Model Theory

Mark Sweeney

1 Introduction

The goal of this paper is to provide a reasonably complete exposition of the Ax-Kochen theorem and its (partial) resolution of Artin's conjecture, in addition to a small selection of other applications of model theory and mathematical logic to algebra. This is based largely on Chang and Keisler's *Model Theory* and Cherlin's *Model Theoretic Algebra: Selected Topics*. Wherever possible, some intuition, motivation, or other plausibility argument accompanies the more difficult results.

2 The Completeness Theorem

This section aims to introduce the fundamental definitions of model theory and apply them to a brief proof of the Completeness Theorem for first-order logic, largely following the presentation in [1].

2.1 Basic Definitions

Definition 2.1. A **language** is a collection $\mathcal{L}$ of symbols and associated natural numbers of the following kinds:

1. a set $\mathcal{C}$ of constants
2. a set $\mathcal{F}$ of functions, and a natural number n_f for each function f (its arity)
3. a set $\mathcal{R}$ of relations, and another natural number n_R for each relation R (also its arity)

Additionally, we define $||\mathcal{L}||$ to be ω if $\mathcal{L}$ is finite, and $|\mathcal{C} \cup \mathcal{F} \cup \mathcal{R}|$ otherwise.
On top of languages, we can build mathematical objections:

Definition 2.2. For a fixed language $\mathcal{L}$, an **$\mathcal{L}$-structure** $\mathcal{M}$ is given by a set M (the **universe**) with an *interpretation*, which gives the "meaning" of symbols in the language $\mathcal{L}$ in M (e.g. constants are assigned to members of M). Typically, we use $c^{\mathcal{M}}$, $f^{\mathcal{M}}$, and $R^{\mathcal{M}}$ to denote the interpretation of constants, functions, and relations on M (in the natural way).
Generally: $\mathcal{M}$ has M as its underlying set.

Isomorphisms between models are exactly what they sound like: sentence- and truth-preserving functions between two models. A more logically relevant notion is that of elementary equivalence:

Definition 2.3. Two models, M and N, in a language $\mathcal{L}$ are said to be elementarily equivalent, $M \equiv N$, when $M \vDash \phi$ if and only if $N \vDash \phi$ for any sentence ϕ of $\mathcal{L}$.

It is possible for nonisomorphic models to be elementarily equivalent, though certainly isomorphic models are elementarily equivalent.

The distinction between theories (syntax) and models (semantics) is important. Two relations are important for this: syntactic consequence ($\vdash$) and semantic consequence ($\vDash$). Given a set of $\mathcal{L}$-sentences T and sentence ϕ (known as a theory) we say that $T \vdash \phi$ if there is a (formal) proof of ϕ from T. On the other hand, for a model, we say that $\mathcal{M} \vDash \phi$ if ϕ is true in $\mathcal{M}$. One can also say for a theory T that $T \vDash \phi$ if ϕ is true in every model of T.

On the other hand, the connections between syntax and semantics are important too. The language plays a crucial role in model theory because its expressiveness determines what can be said and proven about models. For instance, our goal is to prove the Ax-Kochen theorem in a particular language, that of valued fields, but an interesting alternative proof can be given by quantifier elimination, which is possible if the language is expanded to include a predicate for "is an n^{th} power".

Ideally, we would like syntactic and semantic consequence to agree: everything which is provable should be true, and everything which is true should be provable. The first, soundness, just requires that the derivation rules be truth-preserving, which is easily seen to be true for first order logic. The second, completeness, is somewhat deeper, but we will see that it is true in first-order logic.

2.2 Completeness

One of the first and most substantial theorems of mathematical logic is the completeness theorem, which says that a set of sentences is consistent if and only if it has a model. It has a variety of useful corollaries.

We will prove it here using Henkin's construction, following the proofs in [5] [1].

This proof proceeds in two steps: first, we show that any theory can be extended to a maximal theory, and that any theory can be extended to one with the witness property (all consistent, of course). It is more or less straightforward to show that such a theory has a model. With a small amount of extra work, it can be shown that this model is a model of the original theory (in the second step, the language is expanded, introducing a small subtlety).

Definition 2.4. A theory T is said to be **maximal** if, for any sentence ϕ, either ϕ or $\neg\phi$ is in T.

Lemma 2.1. *Suppose T is a consistent theory. Then there exists a maximal (consistent) theory T' extending T.*

Proof. Let Σ be the set of consistent theories containing T. Since $T \in \Sigma$, it is nonempty. Suppose that we have a chain in Σ indexed by I:

$$T_1 \subseteq T_2 \subseteq ...$$

Let $S = \cup_{i \in I} T_i$. Obviously, S contains T; moreover, S is consistent: since proofs are finite, only finitely many T_i are needed to prove a contradiction, but then the largest of those T_i would not be consistent, a contradiction (as all members of Σ are consistent). By Zorn's lemma, Σ has maximal elements.

Let T' be maximal in Σ, and ϕ some sentence. Suppose $\phi \notin T'$. By maximality, it must be that $T' \cup \{\phi\}$ is inconsistent. Certainly $\neg\phi$ is a consequence of any inconsistent theory, and hence

$$T' \vdash \phi \to \neg\phi$$

$$T' \vdash \neg\phi \vee \neg\phi$$

$$T' \vdash \neg\phi$$

This means $T' \cup \{\neg\phi\}$ is consistent, and hence $\neg\phi \in T'$ by maximality. □

Definition 2.5. A theory T has the **witness property** if there exists a set of constants $C \subseteq \mathcal{C}$ in $\mathcal{L}$ such that: for every sentence ϕ of $\mathcal{L}$, there is a constant $c \in C$ such that

$$T \vdash (\exists x \phi) \to \phi(c)$$

The set C is called a set of witnesses for T.

Lemma 2.2. *Given a (consistent) theory T in the language $\mathcal{L}$, there exist a theory T' and language $\mathcal{L}'$ extending T and $\mathcal{L}$ respectively, such that T' has the witness property.*

Proof. Fix $\alpha = ||\mathcal{L}||$. For each $\beta < \alpha$, add a constant c_β to $\mathcal{L}$ (distinct from any extant constants) to obtain a language $\mathcal{L}'$ - note that $||\mathcal{L}|| = ||\mathcal{L}'||$. Order the sentences ϕ of $\mathcal{L}'$ (of which there are α) such that for each sentence ϕ_β there is a corresponding constant d_β. The constant d_β can (and must) be chosen so that it does not occur in ϕ_γ for any $\gamma < \beta$; because any sentence has at most finitely many constants in it, we may just let d_β be the smallest constant not yet used.

We can then create a chain

$$T = T_0 \subseteq T_1 \subseteq ...$$

Such that:

$$T_{\beta+1} = T_\beta \cup \{(\exists x_\beta \phi_\beta) \to \phi_\beta(d_\beta)\}$$

And for any nonzero limit ordinal ζ, $T_\zeta = \cup_{\beta<\zeta} T_\beta$.

Proceeding inductively, we first show that $T_{\beta+1}$ is consistent when T_β is; if not:

$$T_\beta \vdash \neg\left((\exists x_\beta \phi_\beta) \to \phi_\beta(d_\beta)\right)$$

$$T_\beta \vdash (\exists x_\beta \phi_\beta) \wedge \neg\phi_\beta(d_\beta)$$

$$T_\beta \vdash (\exists x_\beta \phi_\beta) \wedge \neg\phi_\beta(x_\beta)$$

$$T_\beta \vdash (\exists x_\beta \phi_\beta) \wedge \neg(\exists x_\beta \phi_\beta),$$

contradicting the consistency of T. (to move from line 2 to 3, note that d_β does not appear in T_β)

It is true in general that $\cup_{\beta<\zeta} T_\beta$ is consistent if each member of the union is consistent. Otherwise, since only finitely many statements are used to prove a contradiction, some T_β would contain all the statements necessary to prove that contradiction, but all the T_β are consistent, so this is not possible.

Now simply let $T' = \cup_{\beta<\alpha} T_\beta$. Just like the other union, it remains consistent. By construction it has the witness property, as each sentence ϕ_β is assigned a witness d_β in the above construction. □

Note that the witness property is preserved under extensions to a theory, since it is quantified over all sentences of the language. This means we can start with T and $\mathcal{L}$, obtain $T', \mathcal{L}'$ which have the witness property, and finally extend that to $\bar{T}, \bar{\mathcal{L}}$ in which $\bar{T}$ is a maximal consistent theory with the witness property. Since $T \subseteq \bar{T}$, any model of $\bar{T}$ is a model of T. We will see that for theories such as $\bar{T}$, the construction of a model is fairly straightforward.

Lemma 2.3. *If T is a maximal consistent set of sentences in $\mathcal{L}$ with a set of C from $\mathcal{L}$, then T has a model. In fact, every element of the model will be the interpretation of some constant of C.*

Proof. Define the relation $\sim$ on C:

$$a \sim b \text{ if and only if } a = b \in T$$

Since T is maximal, it is easy to verify that it is in fact an equivalence relation.

From this we can define our model $\mathcal{M}$. Let the universe be $M = C/\sim$, using bars to denote equivalence classes. The interpretation is fairly straightforward:

1. For each constant, let $c^{\mathcal{M}} = \bar{c}$
2. Let f be an n-place function in $\mathcal{L}$. For each n-tuple $(a_1^{\mathcal{M}}, ..., a_n^{\mathcal{M}})$, note first that:

   $$T \vdash \exists t = f(a_1, ..., a_n)$$

   Hence by the witness property, there is a constant a such that $f(a_1, ..., a_n) = a$. We will say that

   $$f^{\mathcal{M}}(a_1^{\mathcal{M}}, ..., a_n^{\mathcal{M}}) = a^{\mathcal{M}}$$

   We must still prove f is well-defined; suppose that $(a_1^{\mathcal{M}}, ..., a_n^{\mathcal{M}}) = (b_1^{\mathcal{M}}, ..., b_n^{\mathcal{M}})$. This means that

   $$T \vdash a_1 = b_1, a_2 = b_2, ..., a_n = b_n$$

   From the identity axioms, $a = f(a_1, ..., a_n) = f(b_1, ..., b_n)$ and so $f(a_1^{\mathcal{M}}, ..., a_n^{\mathcal{M}}) = f(b_1^{\mathcal{M}}, ..., b_n^{\mathcal{M}})$.
3. For each n-ary relation R in $\mathcal{L}$, we make a similar definition. Let $(a_1^{\mathcal{M}}, ..., a_n^{\mathcal{M}})$ be an n-tuple. We will say that

   $$(a_1^{\mathcal{M}}, ..., a_n^{\mathcal{M}}) \in R^{\mathcal{M}} \text{ if and only if } T \vdash R(a_1, ..., a_n)$$

   Exactly as we saw for functions, this is well-defined due to the identity axioms and our construction of the equivalence classes.

Finally, it remains to show that $\mathcal{M}$ is truly a model of T. Since T is complete, this amounts to showing that$\mathcal{M} \vDash \phi$ if and only if $\phi \in T$.

1. First we will establish the claim for sentences of the form $t_1 = t_2$, where t_1 and t_2 are closed terms.

   Suppose $T \vDash t_1 = t_2$; since $T \vDash \exists x(x = t_1)$ and $\exists x(x = t_2)$, the witness property gives us constants c_1 and c_2 such that $t_1 = c_1$ and $t_2 = c_2$, so $(c_1 = c_2) \in T$ also. By the remarks on functions above, this is enough to conclude that $M \vDash t_1 = c_1 = c_2 = t_2$

   Conversely, if $T \vDash t_1 \neq t_2$ then we obtain a pair of constants $c_1 \neq c_2$ such that $c_1 = t_1$ and $c_2 = t_2$. Then $\mathcal{M} \vDash t_1 = c_1 \neq c_2 = t_2$, as desired.

2. The same reasoning extends to predicates in general: if $T \vdash P(t_1, ..., t_n)$, the witness property allows us to replace each term with a constant, and then appeal to the construction above, and likewise if $P(t_1, ..., t_n) \notin T$ by completeness.

3. Now we extend this to logical operations by induction on the length of formulas.

   If $\not\phi \in T$, then $\phi \notin T$, which is true if and only if $\mathcal{M} \nvDash \phi$ by induction.

   If $\phi \wedge \psi \in T$, then ϕ and ψ are in T, hence $\mathcal{M}$ satisfies both ϕ and ψ, and so it satisfies $\phi \wedge \psi$.

4. Finally, we must treat quantifiers.

   Suppose $\phi = \exists x\psi$. If $\phi \in T$, then by the witness property, there exists a c such that $\psi(c) \in T$, and so $\mathcal{M} \vDash \psi(c)$ (using the fact that ψ is a shorter sentence), hence $\mathcal{M} \vDash \exists x\psi$.

   Conversely, if $\mathcal{M} \vDash \phi$, then $\mathcal{M} \vDash \psi(c)$ for some c. Here again, ψ is a shorter sentence, so $\psi(c) \in T$, and hence $T \vDash \exists x\psi$.

□

This establishes the results necessary to prove the completeness theorem:

Theorem 2.4. *A set of sentences Σ in $\mathcal{L}$ is consistent if and only if Σ has a model.*

Proof. If Σ were inconsistent, $\Sigma \vdash \exists x(x \neq x)$, and so any model of Σ satisfies that sentence, but there are no models of $\exists x(x \neq x)$.

If Σ is consistent, extend to Σ' and $\mathcal{L}'$ such that Σ' is complete and has witnesses. By the previous lemma, there exists a model $\mathcal{M}$ of Σ' over $\mathcal{L}'$. However, we can always restrict the language to $\mathcal{L}$. Since $\Sigma \subseteq \Sigma'$, and the only symbols we lose in passing from $\mathcal{L}'$ to $\mathcal{L}$ are constants not mentioned in Σ, we see that $\mathcal{M}$ is a model of Σ. □

2.3 Applications and Examples

This result, as well as the particular construction used above, has a number of surprising, interesting, and useful consequences, summarized below. The completeness theorem is useful for constructing models satisfying certain properties.

Theorem 2.5 (Compactness Theorem)**.** *If every finite subset of Σ is satisfiable, then Σ is satisfiable.*

Proof. By completeness, it suffices to show Σ is consistent. Any contradiction derivable from Σ would depend only on some finite subset Δ of Σ. But Δ is satisfiable, and hence cannot prove a contradiction, so Σ is consistent. □

Theorem 2.6 (Downward Löwenheim-Skolem-Tarski)**.** *Every consistent theorem T over a language $\mathcal{L}$ has a model of cardinality at most $||\mathcal{L}||$.*

Proof. In the construction above, we obtain a model whose elements are all equivalence classes of constants of an extension of $\mathcal{L}$ to a language in which T has witnesses. Each constant was either in $\mathcal{L}$ or is associated to a formula of $\mathcal{L}$, of which there are $||\mathcal{L}||$, so taken together there are at most $||\mathcal{L}||$ possible equivalence classes. □

Theorem 2.7 (Upward Löwenheim-Skolem-Tarski)**.** *Every consistent theory T over a language $\mathcal{L}$ with one infinite model has a model of cardinality α for any $\alpha \geq ||\mathcal{L}||$.*

Proof. Enrich $\mathcal{L}$ with α distinct constants $\{c_\beta \mid \beta < \alpha\}$ and expand T by the sentences $\{c_\beta \neq c_\gamma \mid \beta < \gamma < \alpha\}$. The new theory is finitely satisfiable: any finite subset involves only finitely many of the c_βs, and T has an infinite model, so it is always possible to interpret them as distinct elements of that model.

As a result, the larger theory has a model of cardinality at most $||\mathcal{L} \cup \{c_\beta \mid \beta < \alpha\}|| \leq \alpha$, but each of the constants is distinct, so the model has at least cardinality α. □

In some sense the last two theorems above, the Upward and Downward Löwenheim-Skolem-Tarski Theorems, are a negative result about the strength of first order logic: they show that there is no first order sentence which fixes the (infinite) cardinality of a model. Unfortunately, there is a similar result for finiteness; while any particular finite cardinality can be picked out with a first order sentence, there is no sentence which has all and only the finite models.

Theorem 2.8. *If a theory T has finite models of arbitrary order, then it has an infinite model.*

Proof. Add constants $c_1, c_2, ...$ to the language, and let $T' = T \cup \{c_i \neq c_j \mid i \neq j\}$. Any finite subset of T' is realized by a sufficiently large model of T, which can interpret all the c_i as distinct elements, so T' is consistent and has a model. Any model of T' has at least ω elements, and is also a model of T, so T has an infinite model. □

Many other properties involving finiteness also cannot be enforced by a first order sentence, such as the property of being a torsion group:

Theorem 2.9. *Suppose T extends the theory of groups in the language of groups. If T has models with elements of arbitrarily high order, then T has a model with an element of infinite order.*

Proof. For any n, there exists a model G_n of T which contains an element g_n of order at least n. Expand $\mathcal{L}$ to $\mathcal{L} \cup \{c\}$. Let ϕ_n be the sentence given by

$$\underbrace{x * x * x * ... * x}_{n \text{ times}} = e$$

Let $T' = T \cup \{\neg\phi_n(c) \mid n \in \mathbb{N}\}$. Then T' is finitely satisfiable: if $\Delta \subseteq T'$, it involves at most finitely many of the ϕ_n, so there is some N such that $\neg\phi_n(c)$ is not in Δ for $n \geq N$. Then G_N is a model of Δ with c interpreted as g_N.

Thus, T' is consistent, and so it has a model, which will also be a model of T. The element corresponding to 'c' in that model will have infinite order. □

3 Algebraically Closed Fields

One of the nicest theories available for study is that of algebraically closed fields. This field can be axiomatized in the language of rings with identity, $\mathcal{L}_r = \{0, 1, +, \times\}$. The following results are presented in [5]

Clearly, being a field is first-order. Algebraic closure can be axiomatized by the following family of first-order sentences:

$$\phi_n : \forall a_0 \forall a_1 ... \forall a_n \exists x (a_0 + a_1 x + a_2 x^2 + ... + a_n x^n = 0$$

Characteristic is also a first-order property, as if we let

$$\psi_p : \underbrace{1 + 1 + ... + 1}_{p \text{ times}} = 0,$$

then a field with characteristic p satisfies ψ_p, while a field of characteristic zero satisfies $\{\neg\psi_p\}$.

Let ACF be the theory of fields plus the axioms for algebraic closure. For $p \neq 0$, let $ACF_p = ACF \cup \{\psi_p\}$, and let $ACF_0 = ACF \cup \{\neg\psi_p\}$.

One reason algebraically closed fields behave so well is the following algebraic result:

Theorem 3.1. *Any two algebraically closed fields of the same characteristic and transcendence degree α are isomorphic.*

Sketch of proof: take L and K with prime field k, and let L_0 and K_0 be maximal purely transcendental subextensions of L and K respectively. Since they have the same transcendence degree, $L_0 \cong K_0$. Then note that L and K are the algebraic closures of L_0 and K_0, respectively, hence are isomorphic.

Note that for an uncountable algebraically closed field, its transcendence degree equals its cardinality.

Corollary 3.1.1. *For any sentence ϕ, either $ACF_p \vdash \phi$ or $ACF_p \vdash \neg\phi$ (clearly not both)*

Proof. If not, then both ϕ and $\neg\phi$ are consistent with ACF_p, and hence we can find models L and L' of $ACF_p + \phi$ and $ACF_p + \neg\phi$, respectively.

All algebraically closed fields are infinite: since $ACF_p + \phi$ and $ACF_p + \neg\phi$ each have an infinite model, they have infinite models of all cardinalities. Let $\kappa > |L|, |L'|$, and K, K' models of $ACF_p + \phi$ and $ACF_p + \neg\phi$ of order κ. Since K and K' are uncountable algebraically closed fields of the same cardinality, they must be isomorphic. As a consequence, they satisfy the same set of first-order sentences - but that means they each satisfy both ϕ and $\neg\phi$, a contradiction. □

This leads us to a natural correspondence between algebraically closed fields of characteristic $p > 0$ and characteristic zero: as p tends to infinity, the two become more similar.

Theorem 3.2. *Let ϕ be a sentence in $\mathcal{L}_r$. Then TFAE:*

1. *ϕ is true in $\mathbb{C}$*
2. *$ACF_0 \vdash \phi$*
3. *For all sufficiently large p, $ACF_p \vdash \phi$*
4. *For arbitrarily large primes p, $ACF_p \vdash \phi$*

Proof. $(1 \Leftrightarrow 2)$ By the preceeding corollary, anything true in one algebraically closed field of characteristic zero is true in all of them. By the completeness theorem, $ACF_0 \vdash \phi$. The converse is obvious.

$(2 \Rightarrow 3, 4)$ Any proof of ϕ from the axioms of ACF_0 can only use finitely of the statements $\neg\psi_p$, and hence is a valid argument in any other characteristic.

$(\neg 2 \Rightarrow \neg 3, \neg 4)$ By the previous corollary, if $ACF_0 \nvdash \phi$, then $ACF_0 \vdash \neg\phi$. But by the same argument given before, this proof will work for all but finitely many characteristics, and hence for all sufficiently large primes. □

A cute application of that fact is the following theorem:

Theorem 3.3 (Ax)**.** *If $f : \mathbb{C}^n \to \mathbb{C}^n$ is an injective polynomial map, it is also surjective.*

Proof. We will verify that the result holds in finite characteristic, check that it is first-order, and then apply the previous theorem.

Suppose that $f : (F_p^a)^n \to (F_p^a)^n$ is injective. Let $\{a_1, ..., a_m\}$ be all the coefficients appearing in f. Let $(b_1, ..., b_n) \in (F_p^a)^n$. Consider the fields $k = F_p(a_1, ..., a_m, b_1, ..., b_n)$. Since that field is closed under ring operations, we see that f takes k^n to k^n. As it is injective and k^n is finite, it must also be surjective. Hence $(b_1, ..., b_n)$ is in the image of k under f, and hence in the image of $(F_p^a)^n$ under f.

Since the result is true in all finite fields, we can apply the previous theorem as long as the claims are first-order. A sentence of the following form captures "injective implies surjective"

$$(\forall f : K^n \to K^n)\left(\bigwedge_{i=1}^{n} f_i(X) = f_i(X)) \to X = Y\right) \to ((\forall Y \in K^n)(\exists X \in K^n) f(X) = Y$$

Fixing n and the degree of f, It is possible to expand out quantification in f to quantification over its coefficients, and quantification over n-tuples to quantification over their components. This leads to a family of first-order sentences which capture the statement of the theorem. □

4 Ultraproducts

In proving the completeness theorem above, we constructed a model out of constants, and saw that the construction itself had interesting applications (though many of a negative nature). Another important kind of construction in model theory is that of the ultraproduct. Essentially, the ultraproduct will allow us to construct the "average" model from a family of models. This construction can give an alternative proof of the compactness theorem, though the control it gives over the cardinality of the resulting model is not fine enough to prove some of the theorems stated above.

Definition 4.1. Given a set X, a subset F of $\mathcal{P}(X)$ is called a **filter** on X if:

1. F is nonempty.
2. F is closed under finite intersections.
3. For every $x \in F$ and $y \in \mathcal{P}(X)$, $x \subseteq y$ implies $y \in F$

We call F an **ultrafilter** if it is maximal (i.e. the only properly larger filter is all of $\mathcal{P}(X)$). A simple application of Zorn's Lemma shows that any proper filter can be extended to an ultrafilter.

Proposition 4.1. *Equivalently, a filter F is an ultrafilter over X iff $f \in F$ if and only if $X - f \notin F$.*

Proof. Suppose F is an ultrafilter. If $f \in F$, then we cannot have $X - f \in F$, as then $\emptyset = f \cap (X - f) \in F$ (property 2), which would then require that $F = \mathcal{P}(X)$ (property 3) contradicting the fact that F is proper. On the other hand, if $g \notin F$, consider the filter F' generated by F and $X - g$. If we can show that F' is proper, then maximality will require $F = F'$ and hence $X - g \in F$. Suppose otherwise; then there would be a set $f \in F$ such that

$$\emptyset = f \cap (X - g).$$

This would mean $f \subseteq g$, which by property 3 requires that $g \in F$, contrary to assumption.

Conversely, suppose F is such that $f \in F$ iff $X - f \notin F$. Since F is nonempty, it contains X, and hence does not contain $X - X = \emptyset$, so it is a proper filter. Any properly bigger filter would contain both a and $X - a$ for some a, and so that larger filter would contain $\emptyset$, making it improper. □

This whole development is necessary to define the titular subject of this section, the ultraproduct. Ultrafilters behave like measures, and essentially the ultraproduct is a way of turning an indexed collection of models into an "average" model with respect to some ultrafilter over the indexing set. In our proof of the Ax-Kochen theorem, this construction will be used to show that the "average" $\mathbb{Q}_p$ is the same as the "average" $\mathbb{F}_p((t))$, which will imply that any statement true for all but finitely many $\mathbb{Q}_p$ is true for all but finitely many $\mathbb{F}_p((t))$ and vice-versa.

Definition 4.2. Let $\{M_i\}_{i \in I}$ be a collection of models indexed by I, and let D be an ultrafilter on I. The ultraproduct with respect to D is defined as

$$\prod_D M_i = \prod_{i \in I} M_i / \sim_D$$

where the relation $\sim_D$ is defined as

$$(a_i) \sim_D (b_i) \Leftrightarrow \{i \in I \mid a_i = b_i\} \in D$$

Clearly, the language $\mathcal{L}$ has a natural interpretation in the ultraproduct; each constant is interpreted as the corresponding sequence of constants, while functions and relations are defined pointwise.

To make this more concrete, we will consider a particularly useful special case involving fields. Here, ultrafilters will correspond to maximal ideals, and principal ultrafilters with the principal maximal ideals. This example also helps to motivate our next theorem.

Consider the ring $R = \prod_{p \text{ prime}} \mathbb{F}_p$, and let D be an ultrafilter on the set of primes. It turns out that $\prod_D \mathbb{F}_p$ will always be a field, and clearly there is a surjection from R to this field. The kernel will be a maximal ideal. There are two sorts of maximal ideal in R, principal ideals of the form $((1, 1, ..., 1, 0, 1, ...))$

whose quotients are isomorphic to some $\mathbb{F}_p$, and nonprincipal ideals which contain the ideal $\oplus F_p$. Let M be a maximal ideal of R containing $\oplus F_p$, and consider the quotient ring R/M. For any such maximal ideal, we see that R/M is characteristic zero - its identity element is the equivalence class of $(1, 1, ...)$. If it happens that $(p, p, ...) \in M$ for some prime p, then we know that modulo any prime not equal to p, p is a unit, but M will contain some sequence $(0, ..., 0, 1, 0, ...)$ with a 1 in the component corresponding to $\mathbb{F}_p$, but the sum of that and $(p, p, ...)$ will be a unit, contradicting the fact that M is proper.

It turns out that there is a slightly deeper reason that R/M is characteristic zero. For any ultraproduct D, a first-order sentence is true in it if and only if the set of indices of models for which it holds is in D (this is Los's theorem). In the case above, the maximal ideal M corresponds to a ultrafilter containing all the cofinite subsets of sets of primes. Since the first-order statement "the characteristic is not p" is true in all but finitely many cases (all but one) that sentence is true in the ultraproduct, and so the characteristic is not p for any prime p.

Theorem 4.2 (Los)**.** *Let M be an ultraproduct $\prod_D M_i$ with index set I. Then $M \vDash \phi$ if and only if $\{i \in I \mid M_i \vDash \phi\} \in D$.*

The proof is a straightforward proof by induction on formulas (the construction of the ultraproduct ensures than any possible disagreement is absorbed into $\sim_D$), and the only step in which the maximality of the ultrafilter is used is for negation. Considering the previous example involving fields, we can see why extra structure could be needed for this case: we have a surjective homomorphism from the infinite direct product to the ultrafilter, and homomorphisms preserve all negation-free first-order sentences.

Los's Theorem can be used to construct nonstandard models of arithmetic. Take the collection of models $\{\mathbb{N}_i \mid i \in \mathbb{N}\}$, and let D be a nonprincipal ultrafilter. Then $N = \prod_D N_i$ is elementarily equivalent to N by Los's theorem, but whenever D is a nonprincipal ultrafilter, the ultraproduct will always be uncountable, so $N \ncong \mathbb{N}$. While the completeness theorem can also produce nonstandard models of arithmetic, even countable ones, the ultraproduct can feel somewhat more concrete. For instance, the equivalence class of $(1, 2, 3, 4, ...)$ is readily seen to be a nonstandard natural number, as Los's theorem shows that it exceeds $(n, n, ...)$ (which corresponds to a standard natural number) for any n. We can even come up with nostandard primes in this model - $(2, 3, 5, 7, ...)$ is prime in all components, so it is also a prime in this model, but it clearly cannot equal any standard prime. Nonstandard primes can be chosen with interesting number-theoretic properties, such as $(5, 13, 17, ...)$ which is congruent to 1 mod 4 in each component and hence satisfies the same in N.

In fact, Abraham Robinson proves the Ax-Kochen theorem in a special case (for $\mathbb{Q}_p$ and $\mathbb{F}_p((t))$ by proving that $\mathbb{Q}_p \cong \mathbb{F}_p((t))$ for nonstandard primes p using the methods of nonstandard analysis [4, 6]. Recalling the previous remarks that this result is proven by showing that $\prod_D Q_p \cong \prod_D \mathbb{F}_p$, his proof in a sense moves the ultraproduct into the index.

5 Saturation

Saturation is the most model-theoretically challenging notion involved in the proof. Roughly, saturated models are large models which contain any sort of "ideal" element which can be described in the language and are compatible with the theory in question. This property makes it easier to show that saturated models are isomorphic, as we only need to be able to describe the elements we need to extend the map in order to know they exist. To make all of this precise will require some definitions.

Definition 5.1. Let M be a model and X a subset of M. The expansion of M, written $(M, x)_{x \in X}$ is the model M in the larger language $\mathcal{L}_X = \mathcal{L} \cup \{c_x \mid x \in X\}$, interpreting each c_x as x.

This richer language allows the behavior of the model with respect to X to be pinned down more closely by first-order sentences.

Definition 5.2. A 1-type $\Sigma(y)$ is a maximal consistent set of sentences in $\mathcal{L}_X$ with free variable y.

Definition 5.3. Given a model M and a type $\Sigma(y)$, we say that M realizes Σ if there is some $m \in M$ such that $M \vDash \sigma[m]$ for each $\sigma \in \Sigma$. This may also be written $M \vDash \Sigma[m]$.

Definition 5.4. Given a model M, a subset X, and an element m in M, the complete type of m over X, denoted $tp_X(m)$ is the set of all $\mathcal{L}_X$ sentences in one free variable which m satisfies in $(M, x)_{x \in X}$

Types will be the way in which we describe elements we would like to have in our models. Expansions of models will be useful for ensuring that these elements "behave well" with respect to given subsets of the model. Saturated models will realize as many types as possible so that any element we might (consistently) describe actually exists in the model.

Definition 5.5. A model M is α-saturated if or every subset X of M with $|X| < \alpha$, $(M, x)_{x \in X}$ realizes every type $\Sigma(y)$ which is consistent with the set of $\mathcal{L}_X$ sentences true in the expansion of M.

Definition 5.6. A model M is saturated if it is $|M|$-saturated.

Saturated models are one of the friendliest settings in which to find isomorphisms. We give a simple application which also serves to demonstrate the back-and-forth method. The general technique is similar to the proof of the Ax-Kochen theorem, though considerably less difficult.

Theorem 5.1. *Let $M \equiv N$ be saturated models of the same cardinality: then $M \cong N$.*

Proof. By transfinite induction on ordinals $\alpha < |M|$, we will iteratively construct a sentence-preserving map f from M to N.

At each successor ordinal, we will start with $f_\zeta : A_\zeta \to B_\zeta$ which preserves sentences in L_{A_ζ}. Then, we can extend the domain to include some missing member of M. Interchanging the roles of M and N, we expand the range to obtain $f_{\zeta+1} : A_{\zeta+1} \to B_{\zeta+1}$. Taking the union will provide the desired isomorphism.

Base case, $\zeta = 0$: f_0 on $A_i = B_i = \emptyset$.

Limit ordinals: set $A_\zeta = \cup_{\alpha<\zeta} A_\alpha$, $B_\zeta = \cup_{\alpha<\zeta} B_\alpha$, and $f_\zeta = \cup_{\alpha<\zeta} f_\alpha$. If ϕ is a sentence in $\mathcal{L}_{A_\zeta}$, it can only use finitely many of the new constants adjoint to $\mathcal{L}$, so there is some $\alpha < \zeta$ such that A_α contains all of the corresponding elements. By induction, f_ζ extends f_α, and f_α preserves ϕ.

Successor ordinals, $\zeta = \alpha + 1$: pick two elements, $m_\zeta \in M - A_\alpha$ and $n_\zeta \in N - B_\alpha$ - we wish to extend f_α to f_ζ which contains each of these in its domain and range, respectively.

Start with $tp_{A_\alpha}(m_\zeta)$. This is a consistent type in the expansion of M by A_α. Take the corresponding type Σ in N, applying f to each constant in the sentences of $tp_{A_\alpha}(m_\zeta)$. We now have a type in $(N, b)_{b \in B_\alpha}$ - because f preserves sentences, it remains maximal and consistent. As N is saturated, let $y \in N$ realize Σ and define $f'_\zeta(m_\zeta) = y$. Clearly, f'_ζ preserves sentences in $A_\alpha \cup \{m_\zeta\}$ since the types of m_ζ and y correspond. The same argument with the roles of N and M switched, using the function f'^{-1}_ζ allows the range to be expanded to include n_ζ. As such, we obtain a new function f_ζ whose domain A_ζ and range B_ζ include the two stipulated elements. This function remains a bijection: the fact that y satisfies the type corresponding to $tp_{A_\alpha}(m_\zeta)$ ensures that $y \notin B_\alpha$, and likewise for n_ζ with respect to $A_\beta \cup \{m_\zeta\}$ (except when $y = n_\zeta$, but then there is nothing to prove).

By transfinite induction, we obtain a function $f : M \to N$ which is a sentence-preserving bijection - an isomorphism - as desired. □

The general idea of this method is to find an isomorphism $f : M \to N$ by expanding the domain, then expanding the range, and so on (going "back-and-forth") in a sentence-preserving fashion until all of M and N are covered.

The main content of the Ax-Kochen theorem is in proving that two (saturated) elementarily equivalent structures are isomorphic. Unfortunately, the result above cannot be used to do this - our goal is elementary equivalence, which is the hypothesis of this theorem. The last model theoretic result we need is the existence of saturated models.

Theorem 5.2. *Let T be a theory in $\mathcal{L}$, with an infinite model, and let $\alpha = ||\mathcal{L}||$ (ω if $\mathcal{L}$ is finite, and $|\mathcal{L}|$ otherwise). Then T has saturated models of cardinality alpha$^+$, where α^+ is the successor cardinal of α.*

An unfortunate feature of this theorem is that he saturated model's cardinality is α^+, which cannot be more closely pinned down without stronger set theories than ZFC. In the proof of the Ax-Kochen theorem, we will assume the GCH - the proof can be done without the GCH, avoiding saturated models, at the cost of substantial complexity.

6 Ax-Kochen

Our final goal is to prove the Ax-Kochen theorem, closely following the presentation in [1].

Definition 6.1. A valued field F, with cross-section, is a model of the two-sorted language

$$\mathcal{L} = \{F, +, *, 0, 1, V, val, \leq\}$$

satisfying the following axioms:

1. F is a field (using $+, *, 0, 1$),
2. V is a multiplicative group, with respect $*, 1$,
3. $\leq$ is a total ordering of $V \cup 0$ where $0 \geq v$ for all $v \in V$,
4. the group's operation respects the ordering: $(\forall x, y, z \in V)(x \leq y \to x * z \leq y * z)$,
5. val is a multiplicative function from F onto $V \cup \{0\}$ such that $val(x) = 0$ iff $x = 0$, and for any $x, y \in F$:
$$val(x + y) \geq \min(val(x), val(y)),$$
6. the valuation satisfies the cross-section axiom:
$$v \in V \to val(v) = v$$

It is not immediately clear why the cross-section is important, but it will be used at several locations in the proof. The Ax-Kochen theorem can be generalized to valued fields without cross-section (just as it can be proven without the continuum hypothesis) at the cost of greater complexity.

As the reader may recall from algebra, there are two important structures associated with any valued field beyond its value group: the valuation ring and residue field.

Definition 6.2. Given a valued field F, its valuation ring, $R(F)$ is the set of all element in F with valuation at least 1. The valuation ring has a unique maximal ideal, $M(F)$ consisting of those elements with valuation strictly greater than 1. The residue field, F^* is defined as $R(F)/M(F)$.

In particular, the Ax-Kochen theorem applies to valued fields satisfying Hensel's lemma ("Hensel's lemma"). This constraint is of less obvious model-theoretic importance, but it is important to the algebraic parts of the argument. It is a classical result that the fields we are interested in, $\mathbb{Q}_p$ and $\mathbb{F}_p((t))$, satisfy Hensel's lemma.

Definition 6.3 ((Hensel's lemma))**.** Let F be a valued field and take a monic polynomial $f(x) \in R(F)[x]$. Let f^* be the polynomial in $F^*[x]$ obtained by taking f's coefficients mod $M(F)$. If $f^*(x) = g'(x)h'(x)$ in $F^*[x]$, then there exist polynomials $g, h \in R(F)[x]$ such that $g^* = g'$ and $h^* = h'$ and $f(x) = g(x)h(x)$

That Hensel's lemma is a first-order statement is straightforward. Fix the degree of f. Polynomials may be identified with their coefficients, and it is clear that we can pick out $R(F)$ and $M(F)$ with first-order statements about valuation. To say that $f^*(x) = g'(x)h'(x)$ in $F^*[x]$ is the same as saying that the corresponding coefficients are congruent mod $M(F)$, which is to say that their difference is in $M(F)$.

With all of this in place, the Ax-Kochen theorem is the following:

Theorem 6.1 (Ax-Kochen)**.** *Suppose F and G are valued fields with cross-section satisfying Hensel's lemma, $F^* \equiv G^*$, both residue fields are characteristic zero, and $val(F) \equiv val(G)$. Then $F \equiv G$.*

It is not entirely surprising that information about the residue field and value group determine the properties of a valued field. In a sense, we can view a valued field as an extension of the residue field by the value group. This connection is tighter when the residue field has characteristic zero, as then the residue field can be naturally identified with a valued subfield of of $R(F)$. This also suggests why the cross-section simplifies the proof somewhat: for a valued field with cross-section, the value group is truly contained in the field, more tightly constraining these fields to match our intuition about such fields.

Following Chang and Keisler's approach, we isolate the algebraic parts into the following lemma (their lemma 5.4.13). The most important components of the lemma provide conditions under it is possible to extend an isomorphism of valued fields.

A last definition we need for this lemma:

Definition 6.4. If F is a valued field, its henselization is a valued field G extending F such that G is ahensel field and for any other Hensel algebraic extension H of F, there is an embedding of G into H over F.

Lemma 6.2. *1. Let F be a valued field with $x, y \in F$ and $val(x) < val(y)$, then*

$$val(-x) = val(x)$$

$$val(x + y) = val(x)$$

Moreover, for any $v \in val(F)$ and natural number n, there is at most one $w \in val(F)$ such that $v = w^n$.

2. *Let F be a Hensel field whose residue class field has characteristic 0. Then there is a valued subfield F_0 of F contained in $R(F)$ such that the natural homomorphism from $R(F)$ to F^* maps F_0 isomorphically onto F^*.*

3. *Let F_1 and G_1 be valued fields algebraic over Hensel subfields F_0 and G_0. Suppose $f : F_0 \to G_0$ is an isomorphism (of valued fields) which can be extended to an isomorphism (of fields) $g : F_1 \to G_1$. Then g is an isomorphism of valued fields, and if σ is an automorphism of F_1 over F_0 and $x \in F_1$, $val(x) = val(\sigma x)$.*

4. *Every valued field has a henselization. If F and G are valued fields and f is an isomorphism between them, then f extends to an isomorphism of their henselizations.*

5. *If F_0 is the henselization of F then $F_0^* = F*$ and $val(F_0) = val(F)$.*

6. *If F is a valued field with valued subfield F_0, then $val(\bar{F}_0)$ (the value group of the algebraic closure of F_0 in F) is the closure of $val(F_0)$ under roots as a subgroup of $val(F)$. If F is a Hensel field and $F^* = F_0^*$, F^* has charcteristic zero, and $val(F_0)$ is closed under roots in $val(F)$, then $\bar{F}_0$ is the henselization of F_0.*

7. *Suppose F_1 and G_1 are Hensel fields with Hensel subfields F and G, $x \in F_1, y \in G_1$ transcendental over F and G, $f : F \to G$ an isomorphism,*

$$val(F(x)) = val(F), F(x)^* = F^*$$

$$\text{and for all } a \in F, f(val(x - a)) = val(y - f(a)).$$

Then $val(G(y)) = val(G)$, $G(y)^* = G^*$, and f can be extended to an isomorphism between $F(x)$ and $G(y)$.

8. *Suppose F_0 is a Hensel field with a hensel subfield F, $x \in F_1$ is transcendental over F, and $F(x)^* = F^*$. If $val(F)$ is nontrivial, then $|val(F(x))| = |val(F)|$*

The lemma largely consists of conditions under which an isomorphism of valued fields can be extended, as well as facts concerning the behavior of subfields of Hensel fields when extended in various ways. Part 2 provides the base case for inductively constructing our desired isomorphism. Part 7 of the lemma will be most crucial to building our isomorphism. When we come to deal with saturation, Part 8 will provide a valuable cardinality constraint.

We are now ready to prove the main result:

Theorem 6.3 (Ax-Kochen)**.** *Suppose F and G are valued fields with cross-section satisfying Hensel's lemma, $F^* \equiv G^*$, both residue fields are characteristic zero, and $val(F) \equiv val(G)$. Then $F \equiv G$.*

Proof. First, we may assume that F and G are saturated fields of cardinality ω_1 - the set of all first-order sentences true of F is complete and consistent, so it has a saturated model of cardinality $||\mathcal{L}||^+ = |2^\omega| = \omega_1$ (contingent on the GCH), and likewise for G. Clearly, if F and G are saturated then their value groups and residue-fields are saturated. Moreover, outside the case where $val(F) = val(G) = \{1\}$, the value groups and residue fields all have cardinality ω_1 (in that trivial case, $F \cong F^* \equiv G^* \cong G$ and there is nothing to prove).

We will write $f_1 : F_1 \leftrightarrow G_1$ iff f_1 is an isomorphism and

$$(val(F), x)_{x \in val(F_1)} \equiv (val(G), x)_{x \in val(G_1)}$$

Our goal is to show that F and G are isomorphic using a back-and-forth argument like that used to show that any two elementarily equivalent models of the same cardinality are isomorphic. An outline of the induction is as follows:

1. Since $F^* \equiv G^*$ and $|F^*| = |G^*| = \omega_1$, they are isomorphic. It follows from Hensel's lemma that they are algebraically closed in F and G, respectively. This is our base case.

2. Suppose F_1 and G_1 are algebraically closed valued subfields of F and G, containing their respective residue fields, and where $val(F_1) = val(G_1)$ is countable. Let $f_1 : F_1 \leftrightarrow G_1$ extending $f_0 : F^* \leftrightarrow G^*$. For every $x \in F - F_1$ there exist algebraically closed valued subfields F_2 and G_2 with $x \in F$ containing F_1 and G_1, a function $f_2 : F_2 \leftrightarrow G_2$ extending f_1, and such that $val(F_2)$ is countable. There are three subcases:

   (a) When $x \in val(F_1)$,

   (b) When $val(F_1(x)) = val(F_1)$,

   (c) All other cases - but by countably many applications of the second subcase, this is reduced to the first subcase.

3. The second step holds with F_1 andG_1 exchanged. By the same back-and-forth argument, we obtain an isomorphism between F and G.

So it just remains to prove 2a and 2b. The countability hypothesis will be important when applying saturation to find elements.

Case 2a: While apparently longer, this section is conceptually simple. First, we extend the isomorphism to $F_1(x)$, then work carefully to find a Hensel field containing $F_1(x)$ to which the isomorphism extends. This largely comes down to taking the closure of the value group under roots. The cross-section is used heavily in this section.

Consider $tp_{val(F_1)}(x)$. Since we already have $f_1 : F_1 \leftrightarrow G_1$, we can transport this to a type over $val(G_1)$ - $val(G_1)$ is countable and saturated, so there is a y in $val(G)$ satisfying that type, which is to say that:

$$(val(F), x, a)_{a \in val(F_1)} \equiv (val(G, y, b)_{b \in val(G_1)}.$$

Let V be the subgroup of $val(F)$ generated by $val(F_1)$ and x, W the subgroup of $val(G)$ generated by $val(G_1)$ and y. Clearly both V and W are countable. We would like to show that they are the value groups of $F_1(x)$ and $G_1(y)$, respectively. It suffices to verify this that they contain the value groups of $F_1[x]$ resp. $G_1[y]$. Take some polynomial

$$p(t) = e_0 + e_1 t + ... + e_n t^n \in F_1[t].$$

If $r < s \leq n$ are such that $e_r, e_s \neq 0$, then $val(e_r x^r) \neq val(e_s x^s)$. Otherwise, we would have

$$val(e_r)x^r = val(e_r x^r) = val(e_s x^s) = val(e_s)x^s,$$

hence

$$x^{s-r} = val(e_r)/val(e_s) \in val(F_1),$$

But we assumed that F_1 is algebraically closed in F, so its value group is closed under roots in F. Since $val(e_r), val(e_s)$ are in $val(F_1)$, the above would imply that $x \in val(F_1) \subseteq F$ unless $s = r$. As a result, there is a unique $e_r x^r$ of smallest value. By Part 1 of the lemma, $val(p(x)) = val(e_r)x^r$. If $q(t)$ is the image

of $p(t)$ under f_1, then we know that $val(q(y)) = val(f(e_r))y^r$ because f_1 is an isomorphism. This proves $val(F_1(x)) = V$ and $val(G_1(y)) = W$.

Now define a (field) isomorphism $g_1 : F_1(x) \to G_1(y)$ extending f_1 by sending x to y. It is also an isomorphism of the value groups V and W since it preserves products. By Part 4 of the lemma, we can extend g_1 to an isomorphism g_2 of the henselizations F^2 and G^2 of $F_1(x)$ and $G_1(y)$, respectively. Since F and G are hensel fields, we may take F^2 and G^2 to be contained in F and G, respectively. By Part 5 of the lemma, $val(F^2) = val(F_1(x)) = V$ and $val(G^2) = W$.

Let $\bar{V}, \bar{W}$ be the closures of V and W under roots in $val(F), val(G)$, respectively. By Part 1 of the lemma, these can be expressed as countable unions of countable sets, and hence remain countable.

We know that

$(val(F), c)_{c\in V} \equiv (val(G), g_1 c)_{c\in V}$ (there appears to be a typo in the corresponding line in Chang and Keisler)

and so $g_2|_V$ extends to a unique isomorphism $h : \bar{V} \to \bar{W}$, again using saturation. As such, we will have that

$$(val(F), c)_{c\in\bar{V}} \equiv (val(G), g_1 c)_{c\in\bar{V}}$$

Every element in $\bar{V}$ (resp. $\bar{W}$) is algebraic over F^2 (resp. G^2), so consider the algebraic extensions $F^3 = F^2(\bar{V})$ and $G^3 = G^2(\bar{W})$ of F^2 and G^2, respectively. By Part 6 of the lemma, $val(F^3) = \bar{V}$ and $val(G^3) = \bar{W}$, as those value groups are already closed under roots. Then g_2 may be extended to a field isomorphism of F^3 and G^3, which by Part 3 of the lemma will in fact be an isomorphism of valued fields.

All together, we have

$$F^* \subseteq F_1 \subseteq F^3, G^* \subseteq G_1 \subseteq G^3$$

and so $F^* = F^{3*}$ and $G^* = F^{3*}$. From this and Part 6 of the lemma, the algebraic closures of F^3 and G^3 in F and G are their henselizations. So we obtain a final isomorphism $g_4 : \bar{F}^3 \to \bar{G}^3$, where we also know that

$$val(\bar{F}^3) = val(F^3) = \bar{V}$$
$$val(\bar{G}^3) = \bar{W}$$

This, and the statement above about the elementary equivalence of $val(F)$ and $val(G)$ over $\bar{V}$ and $\bar{W}$ means that g_4 satisfies the conditions necessary to conclude that $g_4 : \bar{F}^3 \leftrightarrow \bar{G}^3$. Note that $\bar{V}$ and $\bar{W}$ are countable. This concludes case 2a.

Case 2b: In this case, $val(F_1(x)) = val(F_1)$, setting the stage to apply Part 7 of the lemma. As the reader might guess, saturation will be employed to find an element y of G satisfying Part 7's hypotheses.

It is immediate that $F_1(x)^* = F^* = F_1^*$.

Since F_1 is algebraically closed in F, $val(F_1)$ is closed under roots in $val(F)$ (roots are algebraic, of course) it follows from Part 6 of the lemma that $\bar{F_1} = F_1$ is hensel. The same is true of G_1.

The only unsatisfied hypothesis of Part 7 is the existence of a a $y \in G$ such that for all $a \in F_1$,

$$(*) f_1(val(x - a)) = val(y - f_1(a))$$

Let A be an arbitrary finite subset of F_1. We will find a y in G such that there is a y in G satisfying $(*)$ for all $a \in A$ (recall that the consistency of a type reduces to finite satisfiability).

Let $b \in A$ be such that $c = val(x - b)$ is maximized. Since in this case $val(F_1(x)) = val(F_1)$, $c \in val(F_1)$. Note that for any positive integer n, $val(n) = 1$ - an element of $R(F)$ is a unit iff it has valuation 1, and we are in the case where the characteristic is zero. As such, $val(nc) = val(c)$ for all integers n. Then for all $a \in A$:

$$val(b - nc - a) \geq \min(val(b - x), val(nc), val(x - a)) = val(x - a)$$

since $val(b - x) = val(x - b) = val(nc) = val(c)$ is the largest $val(x - a)$ for each $a \in A$. Further, for a fixed $a \in A$, there is at most one n such that the equality above is strict. If not, let $m < n$ be two such values for some $a \in A$ such that

$$val(b - mc - a), val(b - nc - a) > val(x - a)$$

Of course, this means that their minimum exceeds $x-a$, so:

$$c = val(c) = val((n-m)c) = val((b-mc-a)-(b-nc-a)) \geq \min(val(b-mc-a), val(b-nc-a)) > val(x-a)$$

Then by Part 1 of the lemma,

$$val(b-nc-a) = val((b-x)-nc+(x-a)) = val(x-a),$$

since $val(b-x) = val(-nc) > val(x-a)$. This is a contradiction, so there cannot be such m and n.

As a result, there are only finitely many "exceptional" n, and so there exists some $n > 0$ such that $val(b-nc-a)) = val(x-a)$ for all $a\ inA$. Let $y = f(b-nc)$ for this n. Then for all $a \in A$,

$$f(val(x-a)) = f(val(b-nc-a)) = val(y-f(a))$$

because $val(x-a) = val(b-nc-a)$, on which the isomorphism is already defined.

Since $val(F_1(x)) = val(F_1)$ is countable, there is a countable set $A_1 \subseteq F_1$ (namely, the value group) such that for all $b \in F_1$ there exists an $a \in A_1$ with $val(x-a) = val(x-b)$.

Since we proved the analogue of $(*)$ for all finite subsets of F_1 we at least know that the set of sentences expressing $(*)$ for $a \in A_1$ is finitely satisfiable, hence consistent. We may translate this into a type in G over a countable set. Since G is ω_1-saturated, there exists a $y \in G$ such that $(*)$ holds for all $a \in A_1$.

Now let $b \in F_1$. Since $F_1(x)^* = F_1^*$, there exists a $b' \in F_1$ such that for $c = val(x-b)$:

$$val((x-b)c^{-1} - b') > 1$$

multiply through by $c = val(x-b)$ to obtain

$$val(x-(b+cb')) > val(x-b)$$

so there is an $a \in A_1$ such that $val(x-a) = val(x-(b-cb'))$, and hence $val(x-b) < val(x-a)$.

Apply Part 1 of the lemma to see that

$$val(a-b) = val(x-b-(x-a)) = val(x-b) < val(x-a)$$

and so

$$val(fa-fb) = f(val(a-b)) < f(val(x-a)) = val(y-fa)$$

from which it follows that

$$val(y-fb) = val(y-fa+fa-fb) = val(fa-fb) = f(val(a-b)) = f(val(x-b)).$$

This means we have found a y satisfying the conditions for Part 7 of the lemma, and so $val(G_1(y)) = val(G_1)$ and f extends to an isomorphism $g_1 : F_1(x) \to G_1(y)$.

As F_1 and G_1 are algebraically closed in F, G the value groups are closed under roots in $val(F), val(G)$. As such, Part 6 of the lemma implies that their algebraic closures in F and G, respectively, are their henselizations, and that there is an isomorphism between them extending g_1. These fields satisfy the criteria of the inductive hypothesis, and so this case is complete.

This also concludes the proof. As mentioned previously, this closely follows the proof in section 5.4 of [1], with a few small changes. □

Even with many details abstracted away, this is still a substantial theorem. Still, the general ideas are simple enough. In the easy (but long) case, 2a, we exactly follow our intutition that valued fields look like adjoining the value group to the residue field - most of the complicated work is done to wrangle the field into a form demanded by the inductive step (Hensel and algebraically closed in F) but that largely follows from repeated applications of the isomorphism extension theorems in the lemma. Case 2b uses saturation to specify a desirable element in G to extend an isomorphism. Then, as before, work is done to put the resulting fields in the appropriate form - here, this can be done more easily because the value group is changing, even though it is much harder to find the initial extension of the isomorphism.

Now we can state a powerful transfer principle between $\mathbb{Q}_p$ and $\mathbb{F}_p((t))$ as a corollary:

Corollary 6.3.1. *Let ϕ be a statement in the language of valued fields. Then for all but finitely many primes p, $\mathbb{Q}_p \vDash \phi$ iff $F_p((t)) \vDash \phi$.*

Proof. It is well-known that both $\mathbb{Q}_p$ and $\mathbb{F}_p((t))$ are valued fields satisfying Hensel's lemma. They each have residue fields isomorphic to $\mathbb{F}_p$ and value groups isomorphic to $\mathbb{Z}$ under addition.

Let D be a non-principal ultrafilter on the set of primes. Let $F = \prod_D \mathbb{F}_p((t))$ and $G = \prod_D \mathbb{Q}_p$. Hensel's lemma is first-order, and hence preserved by the ultraproduct. Moreover,

$$F^* \cong \prod_D \mathbb{F}_p \cong G^*$$

As remarked in the example above, when D is non-principal, this will be af ield of characteristic 0. Furthermore, the value groups are

$$val(F) \cong \prod_D \mathbb{Z} \cong val(G)$$

Isomorphic models are certainly elementarily equivalent, so it follows by the previous theorem that for any sentence ϕ, the set $S_\phi = \{p \mid \mathbb{Q}_p \vDash \phi iff \mathbb{F}_p((t))\}$ is in D. This holds for all non-principal ultrafilters, and so all but finitely many primes are contained in that set (if S_ϕ were infinite, there would exist a nonprincipal ultrafilter excluding it, a contradiction). □

This transfer principle allows (first-order) theorems to be shared, for all but finitely many primes, between these two kinds of fields. An important corollary is Artin's conjecture:

Corollary 6.3.2. *For every positive integer d, there is a set Y_d of at most finitely many primes such for all $p \notin Y_d$, every homogeneous form $f(t_1, ..., f_n)$ of degree d in at least $n > d^2$ variables has a nontrivial solution.*

Proof. First, the stated property is first-order. As in earlier examples, quantification over polynomials is replaced by quantification over their coefficients, and nontriviality is asserted by a disjunction. One of the results in Serge Lang's thesis was a proof the analogous statement for all of the fields $\mathbb{F}_p((t))$. The transfer principle above shows that it fails to hold in $\mathbb{Q}_p$ for only finitely many primes. □

This conjecture has an interesting history. In fact, Artin's original statement was false: he had hoped it would be true for all the p-adic fields. In 1965, Ax and Kochen proved the transfer principle above, but it was only strong enough to show that Artin's original conjecture would fail for at most finitely many primes depending on the degree. Then, in 1996, Terjanian came upon a counterexample to the original conjecture in $\mathbb{Q}_2$ of degree 4 several weeks after presenting on the Ax-Kochen theory [3]. Examples are now also known for other p-adic fields (since $\mathbb{Q}_2$ sometimes exhibits unusual behavior, these other examples make the sharpness of the result clearer). As mentioned, there are now effective proofs of this result which can even determine the exceptional set of primes [2] though the computational complexity is significant.

References

[1] C. C. Chang and H. J. Keisler. *Model Theory*, volume 73 of *Studies in Logic*. Elsevier, 2nd edition edition, 2002.

[2] P. Cohen. Decision procedures for real and p-adic fields. 1969.

[3] J. Jorgenson and S. G. Krantz. The mathematical contributions of serge lang, part ii. 2006.

[4] S. Kochen. The model theory of local fields. 1975.

[5] D. Mark. *Model Theory: An Introduction*. Graduate Texts in Mathematics. Springer, 1st edition, 2002.

[6] A. Robinson. Problems and methods of model theory. 1968.