness nor model completeness implies the other (we will see examples shortly), nevertheless the two concepts are closely related.

We define a model X of a theory $\mathcal{T}$ to be **minimal** if for every model Y of $\mathcal{T}$ there exists an embedding of structures $\iota : X \hookrightarrow Y$.

Example: The field $\mathbb{Q}$ is a minimal model for the theory of fields of characteristic 0. The field $\mathbb{F}_p$ is a minimal model for the theory of fields of characteristic $p > 0$.

Example: The field $\mathbb{Q}$ is a minimal model for the theory of ordered fields.

[11]Many standard sources do not require a complete theory to be consistent. Since one is not interested in inconsistent theories – i.e., theories without models – it seems like we are not missing out on anything by excluding them by definition.

[12]Notation in an earlier draft: $\mathcal{T}^{\vdash}$.

Proposition 18. *Let $\mathcal{T}$ be a model complete theory, and let $\mathcal{T}'$ be a satisfiable theory containing $\mathcal{T}$. Then $\mathcal{T}'$ is also model complete.*

Proof. Indeed, let $\iota : X \hookrightarrow Y$ be an embedding between two models of $\mathcal{T}'$. In particular, it is an embedding between two models of $\mathcal{T}$, so by hypothesis is an elementary embedding. (Note that the definition of an elementary embedding, like that of elementary equivalence, refers only to a language $\mathcal{L}$ and not to any $\mathcal{L}$-theory.) $\square$

Proposition 19. *Let $\mathcal{T}$ be a model complete theory which admits a minimal model. Then $\mathcal{T}$ is complete.*

Proof. By Proposition 17, it suffices to show that any two models X_1 and X_2 of $\mathcal{T}$ are elementarily equivalent. But let I be a minimal model, so that there exist embeddings $\iota_1 : I \hookrightarrow X_1$ and $\iota_2 : I \hookrightarrow X_2$. By model completeness, both ι_1 and ι_2 are elementary embeddings, so $X_1 \equiv I \equiv X_2$. $\square$

Our next order of business is to give some of the classical examples of complete and/or model complete theories. The proofs of the completeness require some additional techniques and are deferred until the next chapter. However, we will use the completeness of these theories to derive some interesting theorems in mainstream mathematics.

3.3. **Algebraically closed fields I: model completeness.**

Theorem 20. *The theory of algebraically closed fields is model complete.*

We will prove this result in Chapter 5.

Corollary 21. *a) For any prime number p, the theory of algebraically closed fields of characteristic p is model complete and thus complete.*
b) The theory of algebraically closed fields of characteristic 0 *is model complete and thus complete.*

Exercise 3.3: Prove Corollary 21.

Remark: Corollary 21b) is one way to make precise the **Lefschetz principle** – i.e., that algebraic geometry is "the same" over any algebraically closed field of characteristic 0. A practicing algebraic geometer would recognize this as a rather anemic incarnation of the Lefschetz principle, and this is really the beginning of the story: there has been much work by model theorists on formulating stronger versions.

Example: For those who know some algebraic geometry and elliptic curve theory, here is a "nonexample" of the Lefschetz principle, i.e., a difference in the algebraic geometry of two different algebraically closed fields of the same characteristic. Let p be any prime number. Let E be an elliptic curve defined over $\overline{\mathbb{F}_p}$ (equivalently, over some finite field). Then the endomorphism algebra of E is either an imaginary quadratic field ("ordinary") or the definite rational quaternion algebra of discriminant p ("supersingular"). On the other hand, let $K = \overline{\mathbb{F}_p(t)}$ and let $E_{/K}$ be an elliptic curve with j-invariant t. Then the endomorphism algebra of E is $\mathbb{Q}$.

3.4. **Algebraically closed fields II: Nullstellensätze.**

Our first application of Theorem 21a) is the Nullstellensatz for algebraically closed fields. Owing to the way this was presented in the lectures, we do this in two steps: first a "weak" Nullstellensatz and then Hilbert's Nullstellensatz.

Theorem 22. *Let k be an algebraically closed field, and let $P_1, \ldots, P_k \in k[t] = k[t_1, \ldots, t_n]$ be a finite set of polynomials. Let*

$$V = V(P_1, \ldots, P_m) = \{x = (x_1, \ldots, x_n) \in k^n \mid P_1(x) = \ldots = P_m(x) = 0\}$$

be the locus of simultaneous zeros of the polynomials $P_1, \ldots, P_m$. TFAE:
(i) There exist polynomials $g_1(t), \ldots, g_m(t)$ such that

$$g_1(t)P_1(t) + \ldots + g_m(t)P_m(t) = 1. \tag{1}$$

(ii) $V = \emptyset$.

Proof. That (i) $\implies$ (ii) is obvious: if $x \in V$, then plugging x in to (1) yields $0 = 1$, a contradiction.
Now assume that (i) does not hold: equivalently, the ideal $I = \langle P_1, \ldots, P_k \rangle$ is a proper ideal of the polynomial ring $k[t]$. Therefore I is contained in a maximal ideal $\mathfrak{m}$, so $k[t]/\mathfrak{m}$ is a field, say K. Let $\overline{K}$ be an algebraic closure of K; then the composite map

$$k \to k[t] \to k[t]/\mathfrak{m} = K \to \overline{K}$$

gives an embedding $\iota : k \hookrightarrow \overline{K}$ of algebraically closed fields. By Theorem 20, ι is an elementary embedding, and this means precisely that the system of polynomial equations $P_1 = \ldots = P_m = 0$ has a solution over k iff it has a solution over $\overline{K}$. But the latter is a tautology: indeed, let x be the image of $(t_1, \ldots, t_n)$ in $k[t]/\mathfrak{m} = K$. Since each $P_i(t)$ lies in $\mathfrak{m}$, we have that $P_i(x) = 0$ for all i. In particular x is a common zero of the polynomials over $\overline{K}$. Done! $\square$

Exercise 3.4: Deduce from Theorem 22 the "weak Nullstellensatz": the maximal ideals in $k[t]$ are all of the form $\langle t_1 - a_1, \ldots, t_n - a_n \rangle$ for $(a_1, \ldots, a_n) \in k^n$.

It is well-known that the modern version of the Nullstellensatz – i.e., that the correpondences $I \mapsto V(I)$ and $V \mapsto I(V)$ between radical ideals and Zariski-closed subsets are mutually inverse bijections – follows easily from this by "Rabinowitch's trick": see e.g. `http://www.math.uga.edu/~pete/8320notes3.pdf`.

In fact we *can* give a simple model-theoretic proof of Hilbert's Nullstellensatz. For this we need a preliminary result in commutative algebra: literally, Theorem 1 in Kaplansky's *Commutative Rings*, which the author attributes to W. Krull.

Proposition 23. *(Multiplicative Avoidance) Let R be a commutative ring and $S \subset R$. Suppose:*
1 is in S; 0 is not in S; and S is closed under multiplication: $S \cdot S \subset S$.
Let $\mathcal{I}_S$ be the set of ideals of R which are disjoint from S. Then:
a) $\mathcal{I}_S$ is nonempty;.
b) Every element of $\mathcal{I}_S$ is contained in a maximal element of $\mathcal{I}_S$.
c) Every maximal element of $\mathcal{I}_S$ is prime.

Proof. a) $(0) \in \mathcal{I}_S$. b) Let $I \in \mathcal{I}_S$. Consider the subposet P_I of $\mathcal{I}_S$ consisting of ideals which contain I. Since $I \in P_I$, P_I is nonempty; moreover, any chain in P_I has an upper bound, namely the union of all of its elements. Therefore by Zorn's Lemma, P_I has a maximal element, which is clearly also a maximal element of $\mathcal{I}_S$.
c) Let I be a maximal element of $\mathcal{I}_S$; suppose that $x, y \in R$ are such that $xy \in I$. If x is not in I, then $\langle I, x \rangle$ contains an element s_1 of S, say

$$s_1 = i_1 + ax.$$

Similarly, if y is not in I, then we get an element s_2 of S of the form

$$s_2 = i_2 + by.$$

But then

$$s_1 s_2 = i_1 i_2 + (by) i_1 + (ax) i_2 + (ab) xy \in I \cap S,$$

a contradiction. □

Let k be an arbitrary field, $n \in \mathbb{Z}^+$, and put $R = k[t] = k[t_1, \ldots, t_n]$. For an ideal J of R, put

$$V(J) = \{x = (x_1, \ldots, x_n) \in k^n \mid \forall f \in R \ f(x) = 0\},$$

i.e., the set of simultaneous zeros of all polynomials in J. On the other hand, for $S \subset k^n$, put

$$I(S) = \{f \in R \mid \forall x \in S \ f(x) = 0\}.$$

Then the pair (V, I) is a Galois connection between the set of ideals of R and the set of subsets of k^n (both partially ordered by inclusion).[13] We have associated closure operators $S \mapsto \overline{S} := V(I(S))$ and $J \mapsto \overline{J} = I(V(J))$. The closure operator on subsets also has the following property: for all $S, T \subset k^n$, $\overline{S \cup T} = \overline{S} \cup \overline{T}$. It is therefore a **Kuratowski closure operator**, i.e., the closure operator for a unique topology on k^n, called the **Zariski topology**. In this topology, the closed sets are precisely the zero loci of families of polynomials. Since every ideal of R is finitely generated (Hilbert Basis Theorem), the zero locus of an *a priori* infinite family of polynomials can be rewritten as the zero locus of some finite subfamily.[14]

On the other hand, an explicit description of the closure operator $I \mapsto \overline{I}$ on ideals depends very much on the structure of the ground field k. An explicit description of this closure operator for a field k will be called a **Nullstellensatz** over k.

Exercise 3.5: Let k be any field, and let I be any ideal in $k[t_1, \ldots, t_n]$. Show that $\overline{I}$ contains $\mathrm{rad}(I) = \{x \in R \mid \exists n \in \mathbb{Z}^+ \ x^n \in I\}$.

Theorem 24. *(The Nullstellensatz for Algebraically Closed Fields) Let k an algebraically closed field and I an ideal of $R = k[t_1, \ldots, t_n]$. Then $\overline{I} = \mathrm{rad}(I)$.*

Proof. Let $I = \langle f_1, \ldots, f_m \rangle$ be an ideal of R. By Exercise 3.5, $\overline{I} \supset \mathrm{rad}(I)$. Seeking a contradiction, suppose that there exists $g \in \overline{I} \setminus \mathrm{rad}(I)$. Then the multiplicative subset $S = \{g^n\}_{n=1}^{\infty}$ is disjoint from I. By Proposition 23, there exists a prime

[13] We hope to provide a handout with more details: stay tuned.

[14] Yes, another compactness theorem.

ideal $\mathfrak{p}$ containing I and disjoint from S. Let K be the fraction field of $R/\mathfrak{p}$ and let $\overline{K}$ be an algebraic closure of K. Consider the $\mathcal{L}_k$-sentence

$$\exists x(\bigwedge_{i=1}^{m} f_i(x) = 0) \wedge (g(x) \neq 0).$$

Taking x to be the image of $t = (t_1, \ldots, t_n)$ in K, this sentence is true in K; since it is an existential sentence (i.e., the only quantifiers appearing are existential quantifiers), it is also true in $\overline{K}$. By model completeness of ACF, the embedding $k \hookrightarrow \overline{K}$ is elementary, so that the sentence is also true in k, and the existence of such an $x \in k^n$ shows that $g \notin \overline{I}$. This contradiction completes the proof. $\square$

Remark: We could attain a slightly shorter proof of Theorem 24 by replacing Proposition 23 with the following fact: for any ideal I in a commutative ring, $\mathrm{rad}(I)$ is the intersection of all the prime ideals containing it. In particular, if $I \subset J$ is a proper inclusion of radical ideals, then there exists a prime ideal $\mathfrak{p}$ such that $I \subset \mathfrak{p}$ and there exists $g \in J \setminus \mathfrak{p}$. However, since a similar argument will be given later to prove the Nullstellensatz over a real-closed field, we chose to present this slightly more elementary argument for the sake of variety.

Exercise 3.6: Prove the converse of Theorem 24: let k be a field, and suppose that for all $n \in \mathbb{Z}^+$ and all ideals I of $k[t_1, \ldots, t_n]$, $\overline{I} = \mathrm{rad}(I)$. Show that k is algebraically closed.

Exercise 3.7: Find all fields k with the following property: for each $n \in \mathbb{Z}^+$ and $R = k[t_1, \ldots, t_n]$, the zero ideal in R is closed: $I(V(\{0\})) = \{0\}$.

3.5. **Algebraically closed fields III: Ax's Transfer Principle.**

Theorem 25. *(Ax's Transfer Principle)*
For a sentence φ in the language of fields, TFAE:
(i) φ is true in $\mathbb{C}$, the complex field.
(ii) φ is true in every algebraically closed field of characteristic 0.
(iii) There exists a constant N such that for any algebraically closed field K of characteristic $p > N$, φ is true in K.
(iv) There are infinitely many primes p such that φ is true in every algebraically closed field of characteristic p.

Proof. (i) $\equiv$ (ii) follows from Corollary X.Xa). Suppose (ii), so that $\mathrm{ACF}_0 \models \varphi$. By Gödel's Completeness Theorem, there exists a finite subset $\mathcal{T}$ of ACF_0 such that $\mathcal{T} \models \varphi$. Since $\mathcal{T} \subset \mathrm{ACF}_p$ for all but finitely many prime numbers p, (iii) follows. Trivially (iii) $\implies$ (iv). Suppose (iv) holds and that it is not the case that $\mathrm{ACF}_0 \models \varphi$. By completenesss, $\mathrm{ACF}_0 \models \neg\varphi$ and then by the above work $\mathrm{ACF}_P \models \neg\varphi$ for all but finitely many primes, contradiction. $\square$

Corollary 26. *(Grothendieck, Ax) Let $f_1, \ldots, f_n \in \mathbb{C}[t_1, \ldots, t_n]$ be polynomials, and let $f = (f_1, \ldots, f_n) : \mathbb{C}^n \to \mathbb{C}^n$ be a map. If f is injective, then it is surjective.*

Proof. For each fixed $n \in \mathbb{Z}^+$, the statement is expressible as a sentence in the language of fields. Therefore, by Theorem 25, it is equivalent to prove that for all primes p, an injective polynomial map from $\overline{F}_p^n$ to $\overline{F}_p^n$ is surjective. Let $\mathbb{F}_q$ be an extension of $\mathbb{F}_p$ containing all of the coefficients of all the polynomials. Then P

also induces a polynomial map from $(\mathbb{F}_q)^n \to (\mathbb{F}_q)^n$ and, indeed, for all $a \in \mathbb{Z}^+$, polynomial maps $P_a : (\mathbb{F}_{q^a})^n \to (\mathbb{F}_{q^a})^n$ for all $a \in \mathbb{Z}^+$. A moment's thought shows that the injectivity (resp. surjectivity) of P is equivalent to the injectivity (resp. surjectivity) of P_a for all a. Therefore we may assume that, for all $a \in \mathbb{Z}^+$, P_a is injective. But P_a is a map from a finite set to itself, so, being injective, it must also be surjective. Done! □

Remark: The reader with a background in algebraic geometry may enjoy extending the statement (and proof) of Corollary 26 to regular maps of algebraic varieties.

Remarks: Corollary 26 was first by A. Grothendieck in EGAIII circa 1966 and independently by J. Ax [Ax68]. Ax's proof is indeed the one we have given here, namely as a corollary to Theorem 25 (also due to Ax).

A very good exercise for the reader who is not sure what to make of model-theoretic methods is to try to prove Corollary 26 from scratch. This is indeed possible: leaving aside Grothendieck's proof, other proofs were given by Borel [Bor69], Rudin [Rud], Brian Conrad (unpublished, I believe) and Serre [Ser09]. (Serre's recent article is especially highly recommended, as he essentially "transfers" the transfer argument from model theory to conventional algebra.) Certainly there is no argument that if we are given Theorem 25 (which, to be fair, we point out that we have not yet proven) one can scarcely imagine a simpler proof of Corollary 26. Moreover Theorem 25 is a very general result which builds a bridge between algebraic geometry of characteristic 0 and algebraic geometry in positive characteristic.

3.6. **Ordered fields and formally real fields I: background.**

The structure of real-closed fields is a shining example of the merits of extension of language. This is because, even in the non-model theoretic study of real-closed fields, it is natural to consider them as two different types of structures: as a certain kind of field, and as a certain kind of ordered field. We briefly recall the basic definitions. For more details, the reader may consult (e.g.) [FT, Ch. 11].

A field K is **formally real** if for no $n \in \mathbb{Z}^+$ is it possible to express -1 as a sum of n squares in the field. (Thus this is a first-order property which can be axiomatized by infintely many sentences.) A field is **real-closed** if it is formally real but admits no proper formally real algebraic extension. For instance, the real numbers are formally real. The main result on formally real fields is the following celebrated theorem.

Theorem 27. *(Grand Artin-Schreier Theorem) For a field F, TFAE:*
(i) F is formally real and admits no proper formally real algebraic extension.
(ii) F is formally real, every odd degree polynomial over F has a root, and for each $x \in F^\times$, exactly one of x, $-x$ is a square.
(iii) F is formally real and $F(\sqrt{-1})$ is algebraically closed.
(iv) The absolute Galois group of F is finite and nontrivial.

In particular, property (ii) is equivalent to an infinite union of sentences in the language of fields, so the class of real-closed fields is again first-order.

The "Little Artin-Schreier Theorem" asserts that a field admits a compatible ordering (as defined above) iff it is formally real. One direction is easy: directly from the axioms one sees that in an ordered field, -1 is negative and every sum of squares is non-negative, so certainly an orderable field is formally real. In particular, orderable fields have characteristic 0. So the following result is a strong form of the Little Artin-Schreier Theorem.

Theorem 28. *(Artin) Let F be a a field with* $\operatorname{char}(F) \neq 2$ *and $x \in F$. TFAE:*
a) For every ordering $<$ on F, $x > 0$.
b) x is a sum of squares of elements of F.

Proof. See e.g. [FT, Cor. 94]. □

(Just to be sure: this implies the little Artin-Schreier theorem because if F is formally real, -1 is not a sum of squares, so -1 is not positive for every ordering on F, so there must be at least one ordering on F!)

3.7. **Ordered fields and formally real fields II: the real spectrum.**

So every formally real field may be given the structure of an ordered field. In general, there are many different orderings on a formally real field. For instance, for a number field, the orderings correspond to the real roots of a minimal polynomial on F. In fact, the set of all orderings of a field can be topologized as follows: for $a \in F^{\times}$, define $H(a)$ to be the set of all orderings on F with respect to which a is positive. Then there is a unique topology on the space $\operatorname{RSpec}(F)$ of orderings of F for which $\{H(a) \mid a \in F\}$ is a subbase: i.e., for which the open sets are finite intersections of the sets $H(a)$. This is called the **Harrison topology** and the sets $H(a)$ are called the **Harrison subbase**.

Theorem 29. *Let F be a field. Then the space* $\operatorname{RSpec}(F)$ *of orderings of F endowed with the Harrison topology is a Stone space: i.e., it is compact (Hausdorff!) and totally disconnected.*

Exercise 3.8: This exercise leads you through a proof of the compactness of $\operatorname{RSpec}(F)$.
a) Show that for all $a \in F^{\times}$, $\operatorname{RSpec}(F) = H(a) \coprod H(-a)$.
b) If P, P' are distinct orderings on F, then there exists $a \in F$ such that $a \in P$, $-a \in P'$.
c) Deduce from a) and b) that $\operatorname{RSpec}(F)$ is Hausdorff and totally disconnected.
d) Define an injection $\iota : \operatorname{RSpec}(F) \hookrightarrow \{0,1\}^{F^{\times}}$. (Hint: to each ordering P, associate the subset of positive elements.)
e) Endow $\{0,1\}$ with the discrete topology and $\{0,1\}^{F^{\times}}$ with the product topology. Show that it is a Stone space. Show moreover that the map $\iota : \operatorname{RSpec}(F) \to \iota(\operatorname{RSpec}(F))$ is a homeomorphism: i.e., the topology that $\operatorname{RSpec}(F)$ inherits from its embedding into $\{0,1\}^{F^{\times}}$ coincides with the Harrison topology.
f) Show that $\iota(\operatorname{RSpec}(F))$ is closed in $\{0,1\}^{F^{\times}}$. Deduce that $\operatorname{RSpec}(F)$ is compact.

Exercise 3.9: Define a relation R on $F \times \operatorname{RSpec}(F)$ by $(x, P) \in R$ if $x \in P$.
a) Show that the corresponding closure operator on $\operatorname{RSpec}(F)$ is precisely the closure with respect to the Harrison topology.
b) What can you say about the closure operator on subsets of F?

3.8. **Real-closed fields I: definition and model completeness.**

On the other hand, any real-closed field admits a *unique* ordering: indeed, by Theorem 27(ii) in a real-closed field F, any nonzero element x is either a square – hence necessarily positive – or minus a square – hence necessarily negative – and not both.

Moreover, one has the notion of the real-closure of an **ordered-field** $(F, <)$: this is a real-closed algebraic extension $\mathcal{R}$ of F such that the unique ordering on $\mathcal{R}$ restricts to the given ordering $<$ on F.

Theorem 30. *Let $(F, <)$ be an ordered field.*
a) Then $(F, <)$ admits a real-closure as an ordered field.
b) Any two real-closures of $(F, <)$ are isomorphic over F.

Proof. For a proof of the existence of a real-closure of an ordered field, see e.g. [FT, Thm. 103]. The uniqueness is considerably harder; for that see [Mar, Appendix B] or [Lam]. □

Exercise 3.10: Let $\iota : \overline{\mathbb{Q}} \hookrightarrow \mathbb{C}$ be an embedding.[15] Define $\mathcal{R}_0 = \iota(\overline{\mathbb{Q}}) \cap \mathbb{R}$ to be the field of real algebraic numbers.
a) Show that $\mathcal{R}_0$ is a real-closed field. (Suggestion: use Theorem 27(iii).)
b) Show that $\mathcal{R}_0$ embeds in each real-closed field. (Hint: the field $\mathbb{Q}$ admits a unique ordering.) c) Deduce that the theory of real-closed fields admits minimal models.

So any real-closed field has the canonical structure of an ordered field. Now a key observation: the theory of real-closed fields in the language of fields is model complete iff the theory of real-closed fields in the language of ordered fields is model-complete. This is because every embedding of real-closed fields in the category of fields promotes uniquely to an embdding of real-closed fields in the category of ordered fields, and conversely.

Theorem 31. *(Tarski* [Tar48]*) The theory of real closed fields is model complete.*

We will prove Theorem 31 in Chapter 5. In fact, the proof of this theorem is the main goal of that chapter.

Corollary 32. *The theory of real closed fields is complete.*

Exercise 3.11: Prove Corolllary 32.

3.9. **Real-closed fields II: Nullstellensatz.**

As our first application, we will prove a Nullstellensatz for real-closed fields.

First some motivation: let us determine which prime ideals $\mathfrak{p}$ of $\mathbb{R}[t]$ are closed – i.e., have the property that if a polynomial g vanishes at every point of simultaneous vanishing of the elements of $\mathfrak{p}$, then $g \in \mathfrak{p}$. We leave it to the reader to check that the zero ideal is closed (a special case of Exercise 3.5. The nonzero prime ideals of $\mathbb{R}[t]$ are all principal and in bijective correspondence with monic irreducible polynomials $p(t)$. Of course these come in two flavors: linear polynomials and irreducible quadratic polynomials. Let $a \in \mathbb{R}$, and put $I = (t - a)$. Then

[15]Depending upon your religious convictions, you may or may not believe that the complex numbers come equipped with a standard such embedding, but of course such embeddings exist.

$V(I) = \{a\}$ and $I(V(I))$ is the set of all polynomials vanishing at a. Of course, if $g(a) = 0$, then $t - a \mid g$, so this ideal is indeed $(t - a)$: I is closed. On the other hand, if $I = (t^2 + bt + c)$ is the principal ideal generated by an irreducible quadratic, then $V(I) = \emptyset$ and hence $I(V(I)) = \mathbb{R}[t]$ – every polynomial vanishes at every point of the empty set! These ideals are not closed.

This example suggests that we need a notion of a "real ideal". The correct one for this context is the following.

An ideal I of a commutative ring is **real** if for all $n \in \mathbb{Z}^+$ and $x_1, \ldots, x_n \in R$, $x_1^2 + \ldots + x_n^2 \in I \implies x_1, \ldots, x_n \in I$.

Exercise 3.12: Let I be a real ideal in a commutative ring. Show that I is radical, i.e., $I = \operatorname{rad}(I)$. (Suggestion: suppose $x^n \in I$. If n is even then $x^{\frac{n}{2}} \in I$, whereas if n is odd then $x^{\frac{n+1}{2}} \in I$.)

Exercise 3.13: Let $\{I_i\}$ be a family of real ideals in a commutative ring R. Show that $I = \bigcap_i I_i$ is also a real ideal.

A commutative ring R is real if the zero ideal is real.

Exercise 3.14: Show that an integral domain is real iff its fraction field is formally real.

Proposition 33. *Let k be a formally real field and $R = k[t] = k[t_1, \ldots, t_n]$. If an ideal J of R is closed, then J is real.*

Proof. Let J be a closed ideal of R and let $f_1, \ldots, f_m \in k[t]$ be such that $f_1^2 + \ldots + f_m^2 \in J$. Fix $x \in V(J)$. Then $f_1(x)^2 + \ldots + f_m(x)^2 = 0$; since k is formally real, this implies $f_1(x) = \ldots = f_m(x) = 0$. So for all $1 \leq i \leq m$, $f_i \in I(V(J)) = J$. $\square$

Exercise 3.15: Find a real prime ideal $\mathfrak{p} \in \mathbb{Q}[t]$ which is not closed.

However, if k is real-closed, the converse of Proposition 33 holds: every real ideal is closed. We want a little more than this, namely to identify the closure operator on an arbitrary ideal of $k[t]$.

For an ideal I of a ring R, define the **real radical**

$$\mathbb{R}\operatorname{rad}(I) = \{x \in R \mid \exists n \in \mathbb{Z}^+ \ \exists b_1, \ldots, b_m \in R \mid a^{2n} + b_1^2 + \ldots + b_m^2 \in I\}.$$

The following result is the key piece of algebraic information we need to adapt our previous model-theoretic arguments: it is the "real algebraic" analogue of the fact that the radical of an ideal in a commutative ring is the intersection of the prime ideals containing it.

Proposition 34. [BCR, Prop. 4.1.7] *Let I be an ideal in a commutative ring R.*
a) A real ideal J contains I iff $J \supset \mathbb{R}\operatorname{rad}(I)$ i.e., $\mathbb{R}\operatorname{rad}(I)$ is the unique minimal real ideal containing I.
b) $\mathbb{R}\operatorname{rad}(I)$ is equal to the intersection of all real prime ideals $\mathfrak{p} \supset I$.
c) It follows that every real ideal is equal to the intersection of all the real prime ideals containing it.

Remark: If there are no real prime ideals containing I, then the intersection over this empty set is taken to be R.

Proof. Step 1: we show that $\mathbb{R}\operatorname{rad}(I)$ is an ideal. The only nonobvious part of this is closure under addition. Suppose that

$$a^{2n} + b_1^2 + \ldots + b_m^2,\ A^{2N} + B_1^2 + \ldots + B_M^2 \in I.$$

We may write

$$(a + A)^{2(n+N)} + (a - A)^{2(n+N)} = a^{2m}c + A^{2M}C,$$

with c, C sums of squares in R. Then

$$(a + A)^{2(n+N)} + (a - A)^{2(n+N)} + c(b_1^2 + \ldots + b_m^2) + C(B_1^2 + \ldots + B_M^2) \in I,$$

so $a + A \in \mathbb{R}\operatorname{rad}(I)$.
Step 2: $\mathbb{R}\operatorname{rad}(I)$ is a real ideal. Indeed, if $x_1^2 + \ldots + x_k^2 \in \mathbb{R}\operatorname{rad}(I)$, then there exists $n \in \mathbb{Z}^+$ and $b_1, \ldots, b_m \in R$ such that

$$(x_1^2 + \ldots + x_k^2)^{2m} + b_1^2 + \ldots + b_m^2 \in I;$$

for each $1 \leq i \leq k$, we may rewrite this expression as $x_i^{4m} + B_1^2 + \ldots + B_N^2$, so $x_i \in \mathbb{R}\operatorname{rad}(I)$.
Step 3: Since every real ideal is radical, it is clear that any real ideal containing I also contains $\mathbb{R}\operatorname{rad}(I)$.
Step 4: Let $a \in R \setminus \mathbb{R}\operatorname{rad}(I)$. By Zorn's Lemma, the set of real ideals containing I but not a has a maximal element, say J. We claim that J is prime. If not, there exist $b, b' \in R \setminus J$ such that $bb' \in J$. Then $a \in \mathbb{R}\operatorname{rad}(J + bR)$ and $a \in \mathbb{R}\operatorname{rad}(J + b'R)$, hence there are $j, j' \in J$ such that

$$a^{2m} + c_1^2 + \ldots + c_q^2 = j + bd,\ a^{2m'} + c_1'^2 + \ldots + c_q'^2 = j' + b'd'.$$

It follows that

$$a^{2(m+m')} + \text{ a sum of squares } = jj' + jb'd' + j'bd + bb'dd' \in J,$$

and thus $a \in \mathbb{R}\operatorname{rad}(J) = J$, contradiction. Thus $\mathbb{R}\operatorname{rad}(I)$ is the intersection of all real prime ideals containing I. $\square$

Theorem 35. *(The Nullstellensatz for Real-Closed Fields) Let k be a real-closed field, and I an ideal in $k[t_1, \ldots, t_n]$. Then $\overline{I} = \mathbb{R}\operatorname{rad}(I)$.*

Proof. Let $I = \langle f_1, \ldots, f_m \rangle$ be an ideal of R. By Propositions 33 and 34, $\overline{I}$ is a real ideal containing I, so $\overline{I} \supset \mathbb{R}\operatorname{rad}(I)$.

Seeking a contradiction, suppose $\mathbb{R}\operatorname{rad}(I) \subset \overline{I}$ is a proper inclusion of real ideals. Then by Proposition 34c) there exists a real prime ideal $\mathfrak{p}$ containing $\mathbb{R}\operatorname{rad}(I)$ but not $\overline{I}$; in particular, there exists $g \in \overline{I} \setminus \mathfrak{p}$. We may now proceed exactly as in the proof of Theorem 24: let K be the fraction field of $R/\mathfrak{p}$ and let K^{rc} be a real closure of K. Consider the $\mathcal{L}_k$-sentence

$$\exists x (\bigwedge_{i=1}^{m} f_i(x) = 0) \wedge (g(x) \neq 0).$$

Taking x to be the image of $t = (t_1, \ldots, t_n)$ in K, this sentence is true in K and hence also in K^{rc}. By model completeness of RCF, the embedding $k \hookrightarrow K^{\mathrm{rc}}$ is elementary, so that the sentence is also true in k, and the existence of such an $x \in k^n$ shows that $g \notin \overline{I}$. This contradiction completes the proof. $\square$

3.10. **Real-closed fields III: Hilbert's 17th problem.**

Hilbert's 17th problem asked if every positive semidefinite polynomial with $\mathbb{R}$-coefficients was a sum of squares of rational functions.

Exercise 3.16: Show that Hilbert's 17th problem has an affirmative answer in the case of rational functions of a single variable.[16]

Theorem 36. *(Artin) Let F be a real-closed field and $f \in F(t) = F(t_1, \ldots, t_n)$ be a positive semidefinite rational function. Then there exist $g_1, \ldots, g_m \in F(x)$ such that $f = g_1^2 + \ldots + g_m^2$.*

Proof. Seeking a contradiction, let f be a positive semidefinite rational function which is *not* a sum of squares. By [FT, Cor. 94], there exists an ordering $<$ on $F(t)$ such that $f < 0$. Let $\mathcal{R}$ be the real-closure of $(F, <)$. Then the elements $t_1, \ldots, t_n$ all lie in $\mathcal{R}$, so that we may think of the field element f as the rational function f with F-coefficients evaluated at the element $t = (t_1, \ldots, t_n)$ of $\mathcal{R}$ and thus $f = f(t) < 0$.[17] In other words,

$$\mathcal{R} \models \exists x \ f(x) < 0.$$

But by the model-completeness of RCF, $F \hookrightarrow \mathcal{R}$ is an elementary embedding, hence also

$$F \models \exists x \ f(x) < 0,$$

i.e., there exists $x = (x_1, \ldots, x_n) \in F$ such that $f(x) < 0$. Contradiction! $\square$

Theorem 36 was first proved by Emil Artin in 1927. Artin's original proof was not model-theoretic but rather used more of the theory of formally real fields that he had developed with Schreier. Nowadays we view Artin's technique of proof as the beginning of a branch of mathematics called **real algebraic geometry**. The above model-theoretic proof is due to Abraham Robinson [ARob55]. The subjects of model theory and real algebraic geometry have been closely connected ever since.

4. Categoricity: a condition for completeness

By Löwenheim-Skolem, any theory in a countable language which admits infinite models admits models of every infinite cardinality, and indeed, models of any given cardinality elementarily equivalent to any fixed infinite model. Thus the next step in understanding the relation of elementary e equivalence is to consider models of a fixed cardinality. In this regard, the following definition captures the simplest possible state of affairs.

Let κ be an infinite cardinal. A theory $\mathcal{T}$ is κ-categorical if there exists a unique (up to isomorphism) model of cardinality κ.

Categoricity leads to completeness as follows:

---

[16] This is sort of a function field analogue of Fermat's Two Squares Theorem.

[17] This step of the proof was confusing to me when I was first learning the subject. We are thinking of f at the same time as an element of the abstract field $\mathcal{R}(t)$ and as a rational function with $\mathcal{R}$-coefficients evaluated at the "generic" element $(t_1, \ldots, t_n)$. This makes perfect sense, but it may take some getting used to.

Theorem 37. *(Vaught's Test) Let $\mathcal{T}$ be a satisfiable theory with no finite models which is κ-categorical for some $\kappa \geq |\mathcal{L}|$. Then $\mathcal{T}$ is complete.*

Proof. Suppose $\mathcal{T}$ is not complete, and let φ be a sentence such that $\mathcal{T} \not\models \varphi$ and $\mathcal{T} \not\models \neg\phi$. Then the extended theories $\mathcal{T}_1 := \mathcal{T} \cup \neg\phi$ and $\mathcal{T}_2 := \mathcal{T} \cup \phi$ are both satisfiable. Since they do not admit finite models, they both admit infinite models. By Löweneim-Skolem, each $\mathcal{T}_i$ admits a model X_i of cardinality κ. But X_1 and X_2 disagree about the truth of φ, so they are not even elementarily equivalent – let alone isomorphic – contradicting the κ-categoricity of $\mathcal{T}$. □

Exercise 4.1: For a theory $\mathcal{T}$, let $\mathcal{T}_\infty$ be the theory of infinite models of $\mathcal{T}$, i.e., $\mathcal{T}$ augmented with the infinite family of sentences φ_n, each φ_n expressing that the structure has at least n distinct elements. Prove the following variation of Vaught's Test: let $\mathcal{T}$ be a theory admitting an infinite model which is κ-categorical for some $\kappa \geq \max(\aleph_0, |\mathcal{L}|)$. Then $\mathcal{T}_\infty$ is complete. Immediately after seeing the proof, A. Brunyate pointed out the following strengthening.

Theorem 38. *(Brunyate's Test) Let $\mathcal{T}$ be a satisfiable theory without finite models. Suppose that there exists an infinite cardinal $\kappa \geq |\mathcal{L}|$ such that any two models of $\mathcal{T}$ of cardinality κ are elementarily equivalent. Then $\mathcal{T}$ is complete.*

Exercise 4.2: Prove Brunyate's Test, and also its analogue for $\mathcal{T}_\infty$ as in Exercise 4.1.

One may ask why we use Vaught's Test and not Brunyate's Test since the latter is plainly stronger. Indeed, *every* complete theory satisfies Brunyate's Test. The answer, I believe, is that the hypothesis of Brunyate's Test is model-theoretic in nature, whereas the (stronger!) hypothesis of Vaught's test belongs to mainstream mathematics. Therefore in certain elementary instances we essentially already know that the hypothesis of Vaught's test is satisfied and stating it as a theorem is a clue to keep one's eye open for κ-categorical theories.

We now give some examples of the successful application of Vaught's test.

Proposition 39. *Let $\mathcal{L}$ be the empty language – i.e., the language of naked sets. Let X and Y be $\mathcal{L}$-structures. TFAE:*
(i) Either X and Y are both infinite, or X and Y are both finite with $|X| = |Y|$.
(ii) $X \equiv Y$.

Exercise 4.3: Prove Proposition 39.

Exercise 4.4: Let $\mathcal{L}$ be the language with a single constant symbol, so $\mathcal{L}$-structures are pointed sets. Classify $\mathcal{L}$-structures up to elementary equivalence.

Exercise 4.5: Let $\mathcal{L}$ be the language with a single unary relation, so $\mathcal{L}$-structures are pairs (X, Y) with $Y \subset X$. Try to classify $\mathcal{L}$-structures up to elementary equivalence.

Theorem 40. *The theories ACF_0 and ACF_p (for any prime $p \geq 0$ are each κ-categorical for any uncountable cardinal κ. None of these theories admit finite models, so by Vaught's test they are all complete.*

Proof. In other words, we claim that if K_1 and K_2 are algebraically closed fields of the same characteristic and the same uncountable cardinality, then they are isomorphic. This is a true fact of field theory, a consequence of the following more precise result: two algebraically closed fields are isomorphic iff they have the same characteristic and the same absolute transcendence degree (i.e., the transcendence degree over their prime subfield). But for an uncountable field, the transcendence degree is equal to the cardinality. □

Remark: The theories ACF_0 and ACF_p are *not* $\aleph_0$-categorical: for countable fields, the absolute transcendence degree is an extra invariant. This provides our first example of two structures of the same cardinality which are elementarily equivalent but not isomorphic: say $\overline{\mathbb{Q}}$ and $\overline{\mathbb{Q}(t)}$.

At a deeper level, algebraic geometers have long known that – Lefschetz principle notwithstanding! – a countable algebraically closed field of larger transcendence degree is a "richer" object than $\overline{Q}$. For instance, not every complex algebraic variety may be defined over $\overline{\mathbb{Q}}$. Among countable models of ACF_0, the "richest" – indeed, the maximal one with respect to embeddings – is clearly the one of countably infinite transcendence degree. Such fields played a fundamental role in Weil's formalization of algebraic geometry via *universal domains*. Although this notion is now somewhere between out of fashion and completely forgotten by contemporary algebraic geometers, it is well appreciated by model theorists, being an instance of the notion of a **saturated model** (which I think we will not get to in this course).

4.1. **DLO.**

Recall that DLO (dense linear orders) is the theory (well-defined up to syntactic closure) in the language $\mathcal{L} = \{<\}$ consisting of one binary relation whose models are precisely the nonempty linearly ordered sets without endpoints and for which the order relation is dense: for all $x < y$, there exists z with $x < z < y$.

Exercise 4.6: a) Show that DLO does not admit finite models.
b) Let $(F, <)$ be an ordered field. Show that the underlying ordered set is a DLO.

Theorem 41. *The theory DLO of dense linear orders without endpoints is $\aleph_0$-categorical. Thus – by Exercise 4.6a) and Vaught's Test, DLO is complete.*

Proof. The proof is by what is called a **back and forth** argument. Let $X = \{x_n\}_{n=1}^{\infty}$ and $Y = \{y_n\}_{n=1}^{\infty}$ be countable DLOs. We will build up an order-preserving bijection from X to Y via a sequence of countable steps. At Step $2n-1$, we will ensure that x_n is in the domain of the bijection, and at Step $2n$, we will ensure that y_n is in the codomain of the bijection. If we can do this, we're done!
Step 1: Take x_1 and map it to any element of Y.
Step 2: If y_1 is already in the codomain of f_1, we do nothing. If not, we choose an element x of X and map x to y_1. We choose x such that $x < x_1$ if $y_1 < f(x_1)$ and $x > x_1$ if $y_1 > f(x_1)$.
Step 3: If x_2 is already in the domain of f_2, we do nothing. If not, we choose an element y of Y and map x_2 to y. We do this in such a way to preserve the extant order relations: the elements in the order-preserving bijection f_2 split up both X and Y into finitely many intervals, each of which is nonempty by definition of DLO. So we need only choose y lying in the correspodning interval to x_2.
We continue in this manner. A little thought shows that this strategy succeeds. □

Exericse 4.6: a) Show that DLO is not $2^{\aleph_0}$-categorical. (Suggestion: compare $\mathbb{R}$ with its canonical ordering to the *ordered sum* $\mathbb{R} + \mathbb{Q}$: i.e., we place a copy of $\mathbb{Q}$ "on top of" $\mathbb{R}$ such that every element of $\mathbb{Q}$ is greater than every element of $\mathbb{R}$.)
b)(harder) Show that DLO is not κ-categorical for any uncountable κ.

Exercise 4.7: Let DLOE be the theory of dense linear orders with largest and smallest elements.
a) Show that DLOE is $\aleph_0$-categorical. (Hint: use either the statement or the proof of Theorem 41.)
b) Apply Vaught's Test to show that DLOE is a complete theory.
c) Show that DLOE is not model complete.[18]

The method of proof of Theorem 41 is probably more important than the result itself. The construction of an isomorphism, or elementary embedding, by **back and forth** turns out to be one of the most fundamental notions in model theory. Indeed, in Bruno Poizat's (somewhat idiosyncratic, but extremely insightful) introductory text [Poi], the concept of back-and-forth is taken as a primitive and elementary equivalence is defined in terms of it.

4.2. R-modules.

In this section we provide a glimpse of the model-theoretic study of modules over a (not necessarily commutative) ring. While perhaps not as sexy as the model theory of fields, this is nevertheless an active research area at the border of model theory and algebra.

As motivation, we provide two examples of complete theories in the language of groups. First some (standard) terminology.

Let G be a group. The **exponent** $E(G)$ is the least positive integer E such that for all $g \in G$, $g^E = e$ – equivalently, the least common multiple of all orders of elements of G. (If no such integer exists, we say that the exponent is ∞.) For instance, if G is finite, then by Lagrange's Theorem $E(G) \mid |G|$.

Moreover, for every $n \in \mathbb{Z}^+$, we have a map $[n] : G \to G$, $g \mapsto g^n$. (Note that $[n]$ need not be a group homomorphism. Indeed $[2]$ is a homomorphism iff G is commutative iff $[n]$ is a homomorphism for all $n \in \mathbb{Z}^+$.) We say that G is **torsionfree** if for all $n \in \mathbb{Z}^+$ and all $g \in G$, $[n]g = e \iff g = e$.[19] G is **divisible** if each $[n]$ is surjective and **uniquely divisible** if each $[n]$ is bijective.

Theorem 42. *Let $\mathcal{L} = \{+, -, 0\}$ be the language of commutative monoids. All of the following $\mathcal{L}$-theories are complete:*
(i) For any prime p, the theory of infinite commutative groups of exponent p.
(ii) The theory of nontrivial uniquely divisible abelian groups.

[18] This is a standard example of a complete but not model complete theory. In fact, it is the simplest one I know, although it requires some machinery to show this. If you can think of a more elementary example of a complete but not model complete theory, please let me know!

[19] If G is commutative, this is equivalent to saying that each $[n]$ is injective. For noncommutative G, the latter condition is *a priori* stronger, but I don't have an example to confirm that it is strictly stronger.

Proof. Each of the theories admits only infinite models, so it enough to show that these theories are κ-categorical for some infinite cardinal and then apply Vaught's test.

A commutative group of exponent p has, in a unique way, the structure of an $\mathbb{F}_p$-vector space, and conversely the additive group of any nontrivial $\mathbb{F}_p$-vector space is a commutative group of exponent p. The only invariant of an $\mathbb{F}_p$-vector space is its dimension, and for any infinite $\mathbb{F}_p$-vector space V, its dimension is simply equal to its cardinality (c.f. Lemma 43). Therefore the theory of infinite commutative groups of exponent p is κ-categorical for all infinite κ.

Similarly, a uniquely divisible abelian group has, in a unique way, the structure of a $\mathbb{Q}$-vector space, and conversely the additive group of any $\mathbb{Q}$-vector space is a commutative, uniquely divisible abelian group. The only invariant of a $\mathbb{Q}$-vector space is its dimension, and for any uncountable $\mathbb{Q}$-vector space V, its dimension is simply equal to its cardinality (c.f. Lemma 43). Therefore the theory of nontrivial uniquely divisible commutative groups is κ-categorical for all uncountble κ. $\square$

The following result nails down the relation between the dimension of a vector space and its cardinality, special cases of which were used in the above proof.

Lemma 43. *Let F be a field and V a nontrivial vector space over F. Then*

$$|V| = \max(|F|, \dim(V)).$$

Exercise 4.8: Prove Lemma 43.

These examples suggest a common generalization in terms of vector spaces over a field. However, it is somewhat "lucky" that vector spaces over $\mathbb{F}_p$ and over $\mathbb{Q}$ are characterized by their underlying abelian groups. This cannot be the case in general: e.g. the underlying abelian groups of a $\mathbb{Q}(\sqrt{2})$-vector space are the same as those of a $\mathbb{Q}(\sqrt{3})$-vector space. This suggests a linguistic adjustment: to capture the structure of a vector space over a field F, we include the action of the elements of F as part of the language. Indeed, this can be done more generally.

Let R be a ring (not necessarily commutative, but with multiplicative identity). To avoid trivialities, we exlcude the zero ring. The language of (say, left) R-modules is, by definition, $\mathcal{L}_R = \{+, -, 0\} \cup \{r : r \in R\}$, i.e., the language of commutative monoids augmented by a *unary function* r for each $r \in R$. The class of left R-modules and R-module homomorphisms is easily seen to be an elementary class of $\mathcal{L}_R$-structures: in other words, the usual axioms for a left R-module are expressable as sentences in $\mathcal{L}_R$.

The theory of R-modules is not complete, because the trivial R-module has one element, whereas the R-module R itself has more than one element. However, there is a class of rings – containing $R = \mathbb{Q}$ as above – such that the theory of nontrivial R-modules is complete.

Theorem 44. *Let R be a ring (not the zero ring!) without zero divisors.*
a) The theory of infinite R-modules is complete iff R is a division ring.
b) The theory of nontrivial R-modules is complete iff R is an infinite division ring.

The following exercises lead a reader through a proof of Theorem 44.

Exercise 4.9: Let R be a ring (not the zero ring!) without zero divisors. An element x in a left R-module M is said to be **torsion** if there exists $0 \neq r \in R$ such that $rx = 0$. A left R-module is **torsionfree** if the only torsion element is zero.
a) Show that R itself is a torsionfree left R-module. (We use here that R has no zero divisors.)
b) Let M and N be left R-modules. If M is torsionfree and $M \equiv N$ in the language of R-modules, then N is torsionfree.
c) Suppose that R is a ring which admits nontrivial torsion left R-modules. Show that the theories of nontrivial and infinite left R-modules are not complete.

Exercise 4.10: For a ring R (not the zero ring)without zero divisors, show that the following are equivalent:
(i) The only left ideals of R are $\{0\}$ and R. (ii) R is a division ring.
(iii) Every left R-module is torsionfree.
(iv) Every left R-module is isomorphic to the direct sum of κ copies of R for a uniquely determined cardinal κ.

Exercise 4.11: Prove Theorem 44.

Exercise 4.12: What about the case of rings with zero divisors?[20]

4.3. **Morley's Categoricity Theorem.**

One cannot help but notice the dichotomy between countable and uncountable cardinals in all of our applications of Vaught's test. It is natural to wonder whether there is a theory which is κ-categorical for some but not all uncountable cardinals. The answer is a resounding no.

Theorem 45. *(Morley's Categoricity Theorem) If a theory is categorical for some uncountable cardinal, then it is categorical for every uncountable cardinal.*

As Marker remarks in his book [Mar], "Morley's proof was the beginning of modern model theory." This theorem is too rich for our blood: beautiful and impressive as it is, it is a theorem of pure model theory: it is hard to imagine a mainstream mathematical problem in which distinct uncountable cardinals arise naturally.[21]

4.4. **Complete, non-categorical theories.**

It is important to emphasize that Vaught's test is only a sufficient condition for completeness. (Indeed, it is the "cheapest" criterion for completeness that I know, but as we have seen, it nevertheless has some useful consequences.) There are complete theories which are far from being κ-categorical for any infinite cardinal κ. As usual, the theory RCF of real-closed fields (again, either in the language of fields or in that of ordered fields; it doesn't matter) is an important example.

[20]This section is taken from my notes on model theory from 2003, in which I had overlooked the point that "torsionfree" is not a good notion for rings with zero divisors. To make things easy for myself here, I have simply added that hypothesis throughout, but it seems likely that something can be said in the general case.

[21]To those readers who are offended by this statement, my apologies: in Athens, GA there are no practitioners of set theory, topos theory, general topology...

Exercise 4.13: Let $\mathcal{T}$ be a theory in a countable language. Show that $\mathcal{T}$ has at most $c = 2^{\aleph_0}$ pairwise nonisomorphic countable models.

Exercise 4.14: Show that RCF, the theory of real-closed fields, has c-many pairwise nonisomorphic countable models. Suggestions:
(i) In fact, there are c-many countable Archimedean real-closed fields. To see this:
(ii) Show that every real number α, there is a countable real-closed subfield R of $\mathbb{R}$ such that $\alpha \in R$.
(iii) Show that any two distinct real-closed subfields of $\mathbb{R}$ are nonisomorphic. (Hint: in a previous exercise you were asked to show that every Archimedean ordered fields order embeds into $\mathbb{R}$. Here we want the fact that this embedding is *unique*, which is in fact easier to see.)

Exercise 4.15: Let C be an algebraically closed field of characteristic 0. The point of this exercise is to use real-closed fields to show that the automorphism group $G = \mathrm{Aut}(C)$ is really big.
a) Show that there exists at least one subfield R of C such that $[C : R] = 2$. (By the Grand Artin-Schreier Theorem, R is real-closed.) Fix one such subfield and call it R_0.
b) Let $H = \{\sigma \in \mathrm{Aut}(C) \mid \sigma(R_0) = R_0\}$. Note that H contains $h = \mathrm{Aut}(C/R_0) = \{1, c_R\}$, a group of order 2. If the unique ordering on R_0 is Archimedean, show that $H = h$.
c)* Show that there exist non-Archimedean real-closed fields with $H \supsetneq h$ and also non-Archimedean real-closed fields with $H = h$. (This is quite difficult.)
d) Show that the coset space G/H is naturally in bijection with the set of all index 2 subfields R of C such that $R \cong R_0$.
e) Show that every real-closed field R with $|R| = c$ embeds as an index 2 subfield of $\mathbb{C}$.
f) Apply part e) and the previous exercise to show that there are precisely $2^c = 2^{2^{\aleph_0}}$ conjugacy classes of order 2 elements in G. In particular $|G| = 2^{2^{\aleph_0}}$.

5. Quantifier elimination: a criterion for model-completeness

Having seen that categoricity is a concept of somewhat limited usefulness, we now turn to a more versatile (and historically prior) technique for establishing model completeness, namely quantifier elimination. But more than just giving further criteria for showing that theories are model / completen, quantifier elimination is a fundamental concept in its own right (arguably more so than categoricity, at least in applied model theory). It leads us to the concept of definable subsets, which – in that it presents model theory as a strict generalization of classical algebraic geometry – is probably the aspect of model theory which is most interesting and useful to mainstream mathematicians.

5.1. Constructible and definable sets.

We have held back the following basic definitions until absolutely needed, which is now.

For a structure M, a **definable** subset of M^n is one which can be realized as the locus of validity of a formula with n unbound variables.

For a structure M, a **constructible** subset of M^n is one which can be realized as the locus of validity of a quantifier-free formula with n variables.

Each of these is a generalization of the notion of *algebraic set* in classical algebraic geometry, i.e., the simultaneous zero locus of a finite family of polynomial equations. However, a constructible set is a very mild generalization which is also familiar in algebraic geometry.[22] For a particular structure M, a basic and important question is whether the class of constructible and definable subsets of M^n coincide, and if not, to seek to understand how much richer the latter class may be.

As our basic example, take a field F in the language of rings. Can we find subsets of F^n which are definable but not constructible?

Example 5.1: Consider the formula

$$\psi(b,c) : \exists x \; x^2 + bx + c = 0$$

it defines the subset of F^2 of coefficients for which the quadratic formula has a solution in the field F. Can this subset be defined by a formula without a quantifier?

It depends on F!

Suppose F is algebraically closed field; then the domain of validity is all of F^2, e.g. given by the formula

$$b = b, c = c$$

which is certainly quantifier-free.

Suppose now $F := \mathbb{R}$ the field of real numbers (or more generally, a real-closed field). Then the domain of validity is given by a polynomial inequality in $\mathbb{R}^2$, namely that $b^2 - 4c \geq 0$. We claim that this locus cannot be defined without quantifiers.

To see this, let us first give a more explicitly geometric description of the constructible subsets of F^n. The "basic" constructible sets are the algebraic sets, i.e., $\{x \in F^n \mid P(x) = 0\}$ for some $P \in F[t_1, \ldots, t_n]$.[23] Using the logical operations $\neg, \wedge, \vee$, we see that the class of constructible sets must be closed under complementation, union and intersection, i.e., it must form an algebra of sets.[24] Indeed, the constructible sets are, almost by definition, the least algebra containing the algebraic sets.

Exercise 5.2: Let F be a field. Show that a subset $S \subset F^n$ is constructible iff there exist Zariski closed subsets $V_1, \ldots, V_n$ and Zariski open subsets $U_1, \ldots, U_n$

[22] Indeed, I have taken the liberty of saying "constructible" instead of the more standard "definable without quantifiers" precisely so as to stress the analogy with the classical geometric case.

[23] One can make sense of the notion of an algebraic subset of a structure M^n in the general case. It is, roughly, the locus of a single formula which does not use either quantifiers, $\vee, \wedge$ or $\neg$.

[24] Indeed, for any structure M, the constructible subsets of M^n form an algebra of sets, as do the definable sets. This is not such a good exercise since there is almost nothing to say, but the reader should make sure she believes it.

such that $S = \bigcup_{i=1}^n V_i \cap U_i$.[25] Show also that the union may be taken to be disjoint.

Here is an analogous characterization of definable sets (in any structure).

Theorem 46. *The definable sets $\{D_n \subset 2^{M^n}\}_{n \geq 1}$ are the smallest family of subsets satisfying the following properties:*
- $M^n \in D^n$
- *if f^M is an n-function, its graph is in D^{n+1}*
- *if R^M is an n-ary relation, its graph is in D^n*
- *for $1 \leq i, j \leq n$, $\{(x_1, \ldots, x_n) : x_i = x_j\} \in D_n$*
- *if $X \in D_n$, $M \times X \in D_{n+1}$*
- *D_n is closed under finite union, finite intersection, and complementation*
- *If $X \in D_{n+1}$ then any coordinate projection of X to D_n is definable*
- *If $X \in D_{n+m}$ and $b \in M^m$, then $\{a \in M^n : (a, b) \in X\} \in D_n$*

Proof. See [Mar, Prop. 1.3.4].[26] □

Exercise 5.3: Let F be a a field, and $P : F^n \to F^n$ a polynomial map.
a) If $S \subset F^n$ is definable, show that $P(S) \subset F^n$ is definable.
b) Let $P(x) = x^2 : \mathbb{R} \to \mathbb{R}$. Find a constructible subset $S \subset \mathbb{R}$ such that $P(S)$ is not constructible.

Exercise 5.4: a) Let $S \subset \mathbb{R}^n$ be a constructible set. Show that at least one of S and $\mathbb{R}^n \setminus S$ has zero Lebesgue measure.
b) Show that the subset of $\mathbb{R}^2$ defined by $\psi(b, c)$ is *not* constructible.

Exercise 5.5: Show that for any $\mathcal{L}$-structure X and any $n \in \mathbb{Z}^+$, each one element subset $\{(a_1, \ldots, a_n)\}$ is definable.

By the previous two exercises, for any $\mathcal{L}$-structure M, the smallest possibility for the algebra of definable subsets of $M = M^1$ is the algebra of finite/cofinite sets, i.e., subsets $A \subset M$ such that either A or $M \setminus A$ is finite. An $\mathcal{L}$-structure M is **strongly minimal** if every definable subset of M is finite or cofinite. Of course this is automatic if M itself is finite. More interestingly, let F be any field. Then the Zariski-closed subsets of F are precisely the finite ones, from which it follows that the constructible subsets of F are precisely the co/finite sets. As we have seen above, at least for some fields F, there are definable sets which are not constructible, so F is not strongly minimal. However, the following important theorem shows, in particular, that any algebraically closed field is a strongly minimal structure.

Theorem 47. *(Tarski) Let $\mathcal{L}$ be the language of rings, let F be an algebraically closed field, and let $\varphi(x) = \varphi(x_1, \ldots, x_n)$ be an $\mathcal{L}$-formula, in the language of rings, with n unbound variables. Then there exists a quantifier free $\mathcal{L}$-formula $\psi(x)$ such that*

$$ACF \models (\forall v_1 \ldots \forall v_n \ (\phi(v_1, \ldots, v_n) \iff \psi(v_1, \ldots, v_n))).$$

In particular, a subset of F^n is definable iff it is constructible.

[25] A subset of a topological space which is the intersection of a closed and an open set is called **locally closed**.

[26] It turns out that we do not need this result but only special cases which are easy to prove from scratch. I am including it here because I gave it in my lecture and the attendees may wish to check that they took down the somewhat complicated statement correctly.

Although the proof of this theorem is not especially difficult, before proving it we wish to derive some important consequences.

Corollary 48. *(Chevalley's Theorem) Let F be an algebraically closed field and $P = (P_1, \ldots, P_n) : F^n \to F^n$ be a polynomial map. Then P maps constructible sets to constructible sets.*

Proof. By Exercise 5.3, over any field a polynomial map carries definable sets to definable sets. By Tarski's theorem, the classes of constructible and definable sets coincide over an algebraically closed field. □

Corollary 49. *The theory ACF of algebraically closed fields is model complete.*

Proof. Let $K \subset L$ be an embedding of algebraically closed fields. Let $\varphi(v_1, \ldots, v_n)$ be a formula, so by Tarski's theorem there exists an equivalent quantifier-free formula $\psi(v_1, \ldots, v_n)$. Because quantifier-free formulas are preserved by substructure and extension, for any $(a_1, \ldots, a_n) \in K^n$ we have

$$K \models \varphi(a_1, \ldots, a_n) \iff K \models \psi(a_1, \ldots, a_n)$$
$$\iff L \models \psi(a_1, \ldots, a_n) \iff L \models \varphi(a_1, \ldots, a_n).$$

□

5.2. **Quantifier Elimination: Definition and Implications.**

The property of ACF expressed in Tarski's theorem is of much more general interest.

We say that a theory $\mathcal{T}$ admits **elimination of quantifiers** if for each $\mathcal{L}$-formula $\varphi(x) = \varphi(x_1, \ldots, x_n)$, there exists a quantifier free $\mathcal{L}$-formula $\psi(x)$ such that

$$ACF \models (\forall v_1 \ldots \forall v_n \ (\phi(v_1, \ldots, v_n) \iff \psi(v_1, \ldots, v_n))).$$

Then the proof of Corollary 50 goes through verbatim to give the following result.

Proposition 50. *A theory which admits quantifier elimination is model complete.*

In fact quantifier elimination gives more than just model completeness: it has the *geometric* consequence that the algebras of definable and constructible sets coincide. In the case of ACF, this is expressed via Chevalley's Theorem, which is a indeed a result of classical algebraic geometry.

Along with proving Tarski's Theorem 47, our main order of business in this section is to prove that the theory RCF of real-closed fields is model-complete. Is it possible to prove this using quantifier elimination?

Taken literally, we have already seen that the answer is no. Indeed, we saw that the definable subsets of $\mathbb{R}^n$ include more than just the constructible sets: using quantifiers we can also define the order relation $<$ on $\mathbb{R}$ and thus the algebra of definable sets certainly contains the **semialgebraic sets**, i.e., the algebra generated by polynomial equations and also polynomial inequalities.

However, there is a way out. Namely, we work instead in the language $\mathcal{L} = \{+, -, \cdot, <, 0, 1\}$ of *ordered fields*. As we have discussed, a real-closed field admits a unique ordering so that there is a fundamental equivalence of structures: define

RCOF to be the union of the theory of ordered fields with the theory of real-closed fields. Then every model of RCOF is uniquely a model of RCF and conversely. In particular, the embeddings of structures between two real-closed fields are exactly the same in the language of fields as in the language of ordered fields. One can check that this implies that RCF is model-complete iff RCOF is model-complete.

The following considerations formalize this and clarify the relationship between quantifier elimination and model completeness. Let $\mathcal{L} \subset \mathcal{L}'$ be languages and let $\mathcal{T}$ an $\mathcal{L}$-theory. Consider an $\mathcal{L}'$-theory $\mathcal{T}'$ such that $\mathcal{T}' \cap \mathcal{L} = \mathcal{T}$. We say that $(\mathcal{L}', \mathcal{T}')$ is a **supercool extension** of $(\mathcal{L}, \mathcal{T})$ if both of the following hold:

(SCE1) for every $\mathcal{L}$'-formula $\varphi(\overline{x})$ there exists an $\mathcal{L}$-formula $\psi(\overline{x})$ such that

$$\mathcal{T} \models \forall \overline{x}(\varphi(\overline{x}) \iff \psi(\overline{x}))$$

(SCE2) For every embedding $\iota : X \to Y$ of models of $\mathcal{T}$, we may extend X and Y to models X' and Y' of $\mathcal{T}'$ such that ι is an embedding of $\mathcal{L}'$-structures.

In particular, every model X of $\mathcal{T}'$ is also a model of $\mathcal{T}$ and the $\mathcal{L}'$-definable subsets of X^n are precisely the $\mathcal{L}$-definable subsets of X^n. However, the $\mathcal{L}'$-constructible subsets of X^n may well be richer than the $\mathcal{L}$-constructible subsets of X^n.

Example 5.6: RCOF is a supercool extension of RCF.

Proposition 51. *Let $(\mathcal{L}', \mathcal{T}')$ be a supercool expansion of $(\mathcal{L}, \mathcal{T})$. Then $\mathcal{T}$ is model-complete iff $\mathcal{T}'$ is model-complete.*

Proof. If $\mathcal{T}$ is model complete and $\iota : X \hookrightarrow Y$ is an embedding of $\mathcal{L}'$-structures, it is also an embedding of $\mathcal{L}$-structures, hence an elementary embedding. Since by (SC1) every $\mathcal{L}'$-formula has an equivalent $\mathcal{L}$-formula, it follows that ι is an elementary embedding of $\mathcal{L}$-structures. Conversely, suppose that $\mathcal{L}'$ is model complete and $\iota : X \hookrightarrow Y$ is an embedding of $\mathcal{L}$-structures. By (SC2), we may extend X and Y to models of $\mathcal{T}'$ such that ι is an embedding of $\mathcal{L}'$-structures. By assumption, ι is an elementary embedding of $\mathcal{L}'$-structures, so *a fortiori* it is an elementary embedding of $\mathcal{L}$-structures. □

Theorem 52. *Consider the following conditions on a theory $\mathcal{T}$:*
(i) There exists a supercool extension $\mathcal{T}'$ of $\mathcal{T}$ which admits quantifier elimination.
(ii) $\mathcal{T}$ is model-complete.
Then (i) $\implies$ (ii).

Proof. If (i) holds, $\mathcal{T}'$ is model-complete. By Proposition 51, $\mathcal{T}$ is model-complete. □

Exercise 5.7: a) In the statement of Theorem 52, does (ii) imply (i)? (This would be nice, since then any model-complete theory admits an extension which has quantifier elimination and also the same class of definable sets.)
b) If not, can you tweak the definition of "supercool extension" to make this work?[27]

Theorem 53. *(Tarski-Robinson) The theory RCOF admits quantifier elimination. In particular:*

[27]Admittedly, this is more of an exercise for me than for you. I'll work on it...

a) RCOF (hence also RCF) is model complete.
b) If R is a real-closed field, a subset of R^n is definable iff it is semialgebraic.

This has a geometric consequence analogous to Chevalley's Theorem. First let's consider an example: for a real-closed field R, the subset of R^2 cut out by $x^2+y^2=1$ is an algebraic set – the unit circle. On the other hand, its projection onto the x-axis is the interval $[-1,1]$: this is not algebraic nor even constructible, but it is semialgebraic. This is an instance of the robustness of semialgebraic sets as guaranteed by the following result, an immediate consequence of the Tarski-Robinson Theorem.

Corollary 54. *(Tarski-Seidenberg) Let R be a real-closed field. The image of a semialgebraic set $S \subset R^n$ under a coordinate projection to R^m, $m < n$, is again semialgebraic.*

This result is of foundational importance in the burgeoning subject of real algebraic geometry. Indeed, we recommend that the interested reader consult [**?**], in which the Tarski-Seidenberg theorem plays the starring role in Chapter 1.

5.3. **A criterion for quantifier elimination.**

Acknowledgment: In this section, we are following David Marker's text [Mar] (as well as certain related lecture notes of Marker's) especially closely. Some of the proofs are taken verbatim from [Mar, Ch. 3].

For $\mathcal{L}$ a language, the set of **terms** in $\mathcal{L}$ is the smallest set containing the constant symbols, the fixed countably infinite set $\{x_n\}_{n=1}^{\infty}$ of variables, and, for each n-ary function f in $\mathcal{L}$, all expressions of the form $f(t_1,\ldots,t_n$ where $t_1,\ldots,t_n$ are terms.

An **atomic $\mathcal{L}$-formula** is a formula either expressing an equality of two terms $t_1 = t_2$ or $R(t_1,\ldots,t_n)$, where R is an n-ary relation and $t_1,\ldots,t_n$ are terms.

Let X be an $\mathcal{L}$-structure and, as before, let $\mathcal{L}_X$ be the language $\mathcal{L}$ augmented with constant symbols for each element of X. The **atomic diagram** $D(X)$ is the set of atomic formulas $\varphi(x_1,\ldots,x_n)$ which are true in X together with the negated atomic formulas which are true in X.

Lemma 55. *(Diagram Lemma) Let Y be an $\mathcal{L}_X$-structure which is a model of $D(X)$. Then there exists an $\mathcal{L}$-emebedding $X \hookrightarrow Y$.*

Proof. Since we have an interpretation of each constant symbol of X in Y, this gives a map of sets $\iota : X \to Y$. Moreover, since for distinct elements $x_1 \neq x_2$ of X, $x_1 \neq x_2$ lies in $D(X)$, is is also true in Y. Therefore ι is an injection. That ι preserves functions and relations is an easy exercise involving the definitions which is left to the reader. □

Theorem 56. *Let $\mathcal{L}$ be a language with a constant symbol c, let $\mathcal{T}$ be an $\mathcal{L}$-theory and $\varphi(\overline{v})$ an $\mathcal{L}$-formula with unbound variables $\overline{v} = (v_1,\ldots,v_n)$. TFAE:*
(i) There exists a quantifier-free $\mathcal{L}$-formula $\psi(\overline{v})$ such that

$$\mathcal{T} \models \forall \overline{v}\ (\varphi(\overline{v}) \iff \psi(\overline{v})). \tag{2}$$

(ii) For any two models X, Y of $\mathcal{T}$ and an $\mathcal{L}$-structure A such that $A \subset X$, $A \subset Y$, then for all $a \in A^n$, $X \models \varphi(a) \iff Y \models \psi(a)$.

Proof. (i) $\implies$ (ii): Let $\psi(\overline{v})$ be a quantifier-free formula satisfying (2), and let $\overline{a} \in A^n$. Then, because quantifier-free formulas are preserved by all embeddings of structures, we have that $\varphi(\overline{a})$ is true in X iff $\psi(\overline{a})$ is true in A iff $\psi(\overline{a})$ is true in Y iff $\varphi(\overline{a})$ is true in Y.

(ii) $\implies$ (i): First, if $\mathcal{T} \models \forall \overline{v}\ \varphi(\overline{v})$, then $\mathcal{T} \models \forall \overline{v}\ (\varphi(\overline{v}) \iff c = c)$. Second, if $\mathcal{T} \models \forall \overline{v}\ \neg\varphi(\overline{v})$, then $\mathcal{T} \models \forall \overline{v}\ (\varphi(c) \iff c \neq c)$. Thus we may assume that both $\mathcal{T} \cup \{\exists \overline{v}\ \varphi(\overline{v})\}$ and $\mathcal{T} \cup \{\neg \exists \overline{v} \varphi(\overline{v})\}$ are satisfiable.

Now let $\Gamma(\overline{v})$ be the set of quantifier-free formulas $\psi(\overline{v})$ such that $\mathcal{T} \models \forall \overline{v}\ (\varphi(\overline{v}) \implies \psi(\overline{v}))$. Let $\overline{d} = (d_1, \ldots, d_n)$ be new constant symbols.
Claim: $\varphi(\overline{d}) \in \overline{\mathcal{T} \cap \Gamma(\overline{d})}$.
Suppose first that the claim holds. Then, by completeness and finite character of syntactic implication, there are $\psi_1, \ldots, \psi_n \in \Gamma$ such that

$$\mathcal{T} \models \forall \overline{v} \left(\bigwedge_{i=1}^{m} \psi_i(\overline{v}) \implies \varphi(\overline{v}) \right),$$

and thus

$$\mathcal{T} \models \forall \overline{v} \left(\bigwedge_{i=1}^{m} \psi_i(\overline{v}) \iff \varphi(\overline{v}) \right)$$

and $\bigwedge_{i=1}^m \psi_i(\overline{v})$ is quantifier-free.
Proof of Claim: suppose not, and let X be a model of $\mathcal{T} \cup \Gamma(\overline{d}) \cup \{\neg\varphi(\overline{d})\}$. Let A be the substructure of X generated by $\overline{d}$ (i.e., the intersection of all substructures of X containing $\overline{d}$). Put

$$\Sigma = \mathcal{T} \cup D(A) \cup \varphi(\overline{d}).$$

If Σ is unsatisfiable, then there are quantifier-free formulas $\psi_1(\overline{d}), \ldots, \psi_m(\overline{d}) \in D(A)$ such that

$$\mathcal{T} \models \forall \overline{v}\ \left(\bigwedge_{i=1}^{m} \psi_i(\overline{v}) \implies \neg\varphi(\overline{v}) \right).$$

But then, contrapositively,

$$\mathcal{T} \models \forall \overline{v}\ \left(\varphi(\overline{v}) \implies \bigvee_{i=1}^{m} \neg\psi_i(\overline{v}) \right),$$

so $\bigvee_{i=1}^m \neg\psi_i(\overline{v}) \in \Gamma$ and thus $A \models \bigvee_{i=1}^m \neg\psi_i(\overline{d})$, a contradiction. Thus Σ is satisfiable.

Let Y be a model of Σ, so that in particular $\varphi(\overline{d})$ holds in Y. Moreover, since $\Sigma \supset D(A)$, by the Diagram Lemma (Lemma 55) we may embed $A \hookrightarrow Y$. But $\neg\varphi(\overline{d})$ holds in X, so that by our assumption (ii), $\neg\varphi(\overline{d})$ holds in Y, contradiction. $\square$

The next result says that to prove quantifier elimination it suffices to remove one existential quantifier at a time.

Lemma 57. *Let $\mathcal{T}$ be an $\mathcal{L}$-theory. Suppose that for each quantifier-free $\mathcal{L}$-formula $\theta(\overline{v}, w)$ there is a quantifier-free formula $\psi(\overline{v})$ such that*

$$\mathcal{T} \models (\forall v \exists w\ \theta(\overline{v}, w) \iff \psi(\overline{v})).$$

Then $\mathcal{T}$ admits quantifier elimination.

Proof. Let $\varphi(\overline{v})$ be an $\mathcal{L}$-formula. We wish to show that there exists a quantifier-free $\mathcal{L}$-formula $\psi(\overline{v})$ such that

$$\mathcal{T} \models (\forall \overline{v}\ (\varphi(\overline{v}) \iff \psi(\overline{v}))).$$

We prove this by an induction on the complexity of φ.
Step 0: Of course, if φ is itself quantifier-free, there is nothing to show.
Now suppose that for $i = 0, 1$,

$$\mathcal{T} \models \forall \overline{v}\ (\theta_i(\overline{v}) \iff \psi_i(\overline{v})),$$

where each ψ_i is quantifier-free.
Step 1a: Suppose $\varphi(\overline{v}) = \neg\theta_0(\overline{v})$, then

$$\mathcal{T} \models \forall \overline{v}\ (\varphi(v) \iff \neg\psi_0(\overline{v})).$$

Step 1b: Suppose $\varphi(\overline{v}) = \theta_0(\overline{v}) \wedge \theta_1(\overline{v})$, then

$$\mathcal{T} \models \forall \overline{v}\ (\varphi(v) \iff (\psi_0(\overline{v}) \wedge \psi_1(\overline{v}))).$$

In either case, φ is equivalent to a quantifier-free formula.
Step 2: Suppose that

$$\mathcal{T} \models \forall \overline{v}\ (\theta(\overline{v}, w) \iff \psi_0(\overline{v}, w)),$$

where ψ_0 is quantifier-free and $\varphi(\overline{v}) = \exists w\ \theta(\overline{v}, w)$. Then

$$\mathcal{T} \models \forall \overline{v}\ (\varphi(\overline{v}) \iff \exists w \psi_0(\overline{v}, w)).$$

By our assumptions, there is a quantifier-free formula $\psi(\overline{v})$ such that

$$\mathcal{T} \models \forall \overline{v}\ (\exists w\ \psi_0(\overline{v}, w) \iff \psi(\overline{v})).$$

But then

$$\mathcal{T} \models \forall \overline{v}\ (\varphi(\overline{v}) \iff \psi(\overline{v})).$$

$\square$

Corollary 58. *Let $\mathcal{T}$ be an $\mathcal{L}$-theory. Suppose that for all quantifier-free formulas $\varphi(\overline{v}, w)$, if X and Y are models of $\mathcal{T}$, A is a common substructure of X and Y, and all $\overline{a} \in A^n$ such that there exists $b \in X$ with $X \models \varphi(\overline{a}, b)$, then there exists $c \in Y$ such that $Y \models \varphi(\overline{a}, c)$. Then $\mathcal{T}$ admits quantifier elimination.*

5.4. **Model-completeness of ACF.**

In this section we will use the criterion of Corollary 58 to prove the model-completeness of ACF. We work in the language $\mathcal{L} = \{+, -, \cdot, 0, 1\}$. (The reader will soon see why having $-$ is useful here.) Suppose K and L are algebraically closed fields, and A is a common substructure of K and L. What we must show is that, for a quantifer-free formula $\varphi(v, \overline{w})$, $\overline{a} \in A^n$ and b in K such that $\varphi(b, \overline{a})$ holds in K, there exists $c \in L$ such that $\varphi(c, \overline{a})$ is true in L.

First a litle algebra: as an $\mathcal{L}$-substrcture of K (or L...), A is a subring of a field – here we use that $-$ is part of the structure; otherwise it need only be a semiring! – hence an integral domain. Let k be the fraction field of A and F an algebraic closure. Since K and L are algebraically closed fields containing A, they also contain (up to unique k-algebra isomorphism) F. Therefore it will suffice to show the following: for $\overline{a} \in F^n$ and b in K such that $\varphi(b, \overline{a})$ holds in K, there exists $c \in F$ such that $\varphi(c, \overline{a})$ holds in F. Indeed, since φ is quantifier-free and F is a substructure of L, necessarily then $\varphi(c, \overline{a})$ holds in L. (Thus we have taken L out of the picture entirely. See §5.6 for more perspective on this.)

We may put the formula φ into **disjunctive normal form**: that is, there are positive integers N and M and a family $\{\theta_{i,j}(\overline{v}, w)\}_{1 \leq i \leq N,\ 1 \leq j \leq M}$ such that

$$\forall v \ \left(\varphi(v, \overline{a}) \iff \bigvee_{i=1}^{N} \bigwedge_{j=1}^{M} \theta_{i,j}(v, \overline{a}) \right).$$

Thus we reduce to the case in which $\varphi(x, \overline{y})$ is a conjunction of atomic and negated atomic formulas. But that just means that there are polynomials

$$p_1, \ldots, p_n, q_1, \ldots, q_m \in F[X]$$

such that $\phi(v, \overline{a})$ is equivalent to

$$\left(\bigwedge_{i=1}^{n} p_i(v) = 0 \right) \wedge \left(\bigwedge_{i=1}^{m} q_i(v) \neq 0 \right).$$

If at least one of the p_i is nonzero, then b is algebraic over F, so $b \in F$ and there is nothing to show. Otherwise the formula is equivalent to finitely many one-variable polynomials not vanishing, so almost any element of the (infinite!) field F will do.

5.5. **Model-completeness of RC(O)F.**

In this section we will use the criterion of Corollary 58 to prove the model-completeness of RCOF (real-closed ordered fields). We work in the language $\mathcal{L} = \{+, -, \cdot, < 0, 1\}$ of ordered rings.

Now some algebra of ordered fields: as an $\mathcal{L}$-substructure of K (or L...), A is a subring of an ordered field, hence an ordered integral domain. Let k be the fraction field of A – it is easy to see that the order extends uniquely to k. Now let F be the real-closure of $(k, <)$, unique up to k-algebra isomorphism. Because of this uniqueness property of the real-closure of an *ordered* field, we get that F may be embedded, as an $\mathcal{L}$-structure, in K and L. Therefore it will suffice to show the following: for $\overline{a} \in F^n$ and b in K such that $\varphi(b, \overline{a})$ holds in K, there exists $c \in F$ such that $\varphi(c, \overline{a})$ holds in F. Indeed, since φ is quantifier-free and F is a substructure of L, necessarily then $\varphi(c, \overline{a})$ holds in L. As above, φ is logically equivalent to a disjunction of conjunctions of formulas of the form $p(v, \overline{w}) = 0$, $p(v, \overline{w}) > 0$; arguing as above, there are polynomials $p_1, \ldots, p_n, q_1, \ldots, q_m \in F[X]$ such that

$$\phi(v, \overline{a}) \iff \left(\bigwedge_{i=1}^{n} p_i(v) = 0 \right) \wedge \left(\bigwedge_{i=1}^{m} q_i(v) > 0 \right).$$

If any p_i is not identically zero, b is algebraic over F but $F(b) < K$ is formally real, so F real-closed implies $b \in F$. So it comes down to

$$\phi(v, \overline{a}) \iff \bigwedge_{i=1}^{m} q_i(v) > 0.$$

We are thus given that $q_i(b) > 0$ for all i. I claim that we can find $c_i, d_i \in F$ such that $c_i < b < d_i$ and $q_i(x)$ is positive on the interval (c_i, d_i). We are now in the endgame of the proof of the single most important result of our course. The killing blow comes from an unexpected source!

Theorem 59. *(Intermediate Value Theorem)*
Let F be an ordered field.
a) Endow F with the linear topology obtained by taking $(a, b) := \{x : a < x < b\}$ as a basis for the open sets. Then polynomial functions are continuous on F.
b) Assume F is moreover real-closed, and let $f \in F[x]$ be such that $f(a) < 0$, $f(b) > 0$. Then f vanishes somewhere on (a, b).

Proof. Since addition and multiplication are clearly continuous with respect to the linear topology, part a) is obvious. For part b), we may assume f is irreducible (some factor must change signs). We know that f is then either linear (no problem), or $f(x) = x^2 + cx + d$ where $c^2 - 4d < 0$. But then

$$f(x) = (x + c/2)^2 + (d - c^2/4)^2$$

and indeed f is positive for all x. □

Indeed, by the intermediate value theorem, $q_i(v)$ can only change sign on F by passing through a root of q_i, of which there are only finitely many. This establishes the claim. Now take $c = \max c_i$ and $d = \min d_i$ and let b' be any element in (c, d); this completes the proof.

5.6. **Algebraically Prime Models.**

In the proofs of quantifier elimination in both ACF and RCOF, things turned out to be pleasantly simpler than they could have been, in a common way. Namely, the criterion of Corollary 55 *a priori* requires us to consider models X and Y which are somewhat indirectly related: they have a common *substructure* A, but A need not be a model of $\mathcal{T}$. It would be nice if we had a more direct relationship between X and Y, e.g. if X were a substructure of Y.

But, in essence, the proofs of quantifier elimination for ACF and RCOF reduce to a substructure situation. This occurs because there is a canonical "minimal" way to pass from an integral domain (resp. ordered integral domain) to an algebraically closed field (resp. real-closed ordered field). This algebraic property can be phrased model-theoretically and leads to a useful alternate version of Corollary 55.

A sentence is said to be **universal** if the existential quantitifer does not appear in it. A theory $\mathcal{T}$ has a **universal axiomization** if there exists a set Γ of universal sentences such that $\overline{\Gamma} = \overline{\mathcal{T}}$, i.e., such that Γ and $\mathcal{T}$ have the same models.

Theorem 60. *For a theory $\mathcal{T}$, TFAE:*
(i) $\mathcal{T}$ has a universal axiomization.
(ii) If X is a model of $\mathcal{T}$ and Y is a substructure of X, then Y is a model of $\mathcal{T}$.

Proof. (i) $\implies$ (ii): Suppose $\mathcal{T}$ has a universal axiomatization; without loss of generality we may as well assume that $\mathcal{T}$ consists entirely of universal sentences. It is clear that if φ is a universal sentence and $Y \subset X$ is an $\mathcal{L}$-substructure of a structure X such that $X \models \varphi$, then also $Y \models \varphi$.
(ii) $\implies$ (i): Suppose that the set of models of $\mathcal{T}$ is closed under passage to substructures. Let Γ be the set of all universal sentences in $\overline{\mathcal{T}}$. We want to show that $\overline{\Gamma} = \overline{\mathcal{T}}$. By definition, every model of $\mathcal{T}$ is a model of Γ. Conversely, let Y be

a model of Γ. We want to show that Y is a model of $\mathcal{T}$, and by (ii) it suffices to construct a model X of $\mathcal{T}$ and an embedding of structures $Y \hookrightarrow X$.

Claim: $\mathcal{T} \cup D(Y)$ is a satisfiable $\mathcal{L}_Y$-theory. If not, by compactness, there exists a finite subset $\Delta = \{\psi_1, \ldots, \psi_n\} \subset D(Y)$ such that $\mathcal{T} \cup \Delta$ is not satisfiable. Let $\overline{c}$ be the vector of new constant symbols from Y used in the $\mathcal{L}$-formulas $\psi_1, \ldots, \psi_n$; say $\psi_i = \varphi_i(\overline{c})$, where φ_i is a quantifier-free $\mathcal{L}$-formula. If there were a model of $\mathcal{T} \cup \{\exists \overline{v} \bigwedge_i \varphi_i(\overline{v})\}$, then taking $\overline{v} = \overline{c}$ shows that $\mathcal{T} \cup \Delta$ is satisfiable. Thus

$$\mathcal{T} \models \forall \overline{v} \left(\bigvee_i \neg \varphi_i(\overline{v}) \right).$$

But this is a universal sentence, i.e.,

$$\Gamma \models \forall \overline{v} \left(\bigvee_i \neg \varphi_i(\overline{v}) \right),$$

contradicting the fact that Y is a model of Γ.

By the Diagram Lemma (Lemma 55) there exists a model X of $\mathcal{T}$ and an $\mathcal{L}$-embedding $Y \hookrightarrow X$, qed. □

Example 5.8: In the language $\{+, -, \cdot, 0, 1\}$ of rings, the theory of integral domains has a universal axiomatization, since a substructure of a domain is a domain. (And indeed the usual axiomatization of integral domains is universal.) Since $\mathbb{Z}$ is an $\mathcal{L}$-substructure of $\mathbb{Q}$ which is not a field, the theory of fields does not have a universal axiomatization. Similarly, in the language $\{+, 0, \cdot, <, 0, 1\}$ of ordered rings, the theory of ordered integral domains has a universal axiomatization but the theory of ordered fields does not.

For a theory $\mathcal{T}$, we put $\mathcal{T}_\forall$ to be the set of all universal sentences φ such that $\mathcal{T} \models \varphi$; $\mathcal{T}_\forall$ is said to be the theory of **universal consequences** of $\mathcal{T}$.

Proposition 61. *a) The models of the theory of universal consequences of ACF are precisely the integral domains.*
b) The models of the theory of universal consequences of RCOF are precisely the ordered integral domains.

Exercise 5.9: Prove Proposition 61.

A theory $\mathcal{T}$ has **algebraically prime models** if for each model A of $\mathcal{T}_\forall$ there exists a model X of $\mathcal{T}$ and an embedding $\iota : A \hookrightarrow X$ such that for every model Y of $\mathcal{T}$ and every embedding $j : A \hookrightarrow Y$, there exists an embedding $i : X \hookrightarrow Y$ such that $j = i \circ \iota$.

Lemma 62. *The theories ACF and RCOF each admit algebraically prime models.*

Proof. For ACF this says that to each integral domain A we must find an embedding ι into an algebraically closed field F such that every embedding from A into an algebraically closed field K factors through ι. It is easy to see that we may take F to be an algebraic closure of the fraction field of A. Similarly, for RCOF this says that to an ordered integral domain A we must find an embedding ι into a real-closed field F such that every embedding from A into a real-closed field K factors through ι. The fraction field k of an ordered integral domain is an ordered field; let F be a real-closure of the ordered field $(k, <)$. That $\iota : A \hookrightarrow F$ has the desired property follows from the uniquness of the real-closure of an ordered field: any two real-closures of the same ordered field k are isomorphic as k-algebras. □

This leads to a simplified version of Corollary 58.

Theorem 63. *(Criterion for quantifier elimination) Let $\mathcal{T}$ be a theory such that:*
(i) $\mathcal{T}$ has algebraically prime models, and
(ii) for any models $X \subset Y$ of $\mathcal{T}$, any quantifier-free formula $\varphi(\overline{v}, w)$ and any $\overline{a} \in X^n$, if $Y \models \exists w \varphi(\overline{a}, w)$ then also $X \models \exists w \varphi(\overline{a}, w)$. Then $\mathcal{T}$ admits quantifier elimination.

Exercise 5.10: Deduce Theorem 63 from Corollary 58.

We highly recommend that the reader look back at the proofs of quantifier elimination in ACF and RCOF and verify that, without saying so in so many words, we used Lemma 58 to reduce Corollary 55 to Theorem 59.

Chapter 3 of [Mar] contains other instances of elimination of quantifiers using Theorem 59, e.g. nontrivial ordered torsionfree divisible abelian groups.

6. Ultraproducts and ultrapowers in model theory

6.1. Filters and ultrafilters.

A **filter** $\mathcal{F}$ on a set X is a nonempty family of nonempty subsets of X satisfying the following properties:

(F1) $A_1, A_2 \in \mathcal{F} \implies A_1 \cap A_2 \in \mathcal{F}$, and
(F2) $A_1 \in \mathcal{F}$, $A_2 \supset A_1 \implies A_2 \in \mathcal{F}$.

That is, a filter is a family of nonempty subsets that is stable under finite intersections and passage to supersets.

Example 6.1: For $\emptyset \neq Y \subset X$, define $\mathcal{F}_Y = \{A \subset X \mid Y \subset A\}$ to be the family of all subsets of X containing the fixed nonempty subset Y. This is a filter. Such filters are called **principal**.

Example 6.2: Let X be an infinite set. A subset $Y \subset X$ is said to be **cofinite** if $X \setminus Y$ is finite. The collection of all cofinite subsets of X is a nonprincipal filter, the **Fréchet filter**.

A filter $\mathcal{F}$ on a set X is **free** if $\bigcap_{A \in \mathcal{F}} A = \emptyset$.

Exercise 6.3: Let $\mathcal{F}$ be a free filter on X.
a) Show that $\mathcal{F}$ is not principal.
b) Show that $\mathcal{F}$ contains the Fréchet filter.

Exercise 6.4: a) Let $\{\mathcal{F}_i\}_{i \in I}$ be an indexed family of filters on X. Show that $\mathcal{F} = \bigcap_{i \in I} \mathcal{F}_i$ is a filter, indeed the largest filter which is contained in each $\mathcal{F}_i$.
b) Let X be a set with at least two elements. Exhibit filters $\mathcal{F}_1$ and $\mathcal{F}_2$ on X such that there is no filter $\mathcal{F}$ containing both $\mathcal{F}_1$ and $\mathcal{F}_2$.

The collection of all filters on a set X is partially ordered under containment. By Exercise 6.4a), this poset contains arbitrary **joins** – i.e., any collection of filters admits a greatest lower bound; on the other hand, Exercise 6.4b) shows that when $|X| > 1$ the poset of filters on X is not directed. If $\mathcal{F}_1 \subset \mathcal{F}_2$ we say that $\mathcal{F}_2$ **refines** $\mathcal{F}_1$ or is a **finer filter** than $\mathcal{F}_1$.

Definition: An **ultrafilter** on X is a maximal element in the poset of filters on X, i.e., a filter which is not properly contained in any other filter on X.

The following is probably the single most important property of ultrafilters.

Theorem 64. *Let $\mathcal{F}$ be a filter on X.*
a) $\mathcal{F}$ is an ultrafilter iff: for all $Y \subset X$, exactly one of $Y, X \setminus Y$ lies in $\mathcal{F}$.

Proof. Let $\mathcal{F}$ be an ultrafilter on X and $Y \subset X$. Suppose first that for all $A \in \mathcal{F}$, $(A \cap Y) \neq \emptyset$. Let $F' = \{A \cap Y \mid A \in \mathcal{F}\}$ and let $\mathcal{F}'$ be the collection of all subset of X containing at least one element of F'. It is easy to see that $\mathcal{F}'$ is a filter on X which contains $\mathcal{F}$. Since $\mathcal{F}$ is an ultrafilter, we must have $\mathcal{F} = \mathcal{F}'$ and thus $Y = X \cap Y \in \mathcal{F}' = \mathcal{F}$. Now suppose that there exists $A \in \mathcal{F}$ such that $A \cap Y = \emptyset$. Equivalently, $A \subset X \setminus Y$ and since $A \in \mathcal{F}$, $X \setminus Y \in \mathcal{F}$.
Now suppose that $\mathcal{F}$ is a filter on X which, given any subset of X, contains as an element either that subset or its complement. Suppose $\mathcal{F}'$ is a filter properly containing $\mathcal{F}$, so that there exists some subset $Y \in \mathcal{F}' \setminus \mathcal{F}$. But then $X \setminus Y \in \mathcal{F}' \subset \mathcal{F}$ so that $\mathcal{F}'$ contains both Y and $X \setminus Y$ and thus contains their intersection, the empty set: contradiction. □

Corollary 65. *Let $\mathcal{F}$ be an ultrafilter on X, let $A \in \mathcal{F}$, and let A_1, A_2 be subsets of X such that $A_1 \cup A_2 = A$. Then at least one of A_1 and A_2 lies in $\mathcal{F}$.*

Proof. Assume not. Then by Theorem 64, both $X \setminus A_1$ and $X \setminus A_2$ lie in $\mathcal{F}$, and hence so does

$$(X \setminus A_1) \cap (X \setminus A_2) = X \setminus (A_1 \cup A_2) = X \setminus A.$$

Thus $\mathcal{F}$ contains both A and its complement $X \setminus A$, contradiction. □

Corollary 66. *Let $\mathcal{F}$ be an ultrafilter on X. Then the following are equivalent:*
(i) $\mathcal{F}$ is not free.
(ii) $\mathcal{F}$ is principal.
(iii) There exists $x \in X$ such that $\mathcal{F}$ is the collection of all subsets containing x.

Proof. The imlications (iii) $\implies$ (ii) $\implies$ (i) clearly hold (for arbitrary filters). Suppose that $\mathcal{F}$ is not free, i.e., there exists $x \in \bigcap_{A \in \mathcal{F}} A$. Then $X \setminus \{x\}$ is not an element of $\mathcal{F}$, so by Theorem 64 we have $\{x\} \in \mathcal{F}$, so that $\mathcal{F}$ is the principal filter on the singleton set $\{x\}$. □

Remark: There exist (non ultra)filters which are neither free nor principal, for instance the filter $\{\{0\}, \mathbb{R}\}$ on $\mathbb{R}$. But no matter.

Proposition 67. *a) For a family $\mathcal{A}$ of nonempty subsets of a set X, the following are equivalent:*
(i) I has the ***finite intersection property****: if $A_1, \ldots, A_n \in I$, then $\bigcap_{i=1}^n A_i \neq \emptyset$.*
(ii) There exists a filter $\mathcal{F} \supset \mathcal{A}$.

A family F satisfying these equivalent conditions is called a ***filter subbase****.*[28]
b) For any filter subbase $\mathcal{A}$, there is a unique minimal filter $\mathcal{F}$ containing $\mathcal{A}$, called the ***filter generated by*** *$\mathcal{A}$.*

Proof. a) Certainly the finite intersection property (f.i.p., for short) is necessary for $\mathcal{A}$ to extend to a filter. Conversely, given a family of sets $\mathcal{A}$ satisfying f.i.p., we build the filter it generates in much the same way that we build the topology generated by a subbase. Namely, let F be the family of all finite intersections of elements of $\mathcal{A}$,[29] and let $\mathcal{F}$ be the family of all subsets of X containing some element of F. It is easy to check that $\mathcal{F}$ is a filter.
b) Every filter $\mathcal{G}$ containing every element of $\mathcal{A}$ must contain all supersets of all finite intersections of elements of $\mathcal{A}$, so the filter $\mathcal{F}$ constructed in part a) above is the unique minimal filter containing $\mathcal{A}$. $\square$

Exercise 6.5: Let $\mathcal{F}_1$ and $\mathcal{F}_2$ be two filters on a set X. Show that the following are equivalent:
(i) For all $A \in \mathcal{F}_1$ and all $B \in \mathcal{F}_2$, $A \cap B \neq \varnothing$.
(ii) The set $\mathcal{F}_1 \cup \mathcal{F}_2$ satisfies the finite intersection condition.
(iii) There exists a filter $\mathcal{F}$ containing both $\mathcal{F}_1$ and $\mathcal{F}_2$.
When these equivalent conditions are satisfied, we say that the filters $\mathcal{F}_1$ and $\mathcal{F}_2$ are **compatible**. This should be thought of in analogy to the situation of ideals I_1 and I_2 such that the ideal $I_1 + I_2$ is proper.

The next result collects some further properties of filters, indeed exactly those that we will need for our model-theoretic applications.

Proposition 68.
a) Let $\mathcal{F}$ be a filter on X. Then there exists an ultrafilter containing $\mathcal{F}$.
b) Any infinite set admits a nonprincipal ultrafilter. Indeed, let $Y \subset X$ with Y infinite. Then there exists a nonprincipal ultrafilter $\mathcal{F}$ on X such that $Y \in \mathcal{F}$.

Proof. a) It is easy to see that the union of a chain of filters on X is a filter on X. Therefore Zorn's Lemma applies to give a maximal element in the poset of filters containing a given filter $\mathcal{F}$, i.e., an ultrafilter containing $\mathcal{F}$.
b) Let $\mathcal{F}_0$ be the Fréchet filter (of cofinite subsets of X), and let $\mathcal{F}_Y = \{A \subset X \mid A \supset Y\}$ be the principal filter on Y. Since Y is infinite, if $B \subset X$ is any cofinite set, $Y \cap B \neq \varnothing$. It follows that the filters $\mathcal{F}_0$ and $\mathcal{F}_Y$ are compatible in the sense of Exercise 6.5, so there exists an ultrafilter $\mathcal{F}$ containing both of them. Since $\mathcal{F}$ contains the Fréchet filter, it is nonprincipal. $\square$

Exercise 6.6 (harder; not used later): Show that in fact, for any infinite set X, the number of nonprincipal ultrafilters on X is $2^{2^{|X|}}$.

6.2. Filters in Topology: An Advertisement.

The night before giving the lecture on ultrafilters and ultraproducts, it occurred to me that ultrafilters might not be part of the working vocabulary of my audience. So I sent out an email advising them to book up on them a little bit and providing a link to some notes on general topology. At the lecture itself, I found out that

[28] One also has the notion of a filter base. But we won't use it, so let's skip the definition.
[29] This would be a filter base, had we defined such a thing. (Sorry!)

indeed most of my audience had not studied filters before.[30]

So here is a quick *précis* of the use of filters and ultrafilters in topology. For more details, please see [GT, Ch. II, §5].

Let $f : X \to Y$ be a function and $\mathcal{F}$ a filter on X. Then the family of subsets $\{f(A) \mid A \in \mathcal{F}\}$ of Y satisfies the finite intersection condition, so is the subbase for a unique filter on Y, which we denote $f(\mathcal{F})$.

Let X be a topological space and x a point of X. Then the set $\mathcal{N}_x$ of neighborhoods of x, i.e., of subsets N of x such that x lies in the interior of N, is a filter on X. It is the principal ultrafilter $\mathcal{F}_x$ iff x is an isolated point of X.

A filter $\mathcal{F}$ on X is said to **converge to x** if $\mathcal{F} \supset \mathcal{N}_x$, i.e., if every neighborhood of x lies in $\mathcal{F}$. We write $\mathcal{F} \to x$. Again, for a trivial example, note that the principal ultrafilter $\mathcal{F}_x$ converges to x no matter what the topology on X is. We say that a filter **converges** if it converges to at least one point. (If X is Hausdorff, a filter converges to at most one point.)

A point $x \in X$ is said to be a **limit point** of a filter $\mathcal{F}$ if the filters $\mathcal{N}_x$ and $\mathcal{F}$ are compatible, i.e., are simultaneously contained in some filter. In other words, x is a limit point of $\mathcal{F}$ if every neighborhood of x meets every element $A \in \mathcal{F}$.

With these definitions, get a theory of convergence via filters paralleling that of sequences in a metrizable (or first countable) space. Here some of the most important tenets of this theory.

Theorem 69.
a) Let X be a topological space. The closure of a subset A of X is the set of all $x \in X$ such that there exists a filter $\mathcal{F}$ on X with $A \in \mathcal{F}$ and $\mathcal{F} \to x$.
b) Let X and Y be topological spaces and $f : X \to Y$ be a map of sets. Then f is continuous iff: for all $x \in X$ and all filters $\mathcal{F}$ on X, $\mathcal{F} \to x$ iff $f(\mathcal{F}) \to f(x)$.
c) Let $X = \prod_i X_i$ be a product of spaces and $\pi_i : X \to X_i$ be the projection map, $\mathcal{F}$ a filter on X and $x = (x_i) \in X$. Then $\mathcal{F} \to x$ iff for all $i \in I$, $\pi_i(\mathcal{F}) \to x_i$.
d) A space X is quasi-compact iff every ultrafilter on X converges.

Each of these statements is straightforward to prove. And they have a nonitrivial consequence.

Exercise 6.7: Deduce from Theorem 69 Tychonoff's theorem, that a product $X = \prod_i X_i$ of nonempty spaces if quasi-compact iff each factor X_i is quasi-compact.

There are of course many proofs of Tychonoff's theorem, but this one has the merit of making the result look completely evident and natural.

[30] I had thought that they were covered in a standard undergraduate topology course. In retrospect, I think they were not covered in *my* undergraduate topology course (which used Munkres' book, as many such courses do) and indeed I may have learned about them for the first time when I started studying model theory in late 2002.

6.3. **Ultraproducts and Los' Theorem.**

The notion of a product of structures is a fundamental one in mathematics. For instance, one has the product of sets, groups, rings, topological spaces, schemes... For many (but not all) of these products, the unifying theme is a certain universal mapping property.

Suppose we have a family $\{X_i\}_{i\in I}$ of models of a theory $\mathcal{T}$. It would be nice, wouldn't it, to be able to define some kind of product model $X = \prod_i X_i$? (This is not much in the way of motivation, but we will soon see just how nice it would be!) Unfortunately, this only works halfway: we may define a product of $\mathcal{L}$-structures, but the product of models of a theory $\mathcal{T}$ need not be a model of $\mathcal{T}$.

Indeed, let $\mathcal{L}$ be a language and $\{X_i\}$ a family of $\mathcal{L}$-structures. Put $X = \prod_i X_i$, the Cartesian product. We may endow X with an $\mathcal{L}$-structure, as follows: for every constant symbol $c \in \mathcal{L}$, we put $c_X = \prod_i c_{X_i}$. For every n-ary function symbol $f \in \mathcal{L}$, we define f_X to be the evident function from $(\prod_i X_i)^n \to \prod_i X_i$, i.e., the one whose i-coordinate is f_{X_i}. Similarly, for every n-ary relation symbol $R \in \mathcal{L}$, we define R_X as the product relation, i.e., $\prod_i R_{X_i} \subset \prod_i X_i^n = (\prod_i X_i)^n$.

Exercise 6.8: If you know and care about such things, show that the product we have defined satisfies the universal mapping property in the sense of category theory.

Thus for instance, if $\mathcal{L}$ is the language of rings, we may take a product of rings. For example, take I to be the set of prime numbers and for $p \in I$, put $R_i = \mathbb{F}_p$. Then the product $\prod_i \mathbb{F}_p$ is again an $\mathcal{L}$-structure (and even a ring). However, suppose $\mathcal{T}$ is the theory of fields. Then each $\mathbb{F}_p$ is a model of $\mathcal{T}$ but the product certainly is not: it is not even a domain.

All this is remedied by passing to a certain quotient of the direct product. To do this, we need an extra ingredient – the crazy part. Namely, we "choose" an ultrafilter $\mathcal{F}$ on the index set I. Then, we define the relation $\sim_{\mathcal{F}}$ on the Cartesian product $X = \prod_i X_i$ by $\{x_i\} \sim_{\mathcal{F}} \{y_i\}$ iff the set of indices $i \in I$ such that $x_i = y_i$ is an element of $\mathcal{F}$. We define the **ultraproduct** $X = \prod_{\mathcal{F}} X_i$ to be the quotient $\tilde{X}/\mathcal{F}$.

Exercise 6.9: Check that $\sim_{\mathcal{F}}$ is indeed an equivalence relation and that the ultraporudct $\prod_{\mathcal{F}} X_i$ is indeed an $\mathcal{L}$-structure in a natural way.

So what is going on here? Magic, I say! Actually, there is one case in which the magic isn't real: there is a little man behind the curtain.

Proposition 70. *Let I be an index set, $i_0 \in I$, and let $\mathcal{F}_{i_0}$ be the principal ultrafilter at i_0. Then the ultraproduct $\prod_{\mathcal{F}_{i_0}} X_i$ is isomorphic to X_{i_0}.*

Exercise 6.10: Prove Proposition 70.

However, when we restrict to nonprincipal ultrafilters, the magic is quite real.

Example 6.11: Let $\mathcal{L}$ be the language of rings, I an index set, $\mathcal{F}$ an ultrafilter on I, for each $i \in I$, let R_i be an integral domain. Then the ultraproduct $R = \prod_{\mathcal{F}} X_i$ is a domain. Indeed, let x and y be elements of R such that $xy = 0$. We must show that $x = 0$ or $y = 0$. Represent x by a sequence $\{x_i\}$ and y by a sequence $\{y_i\}$. Then, to say that $xy = 0$ is to say that the set of indices i such that $x_iy_i = 0$ lies in the filter $\mathcal{F}$: let us call this set A. Let A_1 be the set of indices i such that $x_i = 0$ and let A_2 be the set of indices i such that $y_i = 0$. Since each R_i is a domain, we have $A = A_1 \cup A_2$. By Corollary 65, we have either $A_1 \in \mathcal{F}$ or $A_2 \in \mathcal{F}$, that is, $x = 0$ or $y = 0$: qed.

Now let us show that the ultraproduct $K = \prod_{\mathcal{F}} K_i$ of fields is again a field. So, let $0 \neq x \in K$. We need to show that there exists $y \in K$ such that $xy = 1$. Let $\{x_i\} \in \prod K_i$ be any element representing x, and let $A \subset I$ be the set of indices such that $x_i \neq 0$. Define y to be the element whose i coordinate is: x_i^{-1} if $i \in A$ (so x_i is nonzero in the field K_i and thus has an inverse) and 0 otherwise. Then $x_\bullet y_\bullet$ has i coordinate 1 for all $i \in A$ and 0 otherwise. Hence it is equal to the constant element 1 on a set of indices which lies in $\mathcal{F}$, so $xy = 1$ in the quotient. (Note that we have a lot of leeway in the definition of y_{bullet} – it does not matter at all how we define it at coordinates not lying in A – but all of these elements become equal in the quotient.)

Here is a more interesting example. Let F be the ultraproduct of the finite field $\mathbb{F}_p$. By the above, this is a field. So it has a characteristic – what is it?!?

Case 1: Despite what I said above, it's instructive to consider the case of a principal ultrafilter based at a particular prime p_0. In this case, the ultraproduct is just $\mathbb{F}_{p_0}$, so of course the characteristic is p_0.
Case 2: If $\mathcal{F}$ is nonprincipal, we claim that F has characteristic 0. It suffices to show that for any prime ℓ, $1 + \ldots + 1$ (ℓ times) is not zero. Well, consider the diagonal elements $x_\bullet = \ell$ and $y_\bullet = 0$. What does it mean for $x_\bullet$ and $y_\bullet$ to be equal in the ultraproduct? It means that the set A of primes p such that $\ell = 0$ in $\mathbb{F}_p$ lies in the filter $\mathcal{F}$. But $A = \{p\}$, a finite set, which is not an element of any nonprincipal ultrafilter. Done!

The following result is a vast generalization of these observations. It is often called the **Fundamental Theorem of Ultraproducts**. Nor is the proof difficult; rather it is almost as easy as a proof which proceeds by induction on the complexity of a formula can be. Since we have, somewhat disreputably, not given such a proof thus far,[31] we present the proof of Los' Theorem in all its gory detail.

Theorem 71. *(Los) Let I be an index set, $\mathcal{F}$ an ultrafilter on I, $\{X_i\}_{i \in I}$ an indexed family of $\mathcal{L}$-structures, and put $X = \prod_{\mathcal{F}} X_i$. For any formula ϕ in n unbound variables and $\overline{x} \in X^n$,*

$$X \models \phi(\overline{x})) \iff \{i : X_i \models \phi(\overline{x_i}))\} \in \mathcal{F}.$$

Proof. We prove this by an induction on the complexity of the formulas. First recall that a **term** is the set of $\mathcal{L}$-terms is the smallest set containing the constant symbols of $\mathcal{L}$, the variable names $\{x_i\}_{i=1}^{\infty}$, and for each n-ary function, all expressions of the form $f(t_1, \ldots, t_n)$, where the t_i's are terms.

[31]The fact that truth of quantifier-free formulas is preserved by embeddings of structures was given a somewhat handwavy proof earlier in these notes. What is required to formalize it is precisely an induction on formula complexity

Step 1: Suppose φ is of the form $t_1 = t_2$, where t_1 and t_2 are terms involving n variables $x_1, \ldots, x_n$. For $j = 1, 2$, put

$$g_j(i) = t_j(x_1(i), \ldots, x_n(i)).$$

Then $t_1(x_1, \ldots, x_n) = t_2(x_1, \ldots, x_n)$ as elements of X iff the set of $i \in I$ such that $t_1(x_1(i), \ldots, x_n(i)) = t_2(x_1(i), \ldots, x_n(i))$ is an element of $\mathcal{F}$. This *is* Los' Theorem in this case!

Step 2: Suppose φ is a relation $R(t_1, \ldots, t_n) = R(t)$. Then $R(t_1(\overline{x}), \ldots, t_n(\overline{x}))$ holds in X iff the tuple $(t_1(x), \ldots, t_n(x))$ lies in $R_{X^n} \subset X^n$ iff there exists $\overline{y} = (y_1, \ldots, y_n) \in R_{X^n}$ such that $(t_1(x), \ldots, t_n(x)) = (y_1, \ldots, y_n)$ iff for a set of indices I which lies in $\mathcal{F}$ we have $t(x_i) = y(i)$ iff for a set o findices I which lies in $\mathcal{F}$, $R(t(x_i)) \in R_{X_i}$.

Step 3: Suppose Los' Theorem holds for α and β and $\varphi = \alpha \wedge \beta$. Then $\varphi(x) = \alpha(x) \wedge \beta(x)$ holds in X iff both $\alpha(x)$ and $\beta(x)$ hold in X, iff the sets A_1 (resp. A_2) of indices i such that $\alpha(x_i)$ (resp. $\beta(x_i)$) holds in X_i lie in $\mathcal{F}$ iff (since $\mathcal{F}$ is a filter) the set $A = A_1 \cap A_2$ of indices i such that both $\alpha(x_i)$ and $\beta(x_i)$ hold lies in $\mathcal{F}$ iff the set of indices i such that $\alpha(x_i) \wedge \beta(x_i) = \varphi(x_i)$ holds lies in $\mathcal{F}$.

Step 4: Suppose that Los' Theorem holds for $\varphi(x)$. Then it also holds for $\neg\varphi(x)$. Indeed, $\neg\varphi(x)$ holds in X iff $\varphi(x)$ does not hold in X iff the set A of indices i for which $\varphi(x_i)$ holds in X_i is *not* in $\mathcal{F}$. But the set A' of indices i for which $\neg\varphi(x_i)$ holds in X_i is of course $I \setminus A$, and since $\mathcal{F}$ is an ultrafilter and A is not in $\mathcal{F}$, A' must be in $\mathcal{F}$.[32]

Step 5: Write $x = (x_1, \ldots, x_n)$ and $y = (x_2, \ldots, x_n)$, so $x = (x_1, y)$. Suppose Los' Theorem holds $\psi(x)$; we show that it also holds for $\exists v\ \psi(v, y)$.
This time we handle the two implications separately. First suppose that $\exists v\ \psi(v, y)$ holds in X. Then for some $x = (x_1, y) \in X^n$, $\psi(x)$ holds in X. It follows that the set A of indices i such that $\psi(x(i))$ holds in X_i lies in $\mathcal{F}$. Now the set A' of indices i such that $\exists v\ \psi(v, y(i))$ holds in X_i contains A, so A' lies in $\mathcal{F}$.

Conversely, suppose that the set A of indices i such that $\exists v \psi(v, y(i))$ lies in $\mathcal{F}$. For each such i, choose $x_1(i) \in X_i$ such that $\psi(x_1(i), y(i))$ holds in X_i; for all other indices i, define $x_1(i)$ arbitrarily. There is then an induced element $x_1 = \prod_{\mathcal{F}} x_1(i)$ in the ultraproduct, and then $\varphi(x_1, y)$ holds in X hence so does $\exists v\ \psi(v, y)$. $\square$

Corollary 72. *a) In the setup of Los' Theorem, let $\mathcal{T}$ be an $\mathcal{L}$-theory, and suppose that each X_i is a model of $\mathcal{T}$. Then X is a model of $\mathcal{T}$.*
b) In particular, if for all $i, j \in I$, $X_i \equiv X_j$, then $X \equiv X_i$ for all i.

Proof. a) This is a very special case of Theorem 71: for each sentence $\varphi \in \mathcal{T}$, the set of indices i such that φ_i holds in X_i is the entire index set I, so is certainly an element of $\mathcal{F}$. Thus by Los' Theorem, φ holds in X. Part b) follows immediately. $\square$

One way to enforce $X_i \equiv X_j$ for all indices is simply to choose a single model X of $\mathcal{T}$ and take $X_i = X$ for all i. In this case, we abbreviate $\prod_{\mathcal{F}} X$ to $X^{\mathcal{F}}$, and we say that $X^{\mathcal{F}}$ is an **ultrapower** of X. Thus $X \equiv X^{\mathcal{F}}$, but $X^{\mathcal{F}}$ is guaranteed to

[32] Note that this is the only place in the proof where we use that $\mathcal{F}$ is an ultrafilter!

be a "sufficiently rich" model of $\mathcal{T}$ in a sense that we will not have time to make precise. But, for example, if X is an algebraically closed field, then any nontrivial ultrapower of X is an algebraically closed field of infinite transcendence degree.

Exercise 6.12: Let X be an $\mathcal{L}$-structure and $X^{\mathcal{F}}$ an ultrapower. Show that there is a natural embedding of $\mathcal{L}$-structures $\iota : X \hookrightarrow X^{\mathcal{F}}$ and this embedding is elementary.

6.4. **Proof of Compactness Via Ultraproducts.**

Let $\mathcal{L}$ be a language and $\mathcal{T}$ be a theory such that every finite subset of $\mathcal{T}$ has a model. We wish to show that $\mathcal{T}$ has a model. Formerly, we deduced this as an immediate corollary of Gödel's Completeness Theorem and the finite character of syntactic implication. But, aside from using a proof-theoretic result that we are not going to prove (and is generally regarded as being fundamentally "un-model-theoretic" in nature), this was a proof by contradiction. Much more impressive would be the following head-on attack: for each finite subtheory $\mathcal{T}' \subset \mathcal{T}$, let $X_{\mathcal{T}'}$ be a model of $\mathcal{T}'$. Then using the $X_{\mathcal{T}'}$'s as data, we construct a model X of $\mathcal{T}$.

Prepare to be impressed!

We may of course assume that $\mathcal{T}$ is infinite; otherwise there is nothing to prove. Let I be the set of finite subtheories of $\mathcal{T}$. For $\varphi \in \overline{\mathcal{T}}$, let

$$A(\varphi) = \{\mathcal{T}' \in I \mid \varphi \in \overline{\mathcal{T}'}\},$$

and let $A = \{A(\varphi)\}_{\varphi \in \mathcal{T}}$. Evidently A is a nonempty family of nonempty subsets of I. I claim that moreover A satisfies the finite intersection condition: indeed, for any $\varphi_1, \ldots, \varphi_n \in \mathcal{T}$,

$$\bigcap_{i=1}^{n} A(\varphi) = A(\bigwedge_{i=1}^{n} \varphi_i) \neq \emptyset.$$

Thus, in the terminology of Proposition 67, A is a filter subbase on I. In other words, there is some filter containing A and hence some ultrafilter $\mathcal{F}$ containing I. By hypothesis, for each finite $\mathcal{T}' \subset \mathcal{T}$, there exists at least one $\mathcal{L}$-structure modelling $\mathcal{T}'$: choose one, and call it $X_{\mathcal{T}'}$. Thus $X_{\mathcal{T}'}$ is a family of $\mathcal{L}$-structures indexed by the elements of I, and $\mathcal{F}$ is an ultrafilter on I. So we may form the ultraproduct:

$$X = \prod_{\mathcal{F}} X_{\mathcal{T}'}.$$

We claim that X is a model of $\mathcal{T}$. Indeed, for any $\varphi \in \mathcal{T}$, consider the set J of finite subtheories $\mathcal{T}'$ of $\mathcal{T}$ such that $X_{\mathcal{T}'}$ is a model of φ. It is hard to say exactly what J is (since we chose the models $X_{\mathcal{T}'}$ "at random"), but certainly J contains each finite subtheory $\mathcal{T}'$ such that $\varphi \in \overline{\mathcal{T}'}$, since then φ holds in every model of $\mathcal{T}'$. That is, $J \supset A(\varphi)$; since $A(\varphi) \in \mathcal{F}$ and $\mathcal{F}$ is a filter, $J \in \mathcal{F}$. We are done by Los' theorem.

So the use of ultraproducts gives a quick proof of the Compactness Theorem which, recall, was originally deduced from Gödel's Completeness Theorem and the finite character of syntactic implication. We used the Completeness Theorem and the finite character of syntactic implication at one other key juncture, namely in the proof of Ax's Transfer Principle (Theorem 3.10). We urge every reader to **do the following exercise**.

Exercise 6.13: In the proof of Ax's Transfer Principle, replace all appeals to syntactic considerations by an ultraproduct argument. (Suggestion: use Proposition 68b). That's what it's there for!)

This is a typical phenomenon. Indeed, to the best of my knowledge, in the study of model theory one never needs to use Gödel's Completeness Theorem but can always make do with evident ultraproduct-theoretic analogues.

6.5. **Characterization theorems involving ultraproducts.**

First a result which we could have proven long ago, but is especially appropriate now that we have proved the Compactness Theorem.

Proposition 73.
a) Let $\mathcal{L}$ be a language and $\mathcal{T}_1$, $\mathcal{T}_2$ be two $\mathcal{L}$-theories. Suppose that for an $\mathcal{L}$-structure X, X is a model of $\mathcal{T}_1$ iff X is not *a model of $\mathcal{T}_2$. Then $\mathcal{T}_1$ and $\mathcal{T}_2$ are finitely axiomatizable.*
b) In particular, a class $\mathcal{C}$ is finitely axiomatizable iff both $\mathcal{C}$ and its negation are elementary.

Proof. a) We give two proofs, the first using the Compactness Theorem and the second using ultraproducts as in the proof of the Compactness Theorem. Note that either way, by symmetry it suffices to prove that $\mathcal{T}_1$ is finitely axiomatizable.
First proof: Suppose that $\mathcal{T}_1$ is not finitely axiomatizable. In other words, for every finite subtheory $\mathcal{T}'$ of $\mathcal{T}_1$, there exists an $\mathcal{L}$-structure which is a model of $\mathcal{T}'$ but not of $\mathcal{T}_1$. By hypothesis, this means that $X_{\mathcal{T}'}$ is a model of $\mathcal{T}' \cup \mathcal{T}_2$. But every finite subset of $\mathcal{T} := \mathcal{T}_1 \cup \mathcal{T}_2$ is contained in some $\mathcal{T}' \cup \mathcal{T}_2$ for $\mathcal{T}'$ a finite subset of $\mathcal{T}_1$. Thus the theory $\mathcal{T}$ is finitely satisfiable, hence satisfiable by the Compactness Thorem. But this means that there is a structure X which models both $\mathcal{T}_1$ and $\mathcal{T}_2$, contradiction.
Second proof: Again, suppose $\mathcal{T}_1$ is not finitely axiomatizable, so that for every finite subtheory $\mathcal{T}'$ of $\mathcal{T}_1$ there is a model $X_{\mathcal{T}'}$ of $\mathcal{T}'$ but not of $\mathcal{T}_1$. Again, by our hypothesis $X_{\mathcal{T}'}$ is a model of $\mathcal{T}_2$. Letting I be the set of finite subtheories of $\mathcal{T}_1$, as in the proof of the compactness theorem, there exists an ultrafilter $\mathcal{F}$ on I such that $X = \prod_{\mathcal{F}} X_{\mathcal{T}'}$ is a model of $\mathcal{T}$. On the other hand, each $X_{\mathcal{T}'}$ is a model of $\mathcal{T}_2$, so by Los's Theorem X is also a model of $\mathcal{T}_2$: contradiction. Thus $\mathcal{T}_1$ is finitely axiomatizable. Of course, interchanging the roles of $\mathcal{T}_1$ and $\mathcal{T}_2$ we get that $\mathcal{T}_2$ is finitely axiomatizable.
b) If $\mathcal{C}$ is the class of all models of a finite theory, then certainly it is finitely axiomatizable and indeed is the class of all models of a single sentence φ. But then its negation is the class of all models of $\neg\varphi$ so is also finitely axiomatizable, hence elementary. The converse follows immediately from part a). □

Theorem 74. *Let $\mathcal{C}$ be a class of $\mathcal{L}$-structures.*
a) $\mathcal{C}$ is elementary iff it is closed under ultraproducts and elementary equivalence.
b) $\mathcal{C}$ is finitely axiomatizble iff both $\mathcal{C}$ and its negation are closed under ultraproducts and elementary equivalence.
c) The elementary closure of $\mathcal{C}$ – i.e., the least elementary class containing $\mathcal{C}$ – is the class of all $\mathcal{L}$-structures which are elementarily equivalent to some ultraproduct of elements of $\mathcal{C}$.

Proof. a) It is clear from the definition that an elementary class – i.e., the class of all models of some $\mathcal{L}$-theory $\mathcal{T}$ is closed under elementary equivalence; moreover that an elementary class is closed under passage to ultraproducts is Corollary 72a). Conversely, suppose that $\mathcal{C}$ is a class which is closed under elementary equivalence and passage to ultraproducts. We wish to show that $\mathcal{C}$ is an elementary class. Clearly the only candidate theory is the complete theory of $\mathcal{C}$, i.e., the set of all $\mathcal{L}$-sentences which hold in every element of $\mathcal{C}$. Let X be a model of $\mathcal{T}$. What we need to show is that $X \in \mathcal{C}$. Let Σ be the complete theory of X – so $\Sigma \supset \mathcal{T}$ – and as in the proof of the compactness theorem, let I be the family of all finite subsets of Σ. For each $\mathcal{T}' = \{\varphi_1, \ldots, \varphi_n\} \in I$, there exists $X_{\mathcal{T}'} \in \mathcal{C}$ which is a model of $\mathcal{T}'$, for otherwise the sentence $\neg(\varphi_1 \wedge \ldots \wedge \varphi_n)$ would belong to $\mathcal{T} \setminus \Sigma$, a contradiction. Just as in the proof of the compactness theorem, there exists an ultrafilter $\mathcal{F}$ on I such that the ultraproduct $X' = \prod_{\mathcal{F}} X_i$ is a model of $\mathcal{T}$. By hypothesis, $X' \in \mathcal{C}$. Moreover, since $\mathcal{T}$ is the complete theory of X, this means $X \equiv X'$, and thus by hypothesis $X \in \mathcal{C}$.
Part b) follows immediately from part a) together with Proposition 74. The proof of part c) is similar and left to the reader. □

Theorem 75. *(Keisler-Shelah) Let X and Y be $\mathcal{L}$-structures. TFAE:*
(i) $X \equiv Y$.
(ii) There exists an index set I and an ultrafilter $\mathcal{F}$ on I such that the ultrapowers $X^{\mathcal{F}}$ and $Y^{\mathcal{F}}$ are isomorphic.

This theorem involves delicate set-theoretic considerations. Indeed, it was first proved by H.J. Keisler in 1961 under the assumption of the Generalized Continuum Hypothesis (GCH) and then unconditionally by S. Shelah in 1972. See e.g. [CK90, Thm. 6.1.15] for a proof.

Exercise 6.14: By considering a nontrivial ultraproduct of cyclic groups of prime order, show that the class of simple groups is not an elementary class.

Exercise 6.15 (harder):[33] For all $n \in \mathbb{Z}^+$, we may view S_n as a subgroup of $\operatorname{Aut}(\mathbb{Z}^+)$ by viewing it as the subgroup of permutations of $\mathbb{Z}^+$ which pointwise fix every integer greater than n. With this convention, define the infinite alternating group $A_\infty = \bigcup_{n=1}^{\infty} A_n$ as a subgroup of $\operatorname{Aut}(\mathbb{Z}^+)$.
a) Show that A_∞ is a simple group.
b) Show that no nontrivial ultrapower of A_∞ is simple.
c) Deduce that the class of simple groups is not closed under elementary equivalence.

7. A Glimpse of the Ax-Kochen Theorem

Let $d \in \mathbb{Z}^+$ and $i \in \mathbb{R}^{\geq 0}$. We say that a field K has property $C_i(d)$ if every degree d homogeneous polynomial in at least $d^i + 1$ variables has a nontrivial zero. It is clear that this property is equivalent to a sentence φ_d in the language of fields, so the class $C_i(d)$ of fields is finitely axiomatizable. We also define a field to be C_i if it is $C_i(d)$ for every positive number d. This is the conjunction of the infinitely many sentences φ_d, so C_i is an elementary class.

[33]The material for this exercise was furnished by Simon Thomas as an answer to a question on Math Overflow. Thanks very much to him.

Some relatively elementary facts:

a) a field is C_i for some $i < 1$ iff it is C_0 iff it is algebraically closed.
b) A finite field is C_1 (Chevalley).
c) If K is C_i and L/K has transcendence degree j, then L is C_j (Tsen-Lang).
d) A complete discretely valued field with algebraically closed residue field is C_1 (Lang).
e) The field $\mathbb{F}_q((t))$ is C_2 (Lang).
f) If k is C_i, then $k((t))$ is C_{i+1} (Greenberg).

In particular, combining Chevalley and Greenberg, we find that the locally compact fields of positive characteristic, namely $\mathbb{F}_q((t))$, are C_2.

In view of Greenberg's theorem, it is natural to speculate that a complete discretely valued field with C_i residue field is C_{i+1}. The simplest case of this which is left open by Lang's theorem is that of p-adic fields. Indeed, it was conjectured by E. Artin that a p-adic field is C_2.[34]

Lang's seminal paper [Lan52] contains the sentence "If the residue field of [the CDVF] F is finite, it has been conjectured that F is C_2. We can prove this only in the case of power series fields, leaving the question open in the case of p-adic fields." This was part of Lang's thesis work; I can only imagine his consternation at not being able to prove the p-adic case. Lang and many others tried to prove this throughout the 50's and the first half of the 60's, without success. What was known is that p-adic fields are $C_2(2)$; in other words, a quadratic form over a p-adic field in at least 5 variables is isotropic. This is part of the classical theory of quadratic forms over local fields (and is discussed e.g. in the 8410 course notes). It was also known relatively early on that a cubic form in at least 10 variables has a nontrivial zero (due, I believe, to Davenport). And that was that!

Quite dramatically, in 1966 Guy Terjanian exhibited an anisotropic (i.e.., without nontrivial zero) quartic form over $\mathbb{Q}_2$ in 17 variables [Ter66]. Less well-known is a 1980 theorem of Terjanian [Ter80]: let $d > 2$. Then for all primes p with $p(p-1) \mid d$, there exists an anisotropic degree d form in d^2+1 variables over $\mathbb{Q}_p$. In particular, for no prime p is $\mathbb{Q}_p$ C_2!

On the other hand, James Ax and Simon Kochen proved in 1965 that p-adic fields are "almost C_2". More precisely:

Theorem 76. *(Ax-Kochen Diophantine Theorem) For every positive integer d, there exists a constant $P(d)$ such that for all primes $p > d$, $\mathbb{Q}_p$ is $C_2(d)$.*

Note that their proof gives precisely zero information about the constant $P(d)$, but Terjanian's work gives some *lower bounds* on it. To the best of my knowledge, for $d \geq 4$ no explicit upper bounds on $P(d)$ are known.

But this theorem reeks of model theory, and in particular of Ax's Transfer Principle. Here is what they actually proved:

[34] Or so people say; I am not sure if Artin's conjecture appears in written form.

Theorem 77. *(Ax-Kochen Transfer Principle) Let $\mathcal{F}$ be a nonprincipal ultrafilter on the set $\mathcal{P}$ of prime numbers. Then the fields $\prod_{\mathcal{F}} \mathbb{Q}_p$ and $\prod_{\mathcal{F}} \mathbb{F}_p((t))$ are elementarily equivalent.*

Exercise 6.16: Deduce Theorem 76 from Theorem 77. (Use Proposition 68b).)

The proof of Theorem comes from a penetrating analysis of the model theory of Henselian valued fields which is interesting and useful in its own right. (Sample result: the embedding from a Henselian valued field of characteristic 0 to its completion is an elementary embedding.) This would be a nice topic for a second half-course on applied model theory!

References

[Ax68] J. Ax, *The elementary theory of finite fields.* Ann. of Math. (2) 88 (1968), 239–271.

[AK65] J. Ax and S. Kochen, *Diophantine problems over local fields. I.* Amer. J. Math. 87 (1965) 605–630.

[BS] J.L. Bell and A.B. Slomson, *Models and ultraproducts: An introduction.* North-Holland Publishing Co., Amsterdam-London 1969

[BCR] J.Bochnak, M. Coste and M.-F. Roy, *Real algebraic geometry.* Translated from the 1987 French original. Revised by the authors. Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], 36. Springer-Verlag, Berlin, 1998.

[BR66] A.R. Bernstein and A. Robinson, *Solution of an invariant subspace problem of K. T. Smith and P. R. Halmos.* Pacific J. Math. 16 (1966) 421–431.

[Bor69] A. Borel, *Injective endomorphisms of algebraic varieties.* Arch. Math. (Basel) 20 (1969), 531–537.

[CK90] C.C. Chang and H.J. Keisler, *Model theory. Third edition.* Studies in Logic and the Foundations of Mathematics, 73. North-Holland Publishing Co., Amsterdam, 1990.

[FT] P.L. Clark, *Field Theory*, lecture notes available at `http://math.uga.edu/~pete/FieldTheory.pdf`

[GT] P.L. Clark, *General Topology*, lectures notes available at `http://math.uga.edu/~pete/expositions.html`

[Hod] W. Hodges, *A shorter model theory.* Cambridge University Press, Cambridge, 1997.

[Lam] T.Y. Lam, *Introduction to quadratic forms over fields.* Graduate Studies in Mathematics, 67. American Mathematical Society, Providence, RI, 2005.

[Lan52] S. Lang, *On quasi algebraic closure.* Ann. of Math. (2) 55, (1952). 373–390.

[Mar] D. Marker, *Model theory. An introduction.* Graduate Texts in Mathematics, 217. Springer-Verlag, New York, 2002.

[Mat] Ju. V. Matijasevic, *The Diophantineness of enumerable sets.* (Russian) Dokl. Akad. Nauk SSSR 191 1970 279–282.

[Poi] B. Poizat, *A course in model theory.* (English summary) An introduction to contemporary mathematical logic. Translated from the French by Moses Klein and revised by the author. Universitext. Springer-Verlag, New York, 2000.

[ARob55] A. Robinson, *On ordered fields and definite functions.* Math. Ann. 130 (1955), 275–271.

[JRob49] J. Robinson, *Definability and decision problems in arithmetic.* J. Symbolic Logic 14 (1949), 98–114.

[Rud] W. Rudin, *Injective polynomial maps are automorphisms.* Amer. Math. Monthly 102 (1995), no. 6, 540–543.

[Ser09] J.-P. Serre, *How to use finite fields for problems concerning infinite fields.* Arithmetic, geometry, cryptography and coding theory, 183–193, Contemp. Math., 487, Amer. Math. Soc., Providence, RI, 2009. See also `http://arxiv.org/abs/0903.0517`.

[Tar48] A. Tarski, *A Decision Method for Elementary Algebra and Geometry.* RAND Corporation, Santa Monica, Calif., 1948.

[Ter66] G. Terjanian, *Un contre-exemple à une conjecture d'Artin.* (French) C. R. Acad. Sci. Paris Sr. A-B 262 (1966), A612.

[Ter80] G. Terjanian, *Formes p-adiques anisotropes.* (French) J. Reine Angew. Math. 313 (1980), 217–220.