THE AX-KOCHEN THEOREM: AN APPLICATION OF MODEL THEORY TO ALGEBRA

ALEX KRUCKMAN

Abstract. The Ax-Kochen Theorem is a purely algebraic statement about the zeros of homogeneous polynomials over the p-adic numbers, but its proof uses techniques from mathematical logic. This thesis provides an exposition of the algebra and model theory necessary to understand the theorem and its proof.

Contents

1. Introduction

Certain fields have the property, called C_i, that every homogeneous polynomial with enough variables relative to its degree (specifically, $n > d^i$, where n is the number of variables and d is the degree) has a nontrivial zero.

Emil Artin conjectured that for all primes p, the p-adic field $\mathbb{Q}_p$ is C_2. This conjecture turned out to be false; in fact, $\mathbb{Q}_p$ is not C_2 for any p. However, in their paper *Diophantine problems over local fields* [AK65], Ax and Kochen provided a partially positive result.

Theorem 4.3.1 (Ax-Kochen Theorem)**.** *For each degree $d \geq 1$, there exists a finite set of primes $P(d)$ such that for all $p \notin P(d)$, if f is a homogeneous polynomial over $\mathbb{Q}_p$ of degree d in n variables such that $n > d^2$, then f has a nontrivial zero in $\mathbb{Q}_p^n$.*

The methods used by Ax and Kochen come from model theory, a branch of mathematical logic. They were able to prove a much more general result, known as the Ax-Kochen Principle, which allows theorems about the fields $\mathbb{F}_p((t))$ of formal Laurent series over the finite fields $\mathbb{F}_p$ to be transferred to theorems about the fields $\mathbb{Q}_p$.

Theorem 4.2.3 (Ax-Kochen Principle)**.** *Any first-order logical statement about valued fields which is true of all but finitely many of the fields $\mathbb{F}_p((t))$ is true of all but finitely many of the fields $\mathbb{Q}_p$.*

This thesis provides an exposition of the algebra and model theory necessary to understand the Ax-Kochen Theorem and its proof. It should be accessible to any reader with a firm grasp of abstract algebra.

We begin in Section 2.1 by introducing homogeneous polynomials and the C_i properties. In Sections 2.2 and 2.3, we prove C_i properties for finite fields and algebraic and transcendental extension fields. In Section 2.4, we introduce valued fields and the completion of a discrete valued field, constructing the p-adic fields along the way. Finally, we prove that $\mathbb{F}_p((t))$ is C_2 for all p, the result that will be transferred to the p-adics to complete the proof of the Ax-Kochen Theorem. For the material in Chapter 2, I have followed Greenberg [Gre69] closely.

In Chapter 3, we introduce the reader to model theory, with a focus on those techniques and examples relevant to the Ax-Kochen Principle. I have modeled my notation and exposition after that in Marker [Mar02], but some of the details (for example, the material on ultraproducts and the model theory of valued fields) come from Chang and Keisler [CK73].

Chapter 4 is devoted to the proof of of the Ax-Kochen Principle. The proof relies on the result that the $\mathbb{F}_p((t))$ and $\mathbb{Q}_p$ are Henselian valued fields, and we introduce Hensel's Lemma and some of its consequences in Section 4.1. We give the proof of the Ax-Kochen Principle in Section 4.2, the cornerstone of which is Theorem 4.2.2, which implies that the ultraproducts of the fields $\mathbb{F}_p((t))$ and $\mathbb{Q}_p$ are elementarily equivalent. Finally, we derive the Ax-Kochen Theorem as a corollary in Section 4.3. Again, the main reference for the proof is Chang and Keisler [CK73].

Some sections require a familiarity with the transfinite numbers. Their properties are covered in Appendix A. We will also use the resultant, an algebraic tool for comparing the roots of two polynomials. It is introduced in Appendix C. For simplicity, the proof of the Ax-Kochen Principle as given relies on the Continuum Hypothesis. Appendix B describes a method for eliminating the Continuum Hypothesis from the argument.

This thesis was written in partial fulfillment of the requirements for the degree of Bachelor of Science with Honors in Mathematics at Brown University. I would to express my eternal gratitude to my advisors Dan Abramovich and Michael Rosen, who have been extremely generous with their time, suggestions, and support, and to my parents, for their devotion to my education.

2. On Quasi-Algebraic Closure

2.1. The C_i Properties and Algebraically Closed Fields.

Definition 2.1.1. A polynomial f over a field F is *homogeneous* of degree $d \geq 1$ in n variables, $x_1, \ldots, x_n$, if all monomials of f have degree d, that is, if it can be written in the form $f(x_1, \ldots, x_n) = \sum_i a_i x_1^{b_{i,1}} \ldots x_n^{b_{i,n}}$ such that for all i, $\sum_{j=1}^{n} b_{i,j} = d$.

Remark 2.1.2. If f is a homogeneous polynomial of degree d in n variables over F, then for all $c \in F$, $f(cx_1, \ldots, cx_n) = c^d f(x_1, \ldots, x_n)$.

Example 2.1.3. The function which computes the determinant of an $n \times n$ matrix is a homogeneous polynomial of degree n in n^2 variables, the matrix entries.

Since a homogeneous polynomial cannot have a constant term, all homogeneous polynomials have the trivial zero $(0, \ldots, 0)$. It is of interest to explore when homogeneous polynomials have nontrivial zeros.

Example 2.1.4. Let f_n be the polynomial $x_1^2 + x_2^2 + \ldots + x_n^2$. For all $n > 0$, f_n is a homogeneous polynomial of degree 2 in n variables. Over $\mathbb{R}$, f_n has only the trivial zero for all n. But over $\mathbb{C}$, f_n has nontrivial zeros (for example, $(1, i, 0, \ldots, 0)$) for all $n > 1$. It is easy to check that over $\mathbb{F}_7$, f_n has nontrivial zeros for all $n > 2$, and in Section 2.2 we will show that this is the case for all finite fields. The 2 comes from the degree of f_n.

Definition 2.1.5. A field F is called C_i for $i \in \mathbb{N}$ if every homogeneous polynomial over F of degree d in n variables such that $n > d^i$ has a nontrivial zero in F^n.

We can easily characterize the C_0 fields.

Theorem 2.1.6. *A field is C_0 if and only if it is algebraically closed.*

Proof. Suppose F is an algebraically closed field. Let $f(x_1, \ldots, x_n)$ be a homogeneous polynomial over F of degree d in $n > d^0 = 1$ variables. Write f as a polynomial in one variable, x_1, with coefficients in $F[x_2, \ldots, x_n]$, $f = \sum_{i=1}^{d'} f_i(x_2, \ldots, x_n) x_1^i$. The degree of this polynomial, d', is the highest power of x_1 appearing in any term of f.

If $d' = 0$, then no nonzero power of x_1 appears in f, so $f(1, 0, \ldots, 0) = f(0, 0, \ldots, 0) = 0$, and $(1, 0, \ldots, 0)$ is a nontrivial zero of f. Otherwise, if $d' > 0$, consider the leading coefficient, $f_{d'}(x_2, \ldots, x_n)$. We would like to find nontrivial $(\alpha_2, \ldots, \alpha_n) \in F^{n-1}$ which is *not* a zero of $f_{d'}$. All algebraically closed fields are infinite, and a nonzero polynomial cannot have infinitely many zeros, so there exists $(\alpha_2, \ldots, \alpha_n) \in F^{n-1}$ such that $\alpha_j \neq 0$ for some j and $f_{d'}(\alpha_2, \ldots, \alpha_n) \neq 0$.

Let $\overline{f} = f(x_1, \alpha_2, \ldots, \alpha_n)$. We have simply evaluated the coefficients f_i at $(\alpha_2, \ldots, \alpha_n)$, and the leading coefficient is nonzero, so $\overline{f}$ is a polynomial of degree $d' > 0$ in one variable, x_1. Since F is algebraically closed, $\overline{f}$ has a zero, α_1. Then $(\alpha_1, \alpha_2, \ldots, \alpha_n)$ is a zero of f, and this zero is nontrivial, since $\alpha_j \neq 0$.

Conversely, suppose F is a C_0 field. Let $f(x) = a_d x^d + a_{d-1} x^{d-1} + \ldots + a_0$ be a polynomial of degree $d \geq 1$ over F. We would like to show that f has a root in F. Let $\widehat{f}(x_1, x_2) = a_d x_1^d + a_{d-1} x_1^{d-1} x_2 + \ldots + a_1 x_1 x_2^{d-1} + a_0 x_2^d$. Now $\widehat{f}$ is a homogeneous polynomial over F of

degree d in 2 variables, and $2 > d^0 = 1$, so $\widehat{f}$ has a nontrivial zero $(\alpha_1, \alpha_2) \in F^2$. Note that $\alpha_2 \neq 0$, since otherwise $\widehat{f}(\alpha_1, \alpha_2) = \widehat{f}(\alpha_1, 0) = a_d \alpha_1^d = 0$, so $\alpha_1 = 0$, and (α_1, α_2) is trivial.

We have $\widehat{f}(\alpha_1, \alpha_2) = 0$, so by Remark 2.1.2, $\widehat{f}(\alpha_1 \alpha_2^{-1}, 1) = (\alpha_2^{-1})^d \widehat{f}(\alpha_1, \alpha_2) = 0$. But substituting 1 for x_2 in $\widehat{f}$, $\widehat{f}(x_1, 1) = f(x_1)$, so $f(\alpha_1 \alpha_2^{-1}) = \widehat{f}(\alpha_1 \alpha_2^{-1}, 1) = 0$, and $\alpha_1 \alpha_2^{-1}$ is a root of f in F. Thus every polynomial over F of nonzero degree has a root in F, and hence F is algebraically closed. □

This theorem suggests that the C_i properties can be seen as generalizations of the property of algebraic closure. For this reason, C_1 fields are called quasi-algebraically closed. In the following sections, we will show that many frequently encountered fields are C_i for some i. As a special case, we will obtain our first main result: for all primes p, $\mathbb{F}_p((t))$, the field of formal Laurent series in one variable over the finite field with p elements, is C_2.

2.2. **Finite Fields.** Throughout this section, let K be a finite field of characteristic p with $|K| = q$. Recall that

- p is prime,
- $q = p^v$ for some $v > 0$, and
- the multiplicative group $K^* = K \backslash \{0\}$ is cyclic of order $q - 1$.

We will show that all finite fields are C_1. We begin with a simple but useful lemma.

Lemma 2.2.1. *For $m > 0$,*

$$\sum_{a \in K} a^m = \begin{cases} -1 & \text{if } (q-1) \mid m \\ 0 & \text{otherwise} \end{cases}.$$

Proof. Suppose $q - 1 \mid m$. Then for all $a \in K^*$, $a^m = 1$, so

$$\begin{aligned} \sum_{a \in K} a^m &= 0^m + \sum_{a \in K^*} a^m \\ &= \sum_{a \in K^*} 1 \\ &= -1 \end{aligned}$$

since $q - 1 \equiv -1 \pmod{p}$.

Otherwise, if $(q-1) \nmid m$, let b be a generator of the cyclic group K^*. Then $b^m \neq 1$, since $|K^*| = q - 1$. Let $S = \sum_{a \in K} a^m = \sum_{a \in K^*} a^m$. Multiplication by b permutes the elements of K^*, so

$$\begin{aligned} S &= \sum_{a \in K^*} (ba)^m \\ &= b^m \sum_{a \in K^*} a^m \\ &= b^m S, \end{aligned}$$

and thus $(b^m - 1)S = 0$. But $b^m - 1 \neq 0$, so $S = 0$, as was to be shown. □

The next theorem implies that finite fields are C_1, but it is actually a stronger result about the number of zeros of any polynomial (not necessarily homogeneous) with more variables than its degree.

Theorem 2.2.2 (Chevalley-Warning, [Gre69, Theorem 2.3])**.** *Let f be a polynomial over K of degree d in n variables, $x_1, \ldots, x_n$. If $n > d$, then the number of zeros of f in K^n is divisible by p.*

Proof. For $(\alpha_1, \ldots, \alpha_n) \in K^n$,

$$1 - f(\alpha_1, \ldots, \alpha_n)^{q-1} = \begin{cases} 1 & \text{if } f(\alpha_1, \ldots, \alpha_n) = 0 \\ 0 & \text{otherwise} \end{cases}.$$

We will count the number of zeros (mod p) of f by summing the values of this expression over all $(\alpha_1, \ldots, \alpha_n) \in K^n$. There are q^n such n-tuples.

$$\begin{aligned} \sum_{(\alpha_1,\ldots,\alpha_n)\in K^n} (1 - f(\alpha_1, \ldots, \alpha_n)^{q-1}) &= q^n - \sum_{(\alpha_1,\ldots,\alpha_n)\in K^n} f(\alpha_1, \ldots, \alpha_n)^{q-1} \\ &= 0 - \sum_{(\alpha_1,\ldots,\alpha_n)\in K^n} f(\alpha_1, \ldots, \alpha_n)^{q-1}. \end{aligned}$$

Now f^{q-1} has degree $d(q-1)$, and we can write it as a linear combination of monomials of at most that degree. Let $\prod_{i=1}^n x_i^{\mu_i}$ be one such monomial. The degree of this monomial is $\sum_{i=1}^n \mu_i \leq d(q-1)$. By assumption, $n > d$, so for at least one j, $\mu_j < q-1$. Consider the sum $\sum_{(\alpha_1,\ldots,\alpha_n)\in K^n} \prod_{i=1}^n \alpha_i^{\mu_i} = \prod_{i=1}^n \sum_{\alpha_i \in K} \alpha_i^{\mu_i}$. The j^{th} term of this product is $\sum_{\alpha_j \in K} \alpha_j^{\mu_j}$. If $\mu_j = 0$, this is $\sum_{\alpha_j \in K} 1 = q = 0$. Otherwise, $0 < \mu_j < q-1$, so the sum is 0 by Lemma 2.2.1. Hence the product is 0, and the sum over $(\alpha_1, \ldots, \alpha_n) \in K^n$ of each monomial of f^{q-1} is 0, so $\sum_{(\alpha_1,\ldots,\alpha_n)\in K^n} f(\alpha_1, \ldots, \alpha_n)^{q-1} = 0$.

Thus the number of zeros of f in K^n is 0 mod p. □

Corollary 2.2.3. *Finite fields are C_1.*

Proof. Let f be a homogeneous polynomial over the finite field K of degree d in n variables, where $n > d$. By Theorem 2.2.2, the number of zeros of f is divisible by p. Now f has at least one zero (the trivial zero), so it has at least p zeros, and in particular it has at least $p-1$ nontrivial zeros. Thus K is C_1. □

2.3. **Extension Fields.** In this section, we will show that an extension field of a C_i field of finite transcendence degree j is C_{i+j}. The main idea is to expand a homogeneous polynomial according to a basis for the extension field into a vector of homogeneous polynomials over the base field. So we will need a tool (Theorem 2.3.7) for finding nontrivial common zeros of sets of homogeneous polynomials. The proof of this theorem relies on the concept of a normic form.

Definition 2.3.1. A *normic form* is a homogeneous polynomial ϕ of degree d in n variables such that $n = d$ and ϕ has only the trivial zero.

The name normic form comes from the following example.

Definition 2.3.2. Let E be a finite algebraic extension of a field F. For all $x \in E$, let $m_x : E \to E$ be the linear transformation $m_x(y) = xy$. The *norm* of x, denoted $N(x)$, is the determinant of m_x.

Example 2.3.3. Consider $\mathbb{C}$ as an algebraic extension of $\mathbb{R}$ of degree 2. Take $\{1, i\}$ as a basis for $\mathbb{C}$ over $\mathbb{R}$. For any complex numbers a and b, we can write $a = a_1 + a_2 i$ and $b = b_1 + b_2 i$ according to this basis. Then $ab = (a_1 b_1 - a_2 b_2) + (a_1 b_2 + a_2 b_1)i$.

Representing b as a vector and multiplication by a as a matrix, we have

$$m_a(b) = \begin{pmatrix} a_1 & -a_2 \\ a_2 & a_1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} a_1 b_1 - a_2 b_2 \\ a_1 b_2 + a_2 b_1 \end{pmatrix}.$$

Then $N(a) = |m_a| = a_1^2 + a_2^2$. Taking the coordinates a_1 and a_2 as variables, N is a homogeneous polynomial of degree 2 in 2 variables over $\mathbb{R}$, and it has only the trivial zero in $\mathbb{R}^2$, so N is a normic form.

Lemma 2.3.4 ([Gre69, Lemma 3.1])**.** *If E is a finite algebraic extension of F of degree $d > 1$, then the norm $N(x)$ is a normic form over F of degree d, whose variables are the d coordinates of x after choosing a basis for E as a vector space over F.*

Proof. Let $w_1, \ldots, w_d$ be a basis for E. We define the constants $c_j^{k,l}$ by

$$w_k w_l = \sum_{j=1}^{d} c_j^{k,l} w_j$$

for all $1 \leq k, l \leq d$. That is, $c^{k,l}$ is $w_k w_l$ expressed as a vector.

We will write the variable x as a vector in terms of this basis, $x = \sum_{k=1}^{d} x_k w_k$. Then for any $b \in E$, writing $b = \sum_{l=1}^{d} b_l w_l$,

$$\begin{aligned}
m_x(b) &= \left(\sum_{k=1}^{d} x_k w_k \right) \left(\sum_{l=1}^{d} b_l w_l \right) \\
&= \sum_{k=1}^{d} \sum_{l=1}^{d} x_k b_l w_k w_l \\
&= \sum_{j=1}^{d} \sum_{l=1}^{d} \sum_{k=1}^{d} x_k b_l c_j^{k,l} w_j \\
&= \begin{pmatrix} \sum_{l=1}^{d} \sum_{k=1}^{d} x_k b_l c_1^{k,l} \\ \vdots \\ \sum_{l=1}^{d} \sum_{k=1}^{d} x_k b_l c_d^{k,l} \end{pmatrix} \\
&= \begin{pmatrix} \sum_{k=1}^{d} x_k c_1^{k,1} & \cdots & \sum_{k=1}^{d} x_k c_1^{k,d} \\ \vdots & \ddots & \vdots \\ \sum_{k=1}^{d} x_k c_d^{k,1} & \cdots & \sum_{k=1}^{d} x_k c_d^{k,d} \end{pmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_d \end{pmatrix},
\end{aligned}$$

and we have determined the matrix representation of m_x.

The determinant of this matrix, $N(x)$, is a homogeneous polynomial of degree d in the variables $x_1, \ldots, x_d$. For $a \in E$, if $a \neq 0$, a has an inverse in E, so multiplication by a is invertible, m_a is an invertible matrix, and $N(a) \neq 0$. Thus N has only the trivial zero, and N is a normic form. □

Lemma 2.3.5 ([Gre69, Lemma 3.2])**.** *If a field F is not algebraically closed, then there exist normic forms over F of arbitrarily large degree.*

Proof. F is not algebraically closed, so it has some finite algebraic extension of degree $d > 1$. By Lemma 2.3.4, there is a normic form $\phi(x_1, \ldots, x_d)$ of degree d over F. Let $\phi_1(x_{1,1}, \ldots, x_{1,d}), \ldots, \phi_d(x_{d,1}, \ldots, x_{d,d})$ be d copies of ϕ, each with a set of d distinct variables. By substituting each ϕ_i for x_i in ϕ, we obtain $\phi^{(2)} = \phi(\phi_1, \ldots, \phi_d)$, which is a homogeneous polynomial of degree d^2 in d^2 variables.

Now ϕ has only the trivial zero, so at any zero of $\phi^{(2)}$, each ϕ_i must also take the value 0. But each ϕ_i has only the trivial zero, so $\phi^{(2)}$ has only the trivial zero, and thus $\phi^{(2)}$ is a normic form of degree d^2.

For all $m > 2$, we inductively define $\phi^{(m)} = \phi^{(m-1)}(\phi_1, \ldots, \phi_{d^{m-1}})$, where $\phi_1, \ldots, \phi_{d^{m-1}}$ are copies of ϕ, each with a distinct set of d variables. The same argument shows that $\phi^{(m)}$ is a normic form of degree d^m. Taking m arbitrarily large produces normic forms of arbitrarily large degree. □

We will assume the following theorem. It is not necessary to prove the Ax-Kochen Theorem, but it will allow us to state Theorems 2.3.7 and 2.3.9 so that they also cover the C_0 case. The proof can be easily located in a book on algebraic geometry, for example Hartshorne [Har77] Chapter 1, Theorem 7.2 is an equivalent statement.

Theorem 2.3.6. *Let F be an algebraically closed field. If $f_1, \ldots, f_r$ are homogeneous polynomials over F in n variables, where $n > r$, then they have a common nontrivial zero in F^n.*

When F is not algebraically closed, but is C_i for $i > 0$, we can use normic forms to demonstrate the existence of nontrivial common zeros.

Theorem 2.3.7 (Lang-Nagata, [Gre69, Theorem 3.4])**.** *Let F be a C_i field. Let $f_1, \ldots, f_r$ be homogeneous polynomials over F of degree d in n variables. If $n > rd^i$, then they have a nontrivial common zero in F^n.*

Proof. If F is algebraically closed, then F is C_0 by Theorem 2.1.6. So we have $n > rd^0 = r$, and by Theorem 2.3.6, the polynomials have a nontrivial common zero in F^n.

Otherwise, there is a normic form ϕ over F of degree $l \geq r$ in l variables by Lemma 2.3.5. For all $m \geq 1$, we define $\phi^{(m)}$ inductively, and we define D_m and N_m to be the degree and number of variables of $\phi^{(m)}$ respectively:

$$\begin{aligned}
\phi^{(1)} &= \phi(f_{1,1}, \ldots, f_{1,r}, f_{2,1}, \ldots, f_{2,r}, \ldots, f_{\lfloor \frac{l}{r} \rfloor,1}, \ldots, f_{\lfloor \frac{l}{r} \rfloor,r}, 0, \ldots, 0) \\
\phi^{(m)} &= \phi^{(m-1)}(f_{1,1}, \ldots, f_{1,r}, f_{2,1}, \ldots, f_{2,r}, \ldots, f_{\lfloor \frac{N_{m-1}}{r} \rfloor,1}, \ldots, f_{\lfloor \frac{N_{m-1}}{r} \rfloor,r}, 0, \ldots, 0),
\end{aligned}$$

where each set of polynomials $f_{j,1}, \ldots, f_{j,r}$ is a copy of the set $f_1, \ldots, f_r$ with a distinct set of n variables, $x_{j,1}, \ldots, x_{j,n}$. That is, for all j and k, $f_{j,k} = f_k(x_{j,1}, \ldots, x_{j,n})$. Note that we substitute as many complete sets as possible before padding with 0s.

We will prove by induction that for all $m \geq 1$, if $\phi^{(m)}$ has a nontrivial zero, then the $f_1, \ldots, f_r$ have a nontrivial common zero. For the base case, suppose that $\phi^{(1)}$ has a nontrivial zero, $\boldsymbol{\alpha} \in F^{N_1}$. We will denote by $\alpha_{j,k}$ the $x_{j,k}$-coordinate of $\boldsymbol{\alpha}$. Consider the values of the polynomials $f_{j,k}$ substituted into ϕ in the definition of $\phi^{(1)}$ at $\boldsymbol{\alpha}$. Since ϕ is normic, it has only the trivial zero, and thus all of the $f_{j,k}$ are 0. This means that for all j, $\alpha_{j,1}, \ldots, \alpha_{j,n}$ is a common zero for $f_1, \ldots, f_r$. Since $\boldsymbol{\alpha}$ is nontrivial, at least one of the $\alpha_{j,k}$ is nonzero, so for at least one j, $\alpha_{j,1}, \ldots, \alpha_{j,n}$ is a nontrivial common zero of $f_1, \ldots, f_r$.

Now suppose that for $m > 1$, $\phi^{(m)}$ has a nontrivial zero, $\boldsymbol{\alpha} \in F^{N_m}$. Consider the values of the polynomials $f_{j,k}$ substituted into $\phi^{(m-1)}$ at $\boldsymbol{\alpha}$. If they are all 0, then for at least one j, $\alpha_{j,1}, \ldots, \alpha_{j,n}$ is a nontrivial common zero of $f_1, \ldots, f_r$. If the values of the $f_{j,k}$ are not all 0, then these values constitute a nontrivial zero for $\phi^{(m-1)}$, and by induction we have a nontrivial common zero for $f_1, \ldots, f_r$.

Since F is C_i, it remains to show that for some m, $N_m > (D_m)^i$, since then $\phi^{(m)}$ has a nontrivial zero. We have $D_1 = dl$ and $N_1 = n\lfloor \frac{l}{r} \rfloor$, and for all $m > 1$, $D_m = dD_{m-1}$ and $N_m = n\lfloor \frac{N_{m-1}}{r} \rfloor$. Now,

$$\begin{aligned} \frac{N_m}{(D_m)^i} &\geq \frac{n(\frac{N_{m-1}}{r})}{(dD_{m-1})^i} \\ &\geq \left(\frac{n}{rd^i}\right)\left(\frac{N_{m-1}}{(D_{m-1})^i}\right) \end{aligned}$$

Expanding inductively,

$$\begin{aligned} \frac{N_m}{(D_m)^i} &\geq \left(\frac{n}{rd^i}\right)^{m-1}\left(\frac{N_1}{(D_1)^i}\right) \\ &\geq \left(\frac{n}{rd^i}\right)^{m-1}\left(\frac{n\left(\frac{l}{r}\right)}{(dl)^i}\right) \\ &\geq \left(\frac{n}{rd^i}\right)^{m} l^{1-i}. \end{aligned}$$

By assumption, $n > rd^i$, so $\lim_{m\to\infty} \frac{N_m}{(D_m)^i} = \infty$, and in particular $N_m > (D_m)^i$ for m large enough, as was to be shown. □

We are now in a position to prove our results about extension fields.

Theorem 2.3.8 ([Gre69, Theorem 3.5])**.** *If F is a C_i field, then every algebraic extension of F is C_i.*

Proof. It suffices to prove the theorem for any finite extension of F, since the coefficients of any polynomial lie in a finite extension.

Let E be a finite algebraic extension of F of degree e, and let $w_1, \ldots, w_e$ be a basis for E over F. Let f be a homogeneous polynomial over E of degree d in n variables,

$x_1, \ldots, x_n$, where $n > d^i$. We will write each variable in terms of the basis for E, substituting $\sum_{k=1}^{e} x_{j,k} w_k$ for each x_j and letting the $x_{j,k}$ vary over F.

Expanding, and writing f in terms of the basis for E, $f = \sum_{k=1}^{e} f_k w_k$, where the f_k are polynomials in the variables $x_{j,k}$. Each f_k is a linear combination of monomials of degree d, so the f_k are homogeneous polynomials of degree d in en variables over F.

Now f has a nontrivial zero in E^n if and only if the f_k have a nontrivial common zero in F^{en}. Such a zero exists by Theorem 2.3.7, since $en > ed^i$. □

Theorem 2.3.9 ([Gre69, Theorem 3.6])**.** *If F is a C_i field, and E is an extension of F of finite transcendence degree j, then E is C_{i+j}.*

Proof. By Theorem 2.3.8, we can reduce to the case in which E is a purely transcendental extension. Any purely transcendental extension of F of degree j is isomorphic to the field of rational functions in j variables, $F(t_1, \ldots, t_j)$. We will show that when $E = F(t)$, E is C_{i+1}. A simple induction on j then completes the proof.

The coefficients of homogeneous polynomials over $F(t)$ are, in general, rational functions. However, it suffices to consider homogeneous polynomials with coefficients in $F[t]$, the ring of polynomials, since we can clear denominators. That is, for $f \in F(t)[x_1, \ldots, x_n]$, if g is the product of the denominators of the coefficients of f, then $g^d f \in F[t][x_1, \ldots, x_n]$. But if $(a_1, \ldots, a_n)$ is a nontrivial zero of $g^d f$, then $(ga_1, \ldots, ga_n)$ is a nontrivial zero of f by Remark 2.1.2.

Let f be a homogeneous polynomial over $F[t]$ of degree d in n variables, $x_1, \ldots, x_n$, where $n > d^{i+1}$. For some $s > 0$, which we leave unspecified for now, substitute $\sum_{k=0}^{s} x_{j,k} t^k$ for each x_j, where the $x_{j,k}$ vary over F. If r is the highest degree (in terms of t) of the coefficients of f, then combining like powers of t, $f = \sum_{k=0}^{ds+r} f_k t^k$, where the f_k are polynomials in the variables $x_{j,k}$. Each f_k is a linear combination of monomials of degree d, so the f_k are homogeneous polynomials of degree d in $n(s+1)$ variables over F.

We can apply Theorem 2.3.7 if $n(s+1) > (ds + r + 1)d^i$, or equivalently, if $(n - d^{i+1})s > (r+1)d^i - n$. By assumption, $n > d^{i+1}$, so this inequality is satisfiable by picking s large enough. Then the theorem gives us a nontrivial common zero for the f_k in $F^{n(s+1)}$, which supplies a nontrivial zero of f in E^n. □

2.4. **Valued fields.** The goal of this section is to prove that the field of formal Laurent series over any finite field is C_2. Along the way we will develop the theory of discrete valued fields and their completions, which will allow us to define the p-adic fields. We begin with some definitions.

Definition 2.4.1. A *linearly ordered abelian group* is an abelian group G, together with an order relation $\leq$, such that for all $a, b, c \in G$,

(1) $\leq$ is a linear order on G, that is,
   (a) $a \leq b$ or $b \leq a$,
   (b) if $a \leq b$ and $b \leq a$, then $a = b$,
   (c) if $a \leq b$ and $b \leq c$, then $a \leq c$, and
(2) if $a \leq b$, then $a + c \leq b + c$.

We will sometimes write $b \geq a$ instead of $a \leq b$, and we will write $a < b$ to mean $a \leq b$ and $a \neq b$.

Definition 2.4.2. Let G be a linearly ordered abelian group, where we extend the order and group operation on G to include ∞, so that for all $b \in G \cup \{\infty\}$, $b \leq \infty$ and $b + \infty = \infty$. Given a field F and a map $\mathfrak{v} : F \to G \cup \{\infty\}$, such that for all $a, b \in F$,

(1) $\mathfrak{v}(a) = \infty$ if and only if $a = 0$,
(2) $\mathfrak{v}(ab) = \mathfrak{v}(a) + \mathfrak{v}(b)$, and
(3) $\mathfrak{v}(a + b) \geq \min(\mathfrak{v}(a), \mathfrak{v}(b))$,

we call F a *valued field* and $\mathfrak{v}$ a *valuation* on F.

Example 2.4.3. Given a field F and a linearly ordered abelian group G, F can be equipped with the trivial valuation $\mathfrak{v} : F \to G \cup \{\infty\}$ which maps 0_F to ∞ and all other elements to 0_G.

Example 2.4.4. For any prime p, define $\mathfrak{v}_p : \mathbb{Z} \setminus \{0\} \to \mathbb{N}$ by $\mathfrak{v}_p(a) = k$, where k is the maximum integer such that $p^k \mid a$. We can extend $\mathfrak{v}_p$ to a function $\mathbb{Q} \to \mathbb{Z} \cup \{\infty\}$ by setting $\mathfrak{v}_p(0) = \infty$ and $\mathfrak{v}_p(\frac{a}{b}) = \mathfrak{v}_p(a) - \mathfrak{v}_p(b)$. It is easy to verify that this extension is well-defined, and that $\mathbb{Q}$ and $\mathfrak{v}_p$ satisfy the conditions given in Definition 2.4.2.

Example 2.4.5. For any field F, we can define a similar valuation on the field of rational functions over F. Define $\mathfrak{v}_t : F[t] \setminus \{0\} \to \mathbb{N}$ by $\mathfrak{v}_t(f(t)) = k$, where k is the maximum integer such that $t^k \mid f(t)$. If $f(t) = a_n t^n + \ldots + a_0$, $\mathfrak{v}_t(f(t))$ is the minimum k such that $a_k \neq 0$. We can extend $\mathfrak{v}_t$ to a function $F(t) \to \mathbb{Z} \cup \{\infty\}$ by setting $\mathfrak{v}_t(0) = \infty$ and $\mathfrak{v}_t(\frac{f}{g}) = \mathfrak{v}_t(f) - \mathfrak{v}_t(g)$. It is easy to verify that this extension is well-defined, and that $F(t)$ and $\mathfrak{v}_t$ satisfy the conditions given in Definition 2.4.2.

We can immediately establish some simple facts about valuations.

Lemma 2.4.6. *Let F be a valued field with valuation $\mathfrak{v} : F \to G \cup \{\infty\}$. For all $a, b \in F$,*

(1) $\mathfrak{v}(1) = 0$,
(2) $\mathfrak{v}(a^{-1}) = -\mathfrak{v}(a)$
(3) $\mathfrak{v}(-a) = \mathfrak{v}(a)$,
(4) *if* $\mathfrak{v}(a) \neq \mathfrak{v}(b)$, *then* $\mathfrak{v}(a + b) = \min(\mathfrak{v}(a), \mathfrak{v}(b))$.

Proof. Property (2) in Definition 2.4.2 shows that $\mathfrak{v}$ is a homomorphism from F^* to G. Thus it maps the identity of F^* to the identity of G: $\mathfrak{v}(1) = 0$. Also, $\mathfrak{v}$ maps inverses in F^* to inverses in G: $\mathfrak{v}(a^{-1}) = -\mathfrak{v}(a)$.

Now by property (1a) in Definition 2.4.1, either $0 \leq \mathfrak{v}(-1)$ or $\mathfrak{v}(-1) \leq 0$. Suppose $0 \leq \mathfrak{v}(-1)$. Then by property (2), $\mathfrak{v}(-1) \leq \mathfrak{v}(-1) + \mathfrak{v}(-1) = \mathfrak{v}(-1 \cdot -1) = \mathfrak{v}(1) = 0$, so $\mathfrak{v}(-1) = 0$ by property (1b). The same argument holds if we assume $\mathfrak{v}(-1) \leq 0$, in which case $\mathfrak{v}(-1) \geq \mathfrak{v}(-1) + \mathfrak{v}(-1) = 0$, so $\mathfrak{v}(-1) = 0$. Hence for all $a \in F$, $\mathfrak{v}(-a) = \mathfrak{v}(-1) + \mathfrak{v}(a) = \mathfrak{v}(a)$.

If $\mathfrak{v}(a) \neq \mathfrak{v}(b)$, say $\mathfrak{v}(a) < \mathfrak{v}(b)$, then $\mathfrak{v}(a + b) \geq \min(\mathfrak{v}(a), \mathfrak{v}(b)) = \mathfrak{v}(a) = \mathfrak{v}(a + b + -b) \geq \min(\mathfrak{v}(a + b), \mathfrak{v}(-b))$. But we have the strict inequality $\mathfrak{v}(a) < \mathfrak{v}(b) = \mathfrak{v}(-b)$, so $\min(\mathfrak{v}(a+b), \mathfrak{v}(-b)) = \mathfrak{v}(a+b)$, and thus we have equality above: $\mathfrak{v}(a+b) = \min(\mathfrak{v}(a), \mathfrak{v}(b))$. The same argument holds if $\mathfrak{v}(b) < \mathfrak{v}(a)$. □

All valued fields come with a few interesting structures:

- $\mathfrak{v}$ is a homomorphism from F^* to G, so the image $\mathfrak{v}(F^*)$ is a linearly ordered subgroup of G, called the *value group*. Note that since $\mathfrak{v}(0) = \infty$, $\mathfrak{v}(F) = \mathfrak{v}(F^*) \cup \{\infty\}$.

- We define $\mathcal{O}_F = \{a \in F \mid \mathfrak{v}(a) \geq 0\}$. The set $\mathcal{O}_F$ contains 1 and 0, and is closed under addition, multiplication, and additive inverse, so it is a subring of F, called the *valuation ring*. We will denote the valuation ring by $\mathcal{O}$ when there is no ambiguity. For all $a \in F$, $a \in \mathcal{O}$ or $a^{-1} \in \mathcal{O}$, since if $\mathfrak{v}(a) < 0$, then $\mathfrak{v}(a) + \mathfrak{v}(a^{-1}) < \mathfrak{v}(a^{-1})$, and thus $\mathfrak{v}(a^{-1}) > \mathfrak{v}(1) = 0$. Hence the field of fractions of $\mathcal{O}$ is F.
- For $a \in \mathcal{O}$, if also $a^{-1} \in \mathcal{O}$, then $\mathfrak{v}(a) + \mathfrak{v}(a^{-1}) = \mathfrak{v}(a) + -\mathfrak{v}(a) = 0$, but both $\mathfrak{v}(a) \geq 0$ and $\mathfrak{v}(a^{-1}) \geq 0$, so $\mathfrak{v}(a) = \mathfrak{v}(a^{-1}) = 0$. Conversely, if $\mathfrak{v}(a) = 0$, then $\mathfrak{v}(a^{-1}) = -\mathfrak{v}(a) = 0$, and $a^{-1} \in \mathcal{O}$. Thus $\{a \in F \mid \mathfrak{v}(a) = 0\}$ is $\mathcal{O}^*$, the group of units of $\mathcal{O}$.
- We define $I_1 = \{a \in \mathcal{O} \mid \mathfrak{v}(a) > 0\}$. The set I_1 is an ideal in $\mathcal{O}$, since it is closed under addition, and if $a \in I_1$, $b \in \mathcal{O}$, then $\mathfrak{v}(ab) = \mathfrak{v}(a) + \mathfrak{v}(b) > 0$, so $ab \in I_1$. Moreover, it is a maximal ideal, since if I is an ideal in $\mathcal{O}$ properly containing I_1, then there is $a \in I$ with $\mathfrak{v}(a) = 0$, so a is a unit, and thus $I = \mathcal{O}$.
- We define $\overline{F} = \mathcal{O}/I_1$. Since I_1 is a maximal ideal, $\overline{F}$ is a field, called the *residue class field*. The residue class of $a \in \mathcal{O}$ mod I_1 is denoted $\overline{a}$.

Discrete Valued Fields. The value group of a discrete valued field is isomorphic to $\mathbb{Z}$. The structure imposed by the valuation axioms allows us to complete such a field in a way analogous to how $\mathbb{R}$ is obtained by a completion of $\mathbb{Q}$. Our two main examples of valued fields, $\mathbb{F}_p(t)$ and $\mathbb{Q}$, are discrete valued fields with completions $\mathbb{F}_p((t))$ and $\mathbb{Q}_p$ respectively.

Definition 2.4.7. A valued field F with valuation $\mathfrak{v}$ is called *discrete* if its value group $\mathfrak{v}(F^*)$ is isomorphic to $\mathbb{Z}$ with its usual ordering. Call the isomorphism ϕ. An element $\pi \in F$ is called a *prime element* if $\phi(\mathfrak{v}(\pi)) = 1$.

For notational convenience, we will suppress the isomorphism ϕ and identify the value groups of our discrete valued fields with the integers.

The fields $\mathbb{Q}$ and $F(t)$ with valuations $\mathfrak{v}_p$ and $\mathfrak{v}_t$ defined in Examples 2.4.4 and 2.4.5 are discrete valued fields.

Lemma 2.4.8. *Let F be a discrete valued field with valuation $\mathfrak{v}$. Let π be a prime element in F. Then for $n \in \mathbb{Z}$, any $a \in F^*$ with $\mathfrak{v}(a) = n$ can be written as $u\pi^n$ with $u \in \mathcal{O}^*$, and for all $n \geq 1$, the set $I_n = \{a \in F \mid \mathfrak{v}(a) \geq n\}$ is a principal ideal of $\mathcal{O}$, generated by π^n.*

Proof. For all $a \in F^*$, let $n = \mathfrak{v}(a)$. Then $\mathfrak{v}(a\pi^{-n}) = \mathfrak{v}(a) + \mathfrak{v}(\pi^{-n}) = n - n = 0$, so $u = a\pi^{-n}$ is a unit in $\mathcal{O}$. We can write $a = u\pi^n$, with $u \in \mathcal{O}^*$.

Now for all $n \geq 1$, I_n is closed under addition, and if $a \in I_n$, $y \in \mathcal{O}$, then $\mathfrak{v}(ay) = \mathfrak{v}(a) + \mathfrak{v}(y) \geq n + 0 = n$, so $ay \in I_n$. Thus I_n is an ideal. For all $a \in I_n$, $a = u\pi^{\mathfrak{v}(a)} = u\pi^n\pi^{\mathfrak{v}(a)-n}$ for some unit u, and $\pi^n \in I_n$, so I_n is generated by π^n. □

For all $n \geq 0$, we define the ring $\mathcal{O}_n = \mathcal{O}/I_{n+1}$, the ring of cosets mod π^{n+1}. Note that $\mathcal{O}_0 = \mathcal{O}/I_1 = \overline{F}$. Choose a set of coset representatives $A = \{\alpha_i\} \subset \mathcal{O}$ for the elements of $\overline{F}$. For any $a \in \mathcal{O}$, if $\overline{a} = \overline{\alpha_{i_0}}$, then $a - \alpha_{i_0} \in M$, so $a = \alpha_{i_0} + a_1\pi$ for some $a_1 \in \mathcal{O}$. Repeating this process, if $\overline{a_1} = \overline{\alpha_{i_1}}$, then $a_1 = \alpha_{i_1} + a_2\pi$ for some $a_2 \in \mathcal{O}$, so $a = \alpha_{i_0} + \alpha_{i_1}\pi + a_2\pi^2$. Thus $a \equiv \alpha_{i_0} + \alpha_{i_1}\pi \pmod{\pi^2}$, and $\alpha_{i_0} + \alpha_{i_1}\pi$ is the image of a in $\mathcal{O}_1$.

Continuing in this way, we can represent any element of $\mathcal{O}_n$ uniquely as $\alpha_{i_0} + \alpha_{i_1}\pi + \ldots + \alpha_{i_n}\pi^n$ for $\alpha_{i_0}, \ldots, \alpha_{i_n} \in A$. For all $n > 0$, let ϕ_n be the canonical homomorphism $\mathcal{O}_n \to \mathcal{O}_{n-1}$

which maps an element of $\mathcal{O}_n$ to its coset mod π^n. Under this representation of $\mathcal{O}_n$, ϕ_n simply omits the leading term $\alpha_{i_n}\pi^n$.

Definition 2.4.9. Let F be a discrete valued field. Define the rings $\mathcal{O}_n$ for all $n \geq 0$ and homomorphisms ϕ_n for all $n > 0$ as above. The *completion* of $\mathcal{O}$, $\widehat{\mathcal{O}}$, is defined by

$$\widehat{\mathcal{O}} = \{(a_0, a_1, \ldots) \in \prod_{n\geq 0} \mathcal{O}_n \,|\, \forall n > 0, \phi_n(a_n) = a_{n-1}\}.$$

$\widehat{\mathcal{O}}$ is a subring of the product ring $\prod_{n\geq 0} \mathcal{O}_n$. The *completion* of F, $\widehat{F}$, is defined to be the field of fractions of $\widehat{\mathcal{O}}$.

Those familiar with category theory will recognize this as the inverse limit construction.

Using the representation of $\mathcal{O}_n$ as $\alpha_{i_0} + \alpha_{i_1}\pi + \ldots + \alpha_{i_n}\pi^n$ for $\alpha_{i_0}, \ldots, \alpha_{i_n} \in A$, an arbitrary element of the completion $a \in \widehat{\mathcal{O}}$ looks like $a = (\alpha_{i_0}, \alpha_{i_0} + \alpha_{i_1}\pi, \alpha_{i_0} + \alpha_{i_1}\pi + \alpha_{i_2}\pi^2, \ldots)$ with $\alpha_{i_0}, \alpha_{i_1}, \ldots \in A$. For convenience, we will express this element as an infinite sum: $\alpha_{i_0} + \alpha_{i_1}\pi + \alpha_{i_2}\pi^2 + \ldots$, which is well-defined, since the k^{th} coordinate of a provides the coefficient α_{i_k} of π^k, while agreeing with the previous coordinates on the coefficients α_{i_j} for all $j < k$.

Remark 2.4.10. The completion of F, which we have defined purely algebraically, is isomorphic to the analytic completion of F under the metric induced by the absolute value $||a||_{\mathfrak{v}} = 2^{-\mathfrak{v}(a)}$. The elements of $\widehat{\mathcal{O}}$ correspond to equivalence classes of Cauchy sequences under this metric.

Example 2.4.11. For any field F, the field of rational functions $F(t)$ with the valuation $\mathfrak{v}_t$ defined in Example 2.4.5 is a discrete valued field. We will see that its completion is $F((t))$, the field of formal Laurent series over F.

We have $\mathfrak{v}_p(t) = 1$, and we will choose $\pi = t$ as a prime element. Writing all fractions in lowest terms, we have $\mathcal{O} = \{\frac{f}{g} \in F(t) \,|\, t \nmid g\}$, with maximal ideal $I_1 = \{\frac{f}{g} \in F(t) \,|\, t \,|\, f, t \nmid g\}$.

Now for any rational function $\frac{f}{g} \in F(t)$, with $t \nmid g$, let $h \in F$ be the inverse of the constant term of g. Then $hg \equiv 1 \pmod{t}$. Let $l = \frac{f(hg-1)}{t} \in F[t]$. Then $\frac{f}{g} + \frac{tl}{g} = \frac{f+tl}{g} = \frac{f+f(hg-1)}{g} = \frac{fhg}{g} = fh \in F[t]$. We chose $\frac{tl}{g} \in I_1$, so this shows that any element of $\mathcal{O}$ is congruent to an an element of $F[t]$ mod I_1. Since $t \in I_1$, any element of $\mathcal{O}$ is congruent to an element of F mod I_1.

Hence the residue class field $\overline{F(t)}$ is isomorphic to F, and we can take F as our set of coset representatives.

Now we will take the completion of $\mathcal{O}$. The resulting ring is $\widehat{\mathcal{O}} = F[[t]]$, the field of formal power series over F. As we saw above, the elements of the completion can be uniquely represented in the form $\alpha_0 + \alpha_1 t + \alpha_2 t^2 + \ldots$, with each $\alpha_i \in F$.

The fraction field of $F[[t]]$ is the completion $\widehat{F(t)}$. Let $x = \frac{\alpha_0+\alpha_1 t+\ldots}{\beta_0+\beta_1 t+\ldots} \in \widehat{F(t)}$. Let k be the least integer such that $\beta_k \neq 0$. Now factoring out the leading term $\beta_k t^k$, we can write

$$x = \left(\frac{1}{\beta_k t^k}\right)\left(\frac{\alpha_0 + \alpha_1 t + \ldots}{1 + \gamma_1 t + \ldots}\right),$$

where $\gamma_i = \beta_{k+i}\beta_k^{-1} \in F$ for all $i \geq 1$.

We claim that the inverse of the denominator, $(1 + \gamma_1 t + \ldots)^{-1}$, is an element of $F[[t]]$. We have

$$\begin{aligned} \frac{1}{1+\gamma_1 t + \ldots} &= \frac{1}{1-(-\gamma_1 t - \ldots)} \\ &= 1 + (-\gamma_1 t - \ldots) + (-\gamma_1 t - \ldots)^2 + \ldots, \end{aligned}$$

applying the geometric series formula. Now for all $n \geq 0$, t^n appears in only finitely many terms of the infinite sum, so the coefficient of each t^n is well-defined, and this is a well-defined element of $F[[t]]$.

Letting $y = \frac{1}{1+\gamma_1 t+\ldots} \in F[[t]]$, we can write $x = \beta_k^{-1} t^{-k} y(\alpha_0 + \alpha_1 t + \ldots)$, and this has the form $c_{-k}t^{-k} + \ldots + c_{-1}t^{-1} + c_0 + c_1 t + \ldots$, with each $c_i \in F$. All elements of $\widehat{F(t)}$ can be uniquely represented in this form. We call the completion the field of formal Laurent series over F and denote it by $F((t))$.

Example 2.4.12. For all primes p, we define the field of p-adic numbers, $\mathbb{Q}_p$ to be the completion of $\mathbb{Q}$ according to the valuation $\mathfrak{v}_p$ defined in Example 2.4.4.

We have $\mathfrak{v}_p(p) = 1$, and we will choose $\pi = p$ as a prime element. Writing all fractions in lowest terms, we have $\mathcal{O} = \{\frac{a}{b} \in \mathbb{Q} \,|\, p \nmid b\}$, with maximal ideal $I_1 = \{\frac{a}{b} \in \mathbb{Q} \,|\, p \,|\, a, p \nmid b\}$.

Now for any $\frac{a}{b} \in \mathbb{Q}$, with $p \nmid b$, there is some $d \in \mathbb{Z}$ such that $db \equiv 1 \pmod{p}$. Let $c = \frac{a(db-1)}{p} \in \mathbb{Z}$. Then $\frac{a}{b} + \frac{pc}{b} = \frac{a+pc}{b} = \frac{a+a(db-1)}{b} = \frac{adb}{b} = ad \in \mathbb{Z}$. We chose $\frac{pc}{b} \in I_1$, so this shows that any element of $\mathcal{O}$ is congruent to an integer mod I_1. Since all integer multiples of p are in I_1, any element of $\mathcal{O}$ is congruent to one of $\{0, 1, \ldots, p-1\}$ mod I_1.

Thus the residue class field has p elements, $\overline{\mathbb{Q}} \cong \mathbb{F}_p$, and we can take as our set of coset representatives $A = \{0, 1, \ldots, p-1\}$.

Now we will take the completion of $\mathcal{O}$. The resulting ring is $\widehat{\mathcal{O}} = \mathbb{Z}_p$, the p-adic integers. Its elements can be uniquely represented in the form $\alpha_0 + \alpha_1 p + \alpha_2 p^2 + \ldots$, with each $\alpha_i \in A$.

The p-adic field $\mathbb{Q}_p$ is the field of fractions of $\mathbb{Z}_p$. By a similar argument to the one in Example 2.4.11, elements of $\mathbb{Q}_p$ can be uniquely represented in the form $c_{-k}p^{-k} + \ldots + c_{-1}p^{-1} + c_0 + c_1 p + \ldots$, with each $c_i \in A$.

The elements of $\mathbb{Q}_p$ look very similar to the elements of $\mathbb{F}_p((t))$. They can be thought of as formal Laurent series in a single "variable", p, with coefficients in $\mathbb{F}_p$. The similarity between the two fields is significant to us because it was the motivation for Artin's conjecture that $\mathbb{Q}_p$ is C_2 (see Theorem 2.4.16). However, the fields $\mathbb{Q}_p$ and $\mathbb{F}_p((t))$ are not isomorphic; their arithmetic is very different. In particular, $\mathbb{Q}_p$ has characteristic 0, while $\mathbb{F}_p((t))$ has characteristic p. Informally speaking, elements of $\mathbb{Q}_p$ add and multiply with carries, while elements of $\mathbb{F}_p((t))$ do not. Nevertheless, the Ax-Kochen Principle demonstrates that the similarity between the fields is not just skin-deep.

Homogeneous Polynomials over Complete Discrete Valued Fields. For any discrete valued field F with valuation ring $\mathcal{O}$, there is a homomorphism $i : \mathcal{O} \to \widehat{\mathcal{O}}$ which maps $x \in \mathcal{O}$ to the images of x in $\mathcal{O}_n$ for all $n \geq 0$. The only element of $\mathcal{O}$ divisible by all powers of π is 0, so i is injective. Thus we regard $\mathcal{O}$ as a subring of $\widehat{\mathcal{O}}$ and F as a subfield of $\widehat{F}$. If i is surjective, then $\mathcal{O} \cong \widehat{\mathcal{O}}$ and $F \cong \widehat{F}$. In this case, we say that $\mathcal{O}$ and F are complete.

Lemma 2.4.13. *As one would hope, the completion of a discrete valued field is a complete discrete valued field.*

Proof. Let F be a discrete valued field with valuation $\mathfrak{v}$, valuation ring $\mathcal{O}$, and prime element π. Let $\widehat{\mathcal{O}}$ be the completion of $\mathcal{O}$, and let $\widehat{F}$ be the completion of F (the field of fractions of $\mathcal{O}$). First we must show that $\widehat{F}$ is a discrete valued field.

Define a function $\widehat{\mathfrak{v}} : \widehat{O} \setminus \{0\} \to \mathbb{N}$ which takes an element of $\widehat{O}$, $(a_0, a_1, \ldots)$ to the least integer k such that $a_k \neq 0$. Note that $\widehat{\mathfrak{v}}$ agrees with $\mathfrak{v}$ on the subring $\mathcal{O}$, since for $a \in \mathcal{O}$, the image of a in $\widehat{\mathcal{O}}$ is $(a_0, a_1, \ldots)$ where a_i is the image of a mod π^{i+1}, and k is the least integer such that $a_k \neq 0$ if and only if k is the greatest integer such that $\pi^k \mid a$.

Now we can extend $\widehat{\mathfrak{v}}$ to a function $\widehat{F} \to \mathbb{Z} \cup \{\infty\}$ by setting $\widehat{\mathfrak{v}}(0) = \infty$ and $\widehat{\mathfrak{v}}(\frac{a}{b}) = \widehat{\mathfrak{v}}(a) - \widehat{\mathfrak{v}}(b)$. It is easy to verify that $\widehat{\mathfrak{v}}$ satisfies the valuation axioms. Thus $\widehat{F}$ is a discrete valuation field.

Since $\widehat{\mathfrak{v}}$ agrees with $\mathfrak{v}$ on elements of $\mathcal{O}$, we can choose the same element π as a prime element of $\widehat{F}$. Then for each $n \geq 1$, the ideal $\widehat{I}_n$ consists of all elements which are 0 in their first n coordinates, and for $n \geq 0$, elements of the ring $\widehat{\mathcal{O}}_n = \widehat{\mathcal{O}}/\widehat{I}_{n+1}$ are cosets consisting of elements which agree on their first $n+1$ coordinates.

The inclusion $i : \mathcal{O} \to \widehat{\mathcal{O}}$ induces homomorphisms $i_n : \mathcal{O}_n \to \widehat{\mathcal{O}}_n$. Included in $\widehat{\mathcal{O}}$, the ideal I_n consists of all elements which are 0 in their first n coordinates, and elements of the ring $\mathcal{O}_n$ are cosets consisting of elements which agree on their first $n+1$ coordinates, so the i_n are bijections, and therefore isomorphisms.

Now $\widehat{\mathcal{O}}_n \cong \mathcal{O}_n$ for all $n > 0$, and since the prime element π is the same, $\widehat{\phi_n}$ and ϕ_n act on $\widehat{\mathcal{O}}_n$ and $\mathcal{O}_n$ in the same way. Hence the completions of $\widehat{\mathcal{O}}$ and $\mathcal{O}$ are isomorphic, that is, the completion of $\widehat{\mathcal{O}}$ is isomorphic to $\widehat{\mathcal{O}}$, and thus $\widehat{\mathcal{O}}$ and $\widehat{F}$ are complete. □

We now return to studying homogeneous polynomials and the C_i properties. Let f be a homogeneous polynomial over a discrete valued field F of degree d, and suppose that all of the coefficients of f are in the valuation ring $\mathcal{O}$. Fixing an $m \geq 0$ and a prime element π, we will denote by $\overline{f}$ the reduction of f mod π^{m+1}, the coefficients of which are in the quotient ring $\mathcal{O}_m$. Note that $\overline{f}$ is either a homogeneous polynomial of degree d or the zero polynomial, if all coefficients are divisible by π^{m+1}.

Definition 2.4.14. Let f be a homogeneous polynomial in n variables over a discrete valued field F with prime element π, and suppose that all of the coefficients of f are in the valuation ring $\mathcal{O}$. We will call a zero $(\alpha_1, \ldots, \alpha_n) \in \mathcal{O}^n$ of f *primitive* if for some j, $\pi \nmid \alpha_j$. Similarly, for $m \geq 0$, we call a zero $(\overline{\alpha_1}, \ldots, \overline{\alpha_n}) \in \mathcal{O}_m^n$ of $\overline{f}$ *primitive* if for some j, $\overline{\pi} \nmid \overline{\alpha_j}$, where $\overline{\pi}$ is the coset of π mod π^{m+1}.

Our goal is to reduce the problem of finding zeros of f in $\mathcal{O}^n$ to the problem of finding zeros of $\overline{f}$ in $\mathcal{O}_m^n$ for all $m \geq 0$. The advantage of working with primitive zeros is that a primitive zero cannot become trivial upon reduction mod π^{m+1}.

Theorem 2.4.15 ([Gre69, Theorem 4.7])**.** *Let F be a complete discrete valued field with prime element π. Suppose that the residue class field $\overline{F}$ is finite. Then a homogeneous polynomial f over $\mathcal{O}$ of degree d in n variables has a primitive zero in $\mathcal{O}^n$ if and only if $\overline{f}$ has a primitive zero $\mathcal{O}_m^n$ for all $m \geq 0$.*

Proof. Suppose $(\alpha_1, \ldots, \alpha_n) \in \mathcal{O}^n$ is a primitive zero of f. Then for all $m \geq 0$, reduction mod π^{m+1} is a homomorphism $\mathcal{O} \to \mathcal{O}_m$, so $\overline{f}(\overline{\alpha_1}, \ldots, \overline{\alpha_n}) = 0$. Since $(\alpha_1, \ldots, \alpha_n)$ is primitive, for some j, $\pi \nmid \alpha_j$. Suppose $\overline{\pi} \mid \overline{\alpha_j}$. Then there is some $\overline{d} \in \mathcal{O}_m$ such that $\overline{\pi}\overline{d} = \overline{\alpha_j}$, so lifting to $\mathcal{O}$, $\pi d - \alpha_j \in I_{m+1} \subset I_1$. Now $\pi d \in I_1$, so $\alpha_j \in I_1$, and thus π divides α_j. This is a contradiction, so $\overline{\pi} \nmid \overline{\alpha_j}$, and $(\overline{\alpha_1}, \ldots, \overline{\alpha_n})$ is a primitive zero in $\mathcal{O}_m^n$.

Conversely, for all $m \geq 0$, define $S_m \subseteq \mathcal{O}_m^n$ to be the set of primitive zeros of $\overline{f}$ in $\mathcal{O}_m^n$, and suppose that S_m is nonempty for all m. If $\alpha \in S_{m+1}$ is a primitive zero mod π^{m+2}, then its image mod π^{m+1} is also a primitive zero; that is, $\phi_{m+1}(\alpha) \in S_m$, so $\phi_{m+1}(S_{m+1}) \subseteq S_m$. For all $j < m$, define $S_{m,j} = \phi_{j+1}(\phi_{j+2}(\ldots \phi_m(S_m))) \subseteq S_j$. The set $S_{m,j}$ is the set of primitive zeros in $\mathcal{O}_j^n$ which lift to primitive zeros in $\mathcal{O}_m^n$. Since all the S_m are nonempty, all the $S_{m,j}$ are nonempty.

For all $k \geq 0$, define $T_k = \bigcap_{m>k} S_{m,k}$. T_k is the set of all solutions in $\mathcal{O}_k^n$ which lift to solutions in $\mathcal{O}_m^n$ for all $m > k$. Since $\overline{F}$ is finite, all of the $\mathcal{O}_m$ are finite. The chain $S_{k+1,k} \supseteq S_{k+2,k} \supseteq \ldots$ must break off at some $l > k$, with $S_{m,k} = S_{l,k}$ for all $m \geq l$, since the decreasing sequence of integers $|S_{k+1,k}| \geq |S_{k+2,k}| \geq \ldots$ is bounded below by 1. Thus $T_k = S_{l,k}$ is nonempty for all k.

Pick a zero $a_0 = (\alpha_{0,0}, \ldots, \alpha_{0,n}) \in T_0$. For all m, a_0 lifts to a solution $a_m = (\alpha_{m,0}, \ldots, \alpha_{m,n})$ in $\mathcal{O}_m^n$. That is, assuming that $a_{i-1} \in T_{i-1}$, we can choose $a_i \in \phi_i^{-1}(a_{i-1}) \cap T_i$. By construction, the sequence $(a_0, a_1, \ldots)$ satisfies $\phi_m(a_m) = a_{m-1}$, so the sequences $(\alpha_{0,i}, \alpha_{1,i}, \ldots)$ are elements of the completion $\widehat{\mathcal{O}}$ for all i.

Since $a_m \in S_m$, a_m is a zero of $\overline{f}$ mod π^{m+1} for all $m \geq 0$. Hence, viewing f as a polynomial in the completion by the isomorphism between the complete ring $\mathcal{O}$ and $\widehat{\mathcal{O}}$,

$$f((\alpha_{0,0}, \alpha_{1,0}, \ldots), \ldots, (\alpha_{0,n}, \alpha_{1,n}, \ldots)) = (\overline{f}(\alpha_{0,0}, \ldots, \alpha_{0,n}), \overline{f}(\alpha_{1,0}, \ldots, \alpha_{1,n}), \ldots) = 0,$$

and this is a primitive zero of f in $\widehat{\mathcal{O}}^n$. But $\mathcal{O} \cong \widehat{\mathcal{O}}$, so each $(\alpha_{0,i}, \alpha_{1,i}, \ldots)$ corresponds to an element of $\mathcal{O}$, and this zero corresponds to a primitive zero of f in $\mathcal{O}^n$. □

Theorem 2.4.15 also holds without the assumption that $\overline{F}$ is finite, but the proof of the general version is more difficult, and we will only need the finite case for the Ax-Kochen Theorem.

Theorem 2.4.16 ([Gre69, Corollary 4.9])**.** *If F is a finite field, then $F((t))$ is C_2.*

Proof. It suffices to consider homogeneous polynomials with coefficients in $F[[t]]$, the valuation ring of $F((t))$, since we can clear denominators. That is, for $f \in F((t))[x_1, \ldots, x_n]$ homogeneous of degree d in n variables, if c is the minimum valuation among the coefficients of f, then $t^{cd} f \in F[[t]][x_1, \ldots, x_n]$ is also homogeneous of degree d in n variables. If $(\alpha_1, \ldots, \alpha_n)$ is a nontrivial zero of $t^{cd} f$, then $f(t^c \alpha_1, \ldots, t^c \alpha_n) = t^{cd} f(\alpha_1, \ldots, \alpha_n) = 0$ by Remark 2.1.2, so $(t^c \alpha_1, \ldots, t^c \alpha_n)$ is a nontrivial zero of f.

Let f be a homogeneous polynomial over $F[[t]]$ of degree d in n variables, where $n > d^2$. The residue class field $\overline{F((t))} \cong F$ is finite, and $F((t))$ is complete by Lemma 2.4.13, so we can apply Theorem 2.4.15. Since any primitive zero in $F[[t]]^n$ is a nontrivial in $F((t))^n$, it suffices to find a primitive zero of $\overline{f}$ in the residue ring mod t^{m+1} for all $m \geq 0$.

Fixing $m \geq 0$, let $\widetilde{f}$ be the polynomial obtained by ignoring the terms of degree greater than m in each coefficient of f. Each coefficient of $\widetilde{f}$ is then a polynomial in t of degree at

most m. Now $\widetilde{f}$ is either the zero polynomial or a homogeneous polynomial of degree d in n variables. In the first case, each coefficient of f is divisible by t^{m+1}, so reducing mod t^{m+1}, $\overline{f}$ is the zero polynomial, which clearly has a primitive zero, and we are done.

Otherwise, we will view $\widetilde{f}$ as a polynomial over $F(t)$. By Corollary 2.2.3, F is C_1, and by Theorem 2.3.9, $F(t)$ is C_2. Since $n > d^2$, $\widetilde{f}$ has a nontrivial zero $(\alpha_1, \ldots, \alpha_n)$ in $F(t)$. Using the homogeneity of $\widetilde{f}$, we can normalize to find another zero in $F[t]$ which is primitive.

Let α_j be the coordinate with minimum (possibly negative) valuation, and let $c = \mathfrak{v}(\alpha_j)$. Now let $(\beta_1, \ldots, \beta_n) = (t^{-c}\alpha_1, \ldots, t^{-c}\alpha_n)$. All of the β_i are elements of $F[t]$, since $\mathfrak{v}(\beta_i) = \mathfrak{v}(t^{-c}) + \mathfrak{v}(\alpha_i) \geq -c + c = 0$.

Now $t^{cd}\widetilde{f}(\beta_1, \ldots, \beta_n) = \widetilde{f}(t^c\beta_1, \ldots, t^c\beta_n) = \widetilde{f}(\alpha_1, \ldots, \alpha_n) = 0$. Now $F[t]$ is a domain so since $t^{cd} \neq 0$, $\widetilde{f}(\beta_1, \ldots, \beta_n) = 0$. Moreover, $\mathfrak{v}(\beta_j) = -c + c = 0$, so $t \nmid \beta_j$, and $(\beta_1, \ldots, \beta_n)$ is a primitive zero.

Finally, since $\widetilde{f}(\beta_1, \ldots, \beta_n) = 0$, and $\widetilde{f}$ corresponds to $\overline{f}$ mod t^{m+1}, viewing $\beta_1, \ldots, \beta_n$ as elements of $F[[t]]$ by the natural inclusion, $(\overline{\beta_1}, \ldots, \overline{\beta_n})$ is a zero of $\overline{f}$ mod t^{m+1}. Since the zero is primitive, its reduction mod t^{m+1} is also primitive, as was to be shown. □

As a special case of Theorem 2.4.16, we have that $\mathbb{F}_p((t))$ is C_2 for all primes p. Because of the resemblance between the fields $\mathbb{F}_p((t))$ and $\mathbb{Q}_p$, Artin conjectured that $\mathbb{Q}_p$ is also C_2 for all primes p. This conjecture turned out to be false, but Ax and Kochen were able to prove a weaker statement: for each degree d there exists a finite set of primes $P(d)$ such that the C_2 property holds for polynomials of degree d in $\mathbb{Q}_p$ for all $p \notin P(d)$.

The Ax-Kochen Theorem is a corollary of a much more general Ax-Kochen Principle: any first-order logical statement which is true of all but finitely many of the fields $\mathbb{F}_p((t))$ is true of all but finitely many of the fields $\mathbb{Q}_p$. This statement is what could be called a "meta-theorem", since it quantifies over logical statements. In order to prove it we will need to develop techniques for reasoning about logical statements in general.

3. The Language of Model Theory

3.1. **Languages, Models, and Theories.** Model theory is concerned with the study of mathematical structures and the logical statements about them. Logical statements about a structure are built from the familiar boolean operators and quantifiers, but they must also refer to the distinguished elements, functions, and relations which are inherent to the structure in question. Thus we work in terms of formal languages of symbols representing these elements, functions, and relations.

Definition 3.1.1. A *language* is the union of

- $\mathcal{C}$, a set of constant symbols,
- $\mathcal{F}$, a set of function symbols, with an integer $n_f > 0$ for each $f \in \mathcal{F}$, and
- $\mathcal{R}$, a set of relation symbols, with an integer $n_R > 0$ for each $R \in \mathcal{R}$.

The integers n_f and n_R are called the arities of the corresponding functions and relations. An n-ary function takes n arguments, and an n-ary relation is a relation on n elements. Most of the function and relation symbols we encounter will have $n = 1$ or $n = 2$, called unary and binary respectively.

Definition 3.1.2. Let $\mathcal{L} = \mathcal{C} \cup \mathcal{F} \cup \mathcal{R}$ be a language. An *$\mathcal{L}$-structure* $\mathcal{M}$ is

- a set $M \neq \emptyset$, the domain,
- an element $c^{\mathcal{M}} \in M$ for all $c \in \mathcal{C}$,
- a function $f^{\mathcal{M}} : M^{n_f} \to M$ for all $f \in \mathcal{F}$, and
- a relation $R^{\mathcal{M}} \subseteq M^{n_R}$ for all $R \in \mathcal{R}$.

The elements, functions, and relations $c^{\mathcal{M}}$, $f^{\mathcal{M}}$, and $R^{\mathcal{M}}$ are called the interpretations of the $\mathcal{L}$-symbols in $\mathcal{M}$. The distinction between a symbol and its interpretation in a given structure is very important. This division between syntax and semantics will allow us to define and study logical theories independently of any particular structure.

Example 3.1.3. Let $\mathcal{L}_G$ be the language $\{\cdot, e\}$, where $\cdot$ is a binary function symbol and e is a constant symbol. These symbols are necessary to describe the theory of groups, and the symbols of $\mathcal{L}_G$ can be interpreted in any group. For instance, the group $\langle \mathbb{Z}, +, 0 \rangle$ is an $\mathcal{L}_G$-structure under the interpretations $\cdot^{\mathbb{Z}} = +$ and $e^{\mathbb{Z}} = 0$. But any nonempty set with any binary function can also be an $\mathcal{L}_G$ structure. For example, if $X = \{a, b, c\}$, then X is an $\mathcal{L}_G$ structure under the interpretations $\cdot^X : (x, y) \mapsto b$ for all $x, y \in X$, and $e^X = c$, despite the fact that $\langle X, \cdot^X, c \rangle$ is clearly not a group.

A valued field is more difficult to formalize as a structure, since its definition relies on an auxiliary structure, the value group. We will use a property called cross section to deal with the value group as a substructure of the field itself.

Definition 3.1.4. A valued field F is called a *valued field with cross section* if there is an injective map $i : \mathfrak{v}(F) \to F$ such that i is a group homomorphism from $\mathfrak{v}(F^*)$ to F^*, and for all $x \in \mathfrak{v}(F)$, $\mathfrak{v}(i(x)) = x$.

Any discrete valued field can be given cross section, once we choose a prime element π, by defining the embedding $i(n) = \pi^n$ for all $n \in \mathbb{Z}$, and $i(\infty) = 0$. For all $n \in \mathbb{Z}$, we have $\mathfrak{v}(i(n)) = \mathfrak{v}(\pi^n) = n$, and $\mathfrak{v}(i(\infty)) = \mathfrak{v}(0) = \infty$. For the remainder of this thesis, we will identify the value group of all discrete valuation fields with the multiplicative group $\{\pi^n \mid n \in \mathbb{Z}\}$ and suppress the embedding i.

Example 3.1.5. In order to write down logical statements about valued fields (with cross section), we will need a number of symbols. Let $\mathcal{L}_{VF}$ be the language $\{+, \cdot, -, 0, 1, V, \leq, \mathfrak{v}\}$, where $+$ and $\cdot$ are binary function symbols, $-$ is a unary function symbol, 0 and 1 are constant symbols, V is a unary relation symbol, $\leq$ is a binary relation symbol, and $\mathfrak{v}$ is a unary function symbol.

The cross section property will be useful so that we can refer to elements of the value group within the domain of the valued field structure. When interpreting the symbols of $\mathcal{L}_{VF}$ in a structure which is a valued field, we will use $+, \cdot, -, 0, 1$ to represent the field operations, the additive inverse function, and the additive and multiplicative identities, V to pick out the elements of the value group (that is, $x \in V$ if and only if x is in the value group), $\leq$ to represent the ordering on the value group, and $\mathfrak{v}$ to represent the valuation.

Note again that these function and relation symbols may be interpreted as any functions and relations of the appropriate arities on any set. In order to require that our $\mathcal{L}_{VF}$ structures be valued fields, we will need some logical statements, the valued field axioms.

Terms, Formulas, and Satisfaction. The building blocks of our logical statements are the symbols of a language $\mathcal{L}$, an infinite set of variables $\mathcal{V} = \{v_1, v_2, \ldots\}$, and the formal symbols $=$, $\wedge$, $\vee$, $\neg$, $\exists$, $\forall$, (, and). The symbols $\wedge$, $\vee$, and $\neg$ (read as "and", "or", and "not") are called Boolean operators, and the symbols $\forall$ and $\exists$ (read as "for all" and "there exists") are called quantifiers. Certain finite strings of these symbols, called $\mathcal{L}$-terms, $\mathcal{L}$-formulas, and $\mathcal{L}$-sentences, can be interpreted to have semantic meaning. Intuitively, given an $\mathcal{L}$-structure, we use $\mathcal{L}$-terms to refer to elements of that structure, $\mathcal{L}$-formulas to express properties of particular elements of the structure, and $\mathcal{L}$-sentences to express properties of the structure itself.

In order to analyze these strings systematically, we define them with a specific inductive structure. The simplest are $\mathcal{L}$-terms, which are constructed from constants and variables by means of function applications.

Definition 3.1.6. A finite string t is an *$\mathcal{L}$-term* if and only if

- it is a constant symbol, $t \in \mathcal{C}$, or
- it is a variable, $t \in \mathcal{V}$, or
- it has the form $f(t_1, \ldots, t_{n_f})$, where $f \in \mathcal{F}$ is a function symbol, and $t_1, \ldots, t_{n_f}$ are $\mathcal{L}$-terms.

Binary function symbols, such as $+$ or $\cdot$, will often be written in the usual (infix) way. That is, when constructing $\mathcal{L}$-terms, we will write $t_1 + t_2$ instead of $+(t_1, t_2)$ and $t_1 \cdot t_2$ instead of $\cdot(t_1, t_2)$.

If an $\mathcal{L}$-term t contains variables from $v_1, \ldots, v_n$, we will often write it as $t(v_1, \ldots, v_n)$. We do not require all of the variables $v_1, \ldots, v_n$ to appear in t. Let $a_1, \ldots, a_n$ be elements of the domain of some $\mathcal{L}$-structure $\mathcal{M}$. We will denote by $t^{\mathcal{M}}(a_1, \ldots, a_n)$ the interpretation of t in $\mathcal{M}$ on the elements $a_1, \ldots, a_n$. The interpretation is obtained by substituting for each v_i the corresponding element a_i, substituting for each constant symbol c its interpretation $c^{\mathcal{M}}$, and substituting for each function symbol f its interpretation $f^{\mathcal{M}}$. Applying functions in the natural way, $t^{\mathcal{M}}(a_1, \ldots, a_n)$ is an element of the domain of $\mathcal{M}$.

Some examples of $\mathcal{L}_{VF}$-terms include 0, $1 + 1$, $v_1 \cdot 1$, and $\mathfrak{v}(v_1 + v_2)$. If t is $\mathfrak{v}(v_1 + v_2)$, then taking $\mathbb{Q}_3$ as an $\mathcal{L}_{VF}$-structure (with elements written as "Laurent series" in 3), where we interpret $+$ as addition and $\mathfrak{v}$ as the valuation $\mathfrak{v}_3$, we have $t^{\mathbb{Q}_3}(1, 2 + 2 \cdot 3 + 3^2) = \mathfrak{v}_3(1 +^{\mathbb{Q}_3} (2 + 2 \cdot 3 + 3^2)) = \mathfrak{v}_3(2 \cdot 3^2) = 3^2$ (recall that the value group in $\mathbb{Q}_3$ as a valued field with cross section is $\{3^n \mid n \in \mathbb{Z}\}$).

Next, we define $\mathcal{L}$-formulas. The simplest of these, called atomic $\mathcal{L}$-formulas, express the properties that two terms are equal or that a collection of terms satisfy a relation. General $\mathcal{L}$-formulas are constructed from atomic $\mathcal{L}$-formulas by means of Boolean operators and quantifiers.

Definition 3.1.7. A finite string is an *$\mathcal{L}$-formula* if and only if

- it has the form $t_1 = t_2$, where t_1 and t_2 are $\mathcal{L}$-terms, or
- it has the form $R(t_1, \ldots, t_{n_R})$, where $R \in \mathcal{R}$ is a relation symbol, and $t_1, \ldots, t_{n_R}$ are $\mathcal{L}$-terms.
- it has the form $\neg\phi$, $\phi \wedge \psi$, $\phi \vee \psi$, $\exists v\, \phi$, or $\forall v\, \phi$, where ϕ and ψ are $\mathcal{L}$-formulas and $v \in \mathcal{V}$ is a variable.

Some binary relation symbols, such as $\leq$, will also be written in the usual (infix) way. Instead of the atomic formula $\leq(t_1, t_2)$, we will write $t_1 \leq t_2$.

We will use parentheses for grouping in the natural way to avoid ambiguity. We will omit the formalization of this, as it is straightforward but rather time consuming.

We will also employ the standard abbreviations $\phi \to \psi$ (read as "ϕ implies ψ" or "if ϕ then ψ") for $\neg\phi \vee \psi$ and $\phi \leftrightarrow \psi$ (read as "ϕ if and only if ψ") for $(\phi \to \psi) \wedge (\psi \to \phi)$. We could have omitted $\vee$ and $\forall$ from our definition of $\mathcal{L}$-formula, since $\phi \vee \psi$ and $\forall v\, \phi$ can be viewed as abbreviations for $\neg(\neg\phi \wedge \neg\psi)$ and $\neg(\exists v\, \neg\phi)$ respectively.

Some examples of $\mathcal{L}_{VF}$-formulas include $(v_1 + v_2) + v_3 = v_1 + (v_2 + v_3)$, $\mathfrak{v}(v_1) \leq \mathfrak{v}(v_1 + v_2)$, $\forall v_1\, V(\mathfrak{v}(v_1))$, and $\neg(v_1 = 0) \to (\exists v_2\, v_1 \cdot v_2 = 1)$.

Upon interpreting an $\mathcal{L}$-formula in a particular $\mathcal{L}$-structure, $\mathcal{M}$, the quantifiers $\forall$ and $\exists$ are understood to quantify over the elements of M, the domain of $\mathcal{M}$. This is what makes the formula "first-order". In first-order logic, we cannot express statements like "every bounded subset has a least upper bound" or "$\forall i \in \mathbb{Z}\,, x^i \neq 0$", since the first quantifies over subsets, not elements, and the second quantifies over a specific structure, the integers. Because of this restriction, first-order logic is less expressive than other logics, but it has more structure which can be exploited mathematically.

A variable v is called bound if it occurs inside a $\forall v$ or $\exists v$ quantifier. Otherwise it is called free. In the $\mathcal{L}_{VF}$-formula $\exists v_1\, v_1 \leq v_2$, v_1 is bound, but v_2 is free. To avoid ambiguity, we will require that no variable occurs both free and bound in a formula, and that no variable is bound by more than one quantifier. When combining formulas, we can ensure these conditions by substituting unused variables for any variable which appears in more than one context.

If an $\mathcal{L}$-formula ϕ contains free variables from $v_1, \ldots, v_n$, we often write it as $\phi(v_1, \ldots, v_n)$. We do not require all of the variables $v_1, \ldots, v_n$ to appear in ϕ. If we substitute elements $a_1, \ldots, a_n$ from the domain of an $\mathcal{L}$-structure $\mathcal{M}$ for the variables $v_1, \ldots, v_n$, and if we interpret the $\mathcal{L}$-symbols in $\mathcal{M}$ and interpret the boolean operators and quantifiers in the natural way, then $\phi(a_1, \ldots, a_n)$ is either true or false in $\mathcal{M}$. If it is true, we write $\mathcal{M} \models \phi(a_1, \ldots, a_n)$ and say that $\mathcal{M}$ satisfies $\phi(a_1, \ldots, a_n)$. Otherwise, we write $\mathcal{M} \not\models \phi(a_1, \ldots, a_n)$.

For example, let $\phi_{\mathcal{O}}$ be the formula $1 \leq \mathfrak{v}(v_1)$. Let F be a discrete valued field with valuation $\mathfrak{v}$ and prime element π, taken as an $\mathcal{L}_{VF}$-structure in the natural way. For all $x \in F$, $F \models \phi_{\mathcal{O}}(x)$ if and only if $\mathfrak{v}(x) \geq \pi^0$, that is, if and only if x is an element of $\mathcal{O}_F$. In this way, the formula expresses a property of elements of a valued field, namely, that an element is in the valuation ring.

Definition 3.1.8. Let ϕ be an $\mathcal{L}$-formula with free variables from $v_1, \ldots, v_n$. Let $\mathcal{M}$ be an $\mathcal{L}$-structure with domain M, and let $a_1, \ldots, a_n \in M^n$ be elements of the domain. We define $\mathcal{M} \models \phi(a_1, \ldots, a_n)$ inductively as follows:

- If ϕ is $t_1(v_1, \ldots, v_n) = t_2(v_1, \ldots, v_n)$, then $\mathcal{M} \models \phi(a_1, \ldots, a_n)$ if and only if
$$t_1^{\mathcal{M}}(a_1, \ldots, a_n) = t_1^{\mathcal{M}}(a_1, \ldots, a_n).$$
- If ϕ is $R(t_1(v_1, \ldots, v_n), \ldots, t_{n_r}(v_1, \ldots, v_n))$, then $\mathcal{M} \models \phi(a_1, \ldots, a_n)$ if and only if
$$(t_1^{\mathcal{M}}(a_1, \ldots, a_n), \ldots, t_{n_R}^{\mathcal{M}}(a_1, \ldots, a_n)) \in R^{\mathcal{M}}.$$

- If ϕ is $\neg\psi(v_1, \ldots, v_n)$, then $\mathcal{M} \models \phi(a_1, \ldots, a_n)$ if and only if
$$\mathcal{M} \not\models \psi(a_1, \ldots, a_n).$$
- If ϕ is $\psi(v_1, \ldots, v_n) \wedge \theta(v_1, \ldots, v_n)$, then $\mathcal{M} \models \phi(a_1, \ldots, a_n)$ if and only if
$$\mathcal{M} \models \psi(a_1, \ldots, a_n) \text{ and } \mathcal{M} \models \theta(a_1, \ldots, a_n).$$
- If ϕ is $\psi(v_1, \ldots, v_n) \vee \theta(v_1, \ldots, v_n)$, then $\mathcal{M} \models \phi(a_1, \ldots, a_n)$ if and only if
$$\mathcal{M} \models \psi(a_1, \ldots, a_n) \text{ or } \mathcal{M} \models \theta(a_1, \ldots, a_n).$$
- If ϕ is $\exists v\, \psi(v_1, \ldots, v_n, v)$, then $\mathcal{M} \models \phi(a_1, \ldots, a_n)$ if and only if there exists $b \in M$ such that
$$\mathcal{M} \models \psi(a_1, \ldots, a_n, b).$$
- If ϕ is $\forall v\, \psi(v_1, \ldots, v_n, v)$, then $\mathcal{M} \models \phi(a_1, \ldots, a_n)$ if and only if for all $b \in M$,
$$\mathcal{M} \models \psi(a_1, \ldots, a_n, b).$$

This definition may seem pedantic, but it is another key separation between syntax and semantics, and it clearly demonstrates the inductive structure of $\mathcal{L}$-formulas.

Sentences and Theories. For each prime p, consider the $\mathcal{L}_{VF}$-formula
$$Char_p : \underbrace{1 + 1 + \ldots + 1}_{p\,times} = 0,$$
which expresses the property that an $\mathcal{L}_{VF}$-structure has characteristic p. Once again taking valued fields as $\mathcal{L}_{VF}$-structures in the natural way, we have $\mathbb{F}_5((t)) \models Char_5$, but $\mathbb{Q}_5 \not\models Char_5$.

Definition 3.1.9. An *$\mathcal{L}$-sentence* is an $\mathcal{L}$-formula which has no free variables.

Note that since $Char_p$ has no free variables, we are able to state whether a structure $\mathcal{M}$ satisfies $Char_p$ without choosing any elements from the domain of $\mathcal{M}$ to substitute. Sentences express properties of structures, not of individual elements.

Suppose that we want to express the property that a structure has characteristic zero as an $\mathcal{L}_{VF}$-sentence. That is, we want to say that a structure does not have characteristic p for any prime p. We can write a sentence which expresses the property that a structure does not have characteristic p for some finite number of primes $p_1, \ldots, p_n : \neg Char_{p_1} \wedge \neg Char_{p_2} \wedge \ldots \wedge \neg Char_{p_n}$. But sentences have finite length by definition, so this approach will not work for all p. It turns out that the only way to express the property characteristic zero in $\mathcal{L}_{VF}$ is with an infinite set of sentences.

Definition 3.1.10. An *$\mathcal{L}$-theory* is a set of $\mathcal{L}$-sentences. For $\mathcal{M}$ an $\mathcal{L}$-structure, and T an $\mathcal{L}$-theory, we say that $\mathcal{M}$ is a *model* of T, written $\mathcal{M} \models T$, if $\mathcal{M} \models \phi$ for all sentences $\phi \in T$.

In order to express the property that a structure has characteristic zero, we can define an $\mathcal{L}_{VF}$-theory, $Char_0 = \{\neg Char_p \,|\, p \,\text{prime}\}$. Then $\mathbb{Q}_5 \models Char_0$, since $\mathbb{Q}_5 \models \neg Char_p$ for all primesm p, but $\mathbb{F}_5((t)) \not\models Char_0$, since $\mathbb{F}_5((t)) \not\models \neg Char_5$.

Of course, every valued field with characteristic zero is a model for the theory $Char_0$, but $Char_0$ has other models which are not even fields. If we add the field axioms expressed as

$\mathcal{L}_{VF}$ sentences to the theory, then the models for this theory will be exactly the class of $\mathcal{L}_{VF}$-structures which are fields with characteristic zero.

Definition 3.1.11. A class of $\mathcal{L}$-structures, $\mathcal{K}$, is called *elementary* if there exists an $\mathcal{L}$-theory T such that $\mathcal{K}$ contains exactly those $\mathcal{L}$-structures which are models for T. The structure T is called a set of *axioms* for $\mathcal{K}$.

Example 3.1.12. We will show that the class of valued fields with cross section is elementary by providing a set of axioms in $\mathcal{L}_{VF}$. We will call this theory VF. As an exercise, make sure you understand what property each of the following $\mathcal{L}_{VF}$-sentences expresses.

- Field axioms:
  (1) $\forall v_1 \forall v_2 \forall v_3 \, (v_1 + v_2) + v_3 = v_1 + (v_2 + v_3)$
  (2) $\forall v_1 \, v_1 + 0 = v_1$
  (3) $\forall v_1 \, v_1 + -(v_1) = 0$
  (4) $\forall v_1 \forall v_2 \, v_1 + v_2 = v_2 + v_1$
  (5) $\forall v_1 \forall v_2 \forall v_3 \, (v_1 \cdot v_2) \cdot v_3 = v_1 \cdot (v_2 \cdot v_3)$
  (6) $\forall v_1 \, v_1 \cdot 1 = v_1$
  (7) $\forall v_1 \, \neg(v_1 = 0) \to (\exists v_2 \, v_1 \cdot v_2 = 1)$
  (8) $\forall v_1 \forall v_2 \, v_1 \cdot v_2 = v_2 \cdot v_1$
  (9) $\forall v_1 \forall v_2 \forall v_3 \, v_1 \cdot (v_2 + v_3) = v_1 \cdot v_2 + v_1 \cdot v_3$
  (10) $\neg(1 = 0)$
- Valuation axioms:
  (1) $\forall v_1 \, \mathfrak{v}(v_1) = 0 \leftrightarrow v_1 = 0$
  (2) $\forall v_1 \forall v_2 \, \mathfrak{v}(v_1 \cdot v_2) = \mathfrak{v}(v_1) \cdot \mathfrak{v}(v_2)$
  (3) $\forall v_1 \forall v_2 \, \mathfrak{v}(v_1) \leq \mathfrak{v}(v_2) \to \mathfrak{v}(v_1) \leq \mathfrak{v}(v_1 + v_2)$
- The value group is a subgroup of the multiplicative group:
  (1) $\forall v_1 \forall v_2 \, (V(v_1) \wedge V(v_2)) \to V(v_1 \cdot v_2)$
  (2) $\forall v_1 \, V(v_1) \to (\exists v_2 \, V(v_2) \wedge v_1 \cdot v_2 = 1)$
- Linear order axioms for the value group:
  (1) $\forall v_1 \forall v_2 \, (V(v_1) \wedge V(v_2)) \to (v_1 \leq v_2 \vee v_2 \leq v_1)$
  (2) $\forall v_1 \forall v_2 \, (V(v_1) \wedge V(v_2) \wedge (v_1 \leq v_2) \wedge (v_2 \leq v_1)) \to (v_1 = v_2)$
  (3) $\forall v_1 \forall v_2 \forall v_3 \, (V(v_1) \wedge V(v_2) \wedge V(v_3) \wedge (v_1 \leq v_2) \wedge (v_2 \leq v_3)) \to (v_1 \leq v_3)$
  (4) $\forall v_1 \forall v_2 \forall v_3 \, (V(v_1) \wedge V(v_2) \wedge V(v_3) \wedge v_1 \leq v_2) \to (v_1 \cdot v_3 \leq v_1 \cdot v_3)$
- Cross section axioms:
  (1) $\forall v_1 \, V(\mathfrak{v}(v_1))$
  (2) $\forall v_1 \, V(v_1) \to (\mathfrak{v}(v_1) = v_1)$

Of course, there are many other $\mathcal{L}_{VF}$-sentences which are true of all valued fields but are not included in the axioms. We call these sentences logical consequences of the theory, and denote this relationship with the already overloaded symbol $\models$.

Definition 3.1.13. Let T be an $\mathcal{L}$-theory and ϕ an $\mathcal{L}$-sentence. We call ϕ a *logical consequence* of T and write $T \models \phi$ if $\mathcal{M} \models \phi$ for all models $\mathcal{M} \models T$. An $\mathcal{L}$-theory T is called *complete* if for all $\mathcal{L}$-sentences ϕ, $T \models \phi$ or $T \models \neg\phi$.

Note that in the definition of $T \models \phi$, we do not claim that one can provide a proof of ϕ given the assumptions in T, merely that in any structure in which the sentences in T hold,

ϕ also holds. One of Gödel's remarkable results was proving that these concepts, provability and model theoretic consequence, are actually equivalent.

The mathematical study of proof and proof systems belongs to another branch of logic, and we will not make the notion of proof completely rigorous in this thesis. But we will note that in this context, a proof means a finite sequence of $\mathcal{L}$-sentences, some of which are introduced as assumptions and some of which follow from previous sentences by rules of inference. If a proof of an $\mathcal{L}$-sentence exists involving only assumptions from an $\mathcal{L}$-theory T, we write $T \vdash \phi$.

Theorem 3.1.14 (Gödel's Completeness Theorem)**.** *Let T be an $\mathcal{L}$-theory and ϕ an $\mathcal{L}$-sentence. Then $T \models \phi$ if and only if $T \vdash \phi$.*

One consequence of the Completeness Theorem is that any theory which does not imply a contradiction has a model.

Definition 3.1.15. An $\mathcal{L}$-theory T is called *inconsistent* if there is an $\mathcal{L}$-sentence ϕ such that $T \vdash (\phi \wedge \neg\phi)$. Otherwise, T is called *consistent*.

Definition 3.1.16. An $\mathcal{L}$-theory T is called *satisfiable* if it has a model.

Corollary 3.1.17 ([Mar02, Corollary 2.1.3])**.** *An $\mathcal{L}$-theory T is satisfiable if and only if is consistent.*

Proof. Suppose T is satisfiable. Then there is a model $\mathcal{M} \models T$. If $T \models (\phi \wedge \neg\phi)$ for some $\mathcal{L}$-sentence ϕ, then $\mathcal{M} \models (\phi \wedge \neg\phi)$, which is impossible, since by definition, we would have $\mathcal{M} \models \phi$ and $\mathcal{M} \not\models \phi$, a contradiction. Thus T is consistent.

Now suppose T is not satisfiable. For any $\mathcal{L}$-sentence ϕ, $(\phi \wedge \neg\phi)$ is true in every model of T trivially, since T has no models. Thus $T \models (\phi \wedge \neg\phi)$, and by the Completeness Theorem, $T \vdash (\phi \wedge \neg\phi)$. So T is inconsistent. □

Another easy consequence of the Completeness Theorem is the Compactness Theorem, a powerful result which is central to Model Theory.

Theorem 3.1.18 (Compactness Theorem [Mar02, Theorem 2.1.4])**.** *An $\mathcal{L}$-theory T is satisfiable if and only if every finite subset of T is satisfiable.*

Proof. One direction is obvious. Any model of T is also a model of any subset of T, so if T is satisfiable, then every finite subset of T is satisfiable.

Conversely, suppose that every finite subset of T is satisfiable. Assume for the sake of contradiction that T is not satisfiable. Then by Corollary 3.1.17, there is some $\mathcal{L}$-sentence ϕ such that $T \vdash (\phi \wedge \neg\phi)$. Now since proofs are finite in length, the proof of $\phi \wedge \neg\phi$ can use as assumptions only finitely many elements of T. Call this finite set Δ.

Then $\Delta \vdash (\phi \wedge \neg\phi)$, and by Corollary 3.1.17, Δ is not satisfiable. But this is a contradiction, and hence T is satisfiable. □

There are several other proofs of the Completeness Theorem which do not rely on the Completeness Theorem. These other methods are in a sense more constructive, and they can give us more information about the satisfying model, including us an upper bound on its cardinality. See Marker [Mar02].

Theorem 3.1.19 (Compactness Theorem with cardinality [Mar02, Theorem 2.1.11])**.** *Let T be an $\mathcal{L}$-theory such that every finite subset of T is satisfiable. Then there is a model of T of cardinality $|\mathcal{L}|$.*

Example 3.1.20. Compactness is a powerful tool for constructing models with desired properties. As an example, consider the language $\mathcal{L} = \{\cdot, +, <, 0, 1\}$ and the $\mathcal{L}$-structure $\mathbb{N}$, with the $\mathcal{L}$-symbols interpreted in the usual way. Let $Th(\mathbb{N})$ be the full $\mathcal{L}$-theory of $\mathbb{N}$, that is, the set of all $\mathcal{L}$-sentences which are true in $\mathbb{N}$.

Now we will extend the language by adding a new constant symbol, c. Let $\mathcal{L}' = \mathcal{L} \cup \{c\}$. We will also extend the theory by adding new sentences expressing that c is larger than every natural number. For each $n \in \mathbb{N}$, let ϕ_n be the $\mathcal{L}'$-sentence

$$\underbrace{1 + 1 + \ldots + 1}_{n \text{ times}} < c.$$

Let $T' = Th(\mathbb{N}) \cup \{\phi_n \mid n = 1, 2, \ldots\}$. We can consider the $\mathcal{L}$-sentences in $Th(\mathbb{N})$ as $\mathcal{L}'$-sentences, since $\mathcal{L} \subset \mathcal{L}'$, so T' is an $\mathcal{L}'$-theory.

We will use Compactness to show that T' has a model. Let Δ be a finite subset of T'. We claim that $\mathbb{N} \models \Delta$ under an appropriate interpretation of c. Since Δ is finite, it consists of finitely many sentences of $Th(\mathbb{N})$ and finitely many ϕ_n. Let M be the greatest integer such that $\phi_M \in \Delta$ (or 0 if Δ contains no ϕ_n). Then consider $\mathbb{N}$ as an $\mathcal{L}'$-structure, with the interpretation $c^{\mathbb{N}} = M + 1$. For each $\phi_n \in T'$, $\mathbb{N} \models \phi_n$, since $n < M + 1$. For each other $\psi \in T'$, $\psi \in Th(\mathbb{N})$, so $\mathbb{N} \models \psi$ by definition.

Thus with this interpretation of c, $\mathbb{N} \models \Delta$. Hence every finite subset of T' is satisfiable, and by Compactness, T' has a model.

This means that there is an $\mathcal{L}'$-structure, $\mathcal{N}$, such that every $\mathcal{L}$-sentence which is true in $\mathbb{N}$ is true in $\mathcal{N}$. That is, $\mathbb{N}$ and $\mathcal{N}$ cannot be distinguished using any first-order $\mathcal{L}$-sentence. But the interpretation of c in $\mathcal{N}$ is greater than every natural number, so $\mathcal{N}$ has "infinite" elements.

Model theory is filled with counterintuitive results like these, and much of the theory is devoted to exploring the properties of unusual models for familiar theories. In fact, part of the proof of the Ax-Kochen Theorem requires the use of very large models for the theory of valued fields (see Section 3.3).

Homomorphisms and Elementary Maps. As usual when defining new mathematical objects, we will define the maps between them which preserve structure.

Definition 3.1.21. Given two $\mathcal{L}$-structures $\mathcal{M}$ and $\mathcal{N}$ with domains M and N, an $\mathcal{L}$-*homomorphism* from $\mathcal{M}$ to $\mathcal{N}$ is a map $\eta : M \to N$ which preserves interpretation of $\mathcal{L}$-symbols. That is,

- if c is a constant symbol, then $\eta(c^{\mathcal{M}}) = c^{\mathcal{N}}$,
- if f is a function symbol, then for all $(a_1, \ldots, a_{n_f}) \in M^{n_f}$, $\eta(f^{\mathcal{M}}(a_1, \ldots, a_{n_f})) = f^{\mathcal{N}}(\eta(a_1), \ldots, \eta(a_{n_f}))$, and
- if R is a relation symbol, then for all $(a_1, \ldots, a_{n_f}) \in M^{n_R}$, $(a_1, \ldots, a_{n_R}) \in R^{\mathcal{M}}$ if and only if $(\eta(a_1), \ldots, \eta(a_{n_R})) \in R^{\mathcal{N}}$.

An $\mathcal{L}$-*isomorphism* is a bijective $\mathcal{L}$-homomorphism. If there is an $\mathcal{L}$-isomorphism from $\mathcal{M}$ to $\mathcal{N}$, we write $\mathcal{M} \cong \mathcal{N}$.

This definition of homomorphism generalizes the notion of homomorphism in many settings. For instance, if we interpret groups as $\mathcal{L}_G$-structures as in Example 3.1.3, then all group homomorphisms are $\mathcal{L}_G$-homomorphisms. Similarly, homomorphisms of valued fields, which must preserve the field structure, the valuation, and the ordering on the value group, are $\mathcal{L}_{VF}$-homomorphisms.

The existence of $\mathcal{L}$-homomorphisms between $\mathcal{L}$-structures does not give us much information about which $\mathcal{L}$-formulas are satisfied in these structures. A stronger notion is that of an elementary homomorphism, a map which preserves not just the interpretation of the language, but also the satisfaction of formulas.

Definition 3.1.22. An *elementary* $\mathcal{L}$-homomorphism is an $\mathcal{L}$-homomorphism $j : \mathcal{M} \to \mathcal{N}$ between $\mathcal{L}$-structures $\mathcal{M}$ and $\mathcal{N}$ such that for all $\mathcal{L}$-formulas $\phi(v_1, \ldots, v_n)$ and elements $a_1, \ldots, a_n$ in the domain of $\mathcal{M}$, $\mathcal{M} \models \phi(a_1, \ldots, a_n)$ if and only if $\mathcal{N} \models \phi(j(a_1), \ldots, j(a_n))$.

Definition 3.1.23. An $\mathcal{L}$-structure $\mathcal{M}$ with domain M is a *substructure* of an $\mathcal{L}$-structure $\mathcal{N}$ with domain N if $M \subseteq N$ and the inclusion map is an $\mathcal{L}$-homomorphism.

If $\mathcal{M}$ is a substructure of $\mathcal{N}$ and the inclusion map is elementary, then $\mathcal{M}$ is an *elementary substructure* of $\mathcal{N}$ and $\mathcal{N}$ is an *elementary extension* of $\mathcal{M}$.

One of the most surprising of the foundational theorems of model theory relates to the existence of elementary extensions and substructures. The Löwenheim-Skolem Theorem, given here without proof, intuitively states that given an infinite structure, there are elementary extensions and elementary substructures of all infinite cardinalities.

Theorem 3.1.24 (Löwenheim-Skolem Theorem Up [Mar02, Theorem 2.3.4])**.** *Let $\mathcal{M}$ be an infinite $\mathcal{L}$-structure with domain M, and let κ be an infinite cardinal such that $\kappa \geq |M| + |\mathcal{L}|$. Then there is an $\mathcal{L}$-structure $\mathcal{N}$ of cardinality κ and an elementary embedding $j : \mathcal{M} \to \mathcal{N}$, so that $\mathcal{N}$ is an elementary extension of $j(\mathcal{M})$.*

Theorem 3.1.25 (Löwenheim-Skolem Theorem Down [Mar02, Theorem 2.3.7])**.** *Let $\mathcal{M}$ be an $\mathcal{L}$-structure with domain M, and let $X \subseteq M$. Then there is an elementary substructure $\mathcal{N}$ of $\mathcal{M}$ with domain N such that $X \subseteq N$ and $|N| \leq |X| + |\mathcal{L}| + \aleph_0$.*

Remark 3.1.26. It is immediate from the definition of elementary $\mathcal{L}$-homomorphism that if there is an elementary $\mathcal{L}$-homomorphism $j : \mathcal{M} \to \mathcal{N}$, then $\mathcal{M}$ and $\mathcal{N}$ satisfy exactly the same $\mathcal{L}$-sentences, since for any $\mathcal{L}$-sentence ϕ, $\mathcal{M} \models \phi$ if and only if $\mathcal{N} \models \phi$. Such structures are called elementarily equivalent.

Given an $\mathcal{L}$-structure $\mathcal{M}$, we define the full $\mathcal{L}$-theory of $\mathcal{M}$, $Th(\mathcal{M}) = \{\phi \,|\, \mathcal{M} \models \phi\}$. By definition, if ϕ is an $\mathcal{L}$-sentence, $\mathcal{M} \models \phi$ or $\mathcal{M} \models \neg\phi$, so $Th(\mathcal{M}) \models \phi$ or $Th(\mathcal{M}) \models \neg\phi$, and thus $Th(\mathcal{M})$ is complete.

Definition 3.1.27. Let $\mathcal{M}$ and $\mathcal{N}$ be $\mathcal{L}$-structures. We say that $\mathcal{M}$ and $\mathcal{N}$ are *elementarily equivalent*, written $\mathcal{M} \equiv \mathcal{N}$, if $Th(\mathcal{M}) = Th(\mathcal{N})$, that is, for any $\mathcal{L}$-sentence ϕ, $\mathcal{M} \models \phi$ if and only if $\mathcal{N} \models \phi$.

The converse to Remark 3.1.26 does not hold in general. That is, it is possible to have $\mathcal{L}$-structures $\mathcal{M}$ and $\mathcal{N}$ such that $\mathcal{M} \equiv \mathcal{N}$, but there is no elementary $\mathcal{L}$-homomorphism from $\mathcal{M}$ to $\mathcal{N}$. The statement that $\mathcal{M}$ and $\mathcal{N}$ are elementarily equivalent only requires that they satisfy the same $\mathcal{L}$-*sentences*, but if there is an elementary $\mathcal{L}$-homomorphism from $\mathcal{M}$ to $\mathcal{N}$, then this homomorphism must respect the satisfaction of all $\mathcal{L}$-*formulas*. By choosing a suitably extended language, we can turn these formulas into sentences.

Let $\mathcal{M}$ be an $\mathcal{L}$-structure with domain M. We will extend the language $\mathcal{L}$ by adding a new constant symbol c_m for every element $m \in M$. Let $\mathcal{L}_{\mathcal{M}} = \mathcal{L} \cup \{c_m \,|\, m \in M\}$. Then $\mathcal{M}$ can be viewed as an $\mathcal{L}_{\mathcal{M}}$-structure, where we interpret each constant in the natural way, $c_m^{\mathcal{M}} = m$.

Definition 3.1.28. The *elementary diagram* of an $\mathcal{L}$-structure $\mathcal{M}$ with domain M, denoted $Diag_{el}(\mathcal{M})$, is the $\mathcal{L}_{\mathcal{M}}$-theory which captures all the information about satisfaction of $\mathcal{L}$-formulas in $\mathcal{M}$.

$$Diag_{el}(\mathcal{M}) = \{\phi(c_{m_1}, \ldots, c_{m_n}) \,|\, m_1, \ldots, m_n \in M \text{ and } \mathcal{M} \models \phi(m_1, \ldots, m_n)\}.$$

Lemma 3.1.29. *Let $\mathcal{M}$ be an $\mathcal{L}$-structure with domain M. Suppose $\mathcal{N}$ is an $\mathcal{L}_{\mathcal{M}}$-structure such that $\mathcal{N} \models Diag_{el}(\mathcal{M})$. Then viewing $\mathcal{N}$ as an $\mathcal{L}$-structure (by "forgetting" the interpretations of the symbols c_m), there is an elementary embedding of $\mathcal{M}$ into $\mathcal{N}$.*

Proof. Define $j : \mathcal{M} \to \mathcal{N}$ by $j(m) = c_m^{\mathcal{N}}$, the interpretation of the corresponding constant symbol in $\mathcal{N}$. Suppose $m_1, m_2 \in M$ with $m_1 \neq m_2$. Then $\neg(c_{m_1} = c_{m_2})$ is in $Diag_{el}(\mathcal{M})$, so $\mathcal{N} \models \neg(c_{m_1} = c_{m_2})$, and in $\mathcal{N}$, $c_{m_1}^{\mathcal{N}} \neq c_{m_2}^{\mathcal{N}}$, so $j(m_1) \neq j(m_2)$. Thus j is injective.

If $\mathcal{M} \models \phi(m_1, \ldots, m_n)$ for an $\mathcal{L}$-formula ϕ, then $\phi(c_{m_1}, \ldots, c_{m_n}) \in Diag_{el}(\mathcal{M})$. Since $\mathcal{N} \models Diag_{el}(\mathcal{M})$, $\mathcal{N} \models \phi(c_{m_1}, \ldots, c_{m_n})$, and thus $\mathcal{N} \models \phi(j(m_1), \ldots, j(m_n))$, so j is an elementary $\mathcal{L}$-homomorphism. □

As one would expect, $\mathcal{L}$-isomorphic structures are elementarily equivalent. We will conclude our whirlwind tour of the basics of model theory with a proof of this fact, which will also serve as a first example of the technique of induction on terms and formulas.

The idea is that all terms and formulas are built from atomic elements (terms from variables and constants, formulas from atomic formulas) in a finite number of steps. To prove a claim, we show that it is true for all atomic elements. Then we show that if we construct a new term (or formula) from a set of terms (or formulas) for which our claim is true, then our claim is also true on the new term (or formula). This shows that the claim is true for all terms (or formulas).

Theorem 3.1.30 ([Mar02, Theorem 1.1.10])**.** *For $\mathcal{L}$-structures $\mathcal{M}$ and $\mathcal{N}$, if $\mathcal{M} \cong \mathcal{N}$, then $\mathcal{M} \equiv \mathcal{N}$.*

Proof. Let $j : \mathcal{M} \to \mathcal{N}$ be an $\mathcal{L}$-isomorphism mapping M, the domain of $\mathcal{M}$, bijectively to N, the domain of $\mathcal{N}$. If $\overline{a} = (a_1, \ldots, a_n) \in M^n$, let $j(\overline{a}) = (j(a_1), \ldots, j(a_n)) \in N^n$.

We will prove by induction on terms the following claim: if t is an $\mathcal{L}$-term with free variables from $\overline{v} = (v_1, \ldots, v_n)$, then for all $\overline{a} = (a_1, \ldots, a_n) \in M^n$, $j(t^{\mathcal{M}}(\overline{a})) = t^{\mathcal{N}}(j(\overline{a}))$.

If t is a constant symbol c, then $j(t^{\mathcal{M}}(\overline{a})) = j(c^{\mathcal{M}}) = c^{\mathcal{N}} = t^{\mathcal{N}}(j(\overline{a}))$.

If t is a variable v_i, then $j(t^{\mathcal{M}}(\overline{a})) = j(a_i) = t^{\mathcal{N}}(j(\overline{a}))$.

If t is $f(t_1(\overline{v}), \ldots, t_{n_f}(\overline{v}))$, where f is a function symbol and $t_1, \ldots, t_{n_f}$ are $\mathcal{L}$-terms for which the claim is true, then

$$\begin{aligned} j(t^{\mathcal{M}}(\overline{a})) &= j(f^{\mathcal{M}}(t_1^{\mathcal{M}}(\overline{a}), \ldots, t_{n_f}^{\mathcal{M}}(\overline{a}))) \\ &= f^{\mathcal{N}}(j(t_1^{\mathcal{M}}(\overline{a})), \ldots, j(t_{n_f}^{\mathcal{M}}(\overline{a}))) \text{ since } j \text{ is an } \mathcal{L}\text{-homomorphism} \\ &= f^{\mathcal{N}}(t_1^{\mathcal{N}}(j(\overline{a})), \ldots, t_{n_f}^{\mathcal{N}}(j(\overline{a}))) \text{ by induction} \\ &= t^{\mathcal{N}}(j(\overline{a})). \end{aligned}$$

This completes the induction on terms and the proof of the claim.

Next we will prove by induction on formulas that if ϕ is an $\mathcal{L}$-formula with free variables from $\overline{v} = (v_1, \ldots, v_n)$, then for all $\overline{a} = (a_1, \ldots, a_n) \in M^n$, $\mathcal{M} \models \phi(\overline{a})$ if and only if $\mathcal{N} \models \phi(j(\overline{a}))$.

If $\phi(\overline{v})$ is $t_1(\overline{v}) = t_2(\overline{v})$, where t_1 and t_2 are $\mathcal{L}$-terms, then

$$\begin{aligned} \mathcal{M} \models \phi(\overline{a}) \text{ iff } & t_1^{\mathcal{M}}(\overline{a}) = t_2^{\mathcal{M}}(\overline{a}) \\ \text{iff } & j(t_1^{\mathcal{M}}(\overline{a})) = j(t_2^{\mathcal{M}}(\overline{a})) \text{ because } j \text{ is injective} \\ \text{iff } & t_1^{\mathcal{N}}(j(\overline{a})) = t_2^{\mathcal{N}}(j(\overline{a})) \text{ applying the claim} \\ \text{iff } & \mathcal{N} \models \phi(j(\overline{a})). \end{aligned}$$

If $\phi(\overline{v})$ is $R(t_1(\overline{v}), \ldots, t_{n_R}(\overline{v}))$, where R is a relation symbol and $t_1, \ldots, t_{n_R}$ are $\mathcal{L}$-terms, then

$$\begin{aligned} \mathcal{M} \models \phi(\overline{a}) \text{ iff } & (t_1^{\mathcal{M}}(\overline{a}), \ldots, t_{n_R}^{\mathcal{M}}(\overline{a})) \in R^{\mathcal{M}} \\ \text{iff } & (j(t_1^{\mathcal{M}}(\overline{a})), \ldots, j(t_{n_R}^{\mathcal{M}}(\overline{a}))) \in R^{\mathcal{N}} \text{ because } j \text{ is a homomorphism} \\ \text{iff } & (t_1^{\mathcal{N}}(j(\overline{a})), \ldots, t_{n_R}^{\mathcal{N}}(j(\overline{a}))) \in R^{\mathcal{N}} \text{ applying the claim} \\ \text{iff } & \mathcal{N} \models \phi(j(\overline{a})). \end{aligned}$$

If $\phi(\overline{v})$ is $\neg\psi(\overline{v})$, where ψ is an $\mathcal{L}$-formula for which our assertion is true, then

$$\begin{aligned} \mathcal{M} \models \phi(\overline{a}) \text{ iff } & \mathcal{M} \not\models \psi(\overline{a}) \\ \text{iff } & \mathcal{N} \not\models \psi(j(\overline{a})) \text{ by induction} \\ \text{iff } & \mathcal{N} \models \phi(j(\overline{a})). \end{aligned}$$

If $\phi(\overline{v})$ is $\psi(\overline{v}) \wedge \theta(\overline{v})$, where ψ and θ are $\mathcal{L}$-formulas for which our assertion is true, then

$$\begin{aligned} \mathcal{M} \models \phi(\overline{a}) \text{ iff } & \mathcal{M} \models \psi(\overline{v}) \text{ and } \mathcal{M} \models \theta(\overline{v}) \\ \text{iff } & \mathcal{N} \models \psi(j(\overline{v})) \text{ and } \mathcal{N} \models \theta(j(\overline{v})) \text{ by induction} \\ \text{iff } & \mathcal{N} \models \phi(j(\overline{v})). \end{aligned}$$

If $\phi(\overline{v})$ is $\exists w\, \psi(\overline{v}, w)$, where w is a variable and ψ is an $\mathcal{L}$-formula for which our assertion is true, then $\mathcal{M} \models \phi(\overline{a})$ if and only if there exists some $b \in M$ such that $\mathcal{M} \models \psi(\overline{a}, b)$. Now if there exists such a b, then by induction, $\mathcal{N} \models \psi(j(\overline{a}), j(b))$, so there exists $c \in N$ (take $c = j(b)$) such that $\mathcal{N} \models \psi(j(\overline{a}), c)$, and thus $\mathcal{N} \models \phi(j(\overline{a}))$. Conversely, if $\mathcal{N} \models \phi(j(\overline{a}))$, then there exists $c \in N$ such that $\mathcal{N} \models \psi(j(\overline{a}), c)$. j is surjective, so there exists $b \in M$ such that $j(b) = c$, and by induction $\mathcal{M} \models \psi(\overline{a}, b)$.

This completes the proof by induction on formulas. We do not need to consider formulas constructed using $\vee$ or $\forall$, since these can be re-written to use only $\neg$, $\wedge$, and $\exists$. This also completes the proof of the theorem, since we have shown that if ϕ is an $\mathcal{L}$-sentence, then $\mathcal{M} \models \phi$ if and only if $\mathcal{N} \models \phi$. So $\mathcal{M} \equiv \mathcal{N}$. $\square$

3.2. **Ultraproducts.** The fields $\mathbb{Q}_p$ and $\mathbb{F}_p((t))$ are not elementarily equivalent for any p (for instance, the sentence $Char_p$ is true in $\mathbb{F}_p((t))$ but false in $\mathbb{Q}_p$). However, we will prove that the theory of $\mathcal{L}_{VF}$-sentences which are true in $\mathbb{Q}_p$ for all but finitely many p is the same as the theory of $\mathcal{L}_{VF}$-sentences which are true in $\mathbb{F}_p((t))$ for all but finitely many p.

A construction called the ultraproduct will allow us to build new structures from the $\mathbb{Q}_p$ and $\mathbb{F}_p((t))$ in which a sentence is true if and only if it is true in "almost all" of these fields. The precise meaning of "almost all" is described by the definition of a filter on a set.

Definition 3.2.1. Given a set I, a *filter* on I is a subset of the power set $\mathcal{D} \subseteq \mathcal{P}(I)$ such that

- $\emptyset \notin \mathcal{D}$ and $I \in \mathcal{D}$,
- if $A \in \mathcal{D}$ and $B \in \mathcal{D}$, then $A \cap B \in \mathcal{D}$, and
- if $A \in \mathcal{D}$ and $A \subseteq B \subseteq I$, then $B \in \mathcal{D}$.

Example 3.2.2. The following are examples of filters on a set I:

- $\mathcal{D}_T = \{I\}$. $\mathcal{D}_T$ is called the trivial filter.
- For $j \in I$, $\mathcal{D}_j = \{X \subseteq I \,|\, j \in X\}$. $\mathcal{D}_j$ is called the principal filter generated by j.
- $\mathcal{D}_F = \{X \subseteq I \,|\, I \setminus X \text{ is finite}\}$. $\mathcal{D}_F$ is called the Frechet filter. Note that $\mathcal{D}_F$ is a filter only when I is infinite, since otherwise $\emptyset \in \mathcal{D}_F$.

Some intuition for the notion of a filter can be built by thinking of the sets in the filter as those containing "almost all" elements of I. If two sets both contain almost all elements, their intersection should also contain almost all elements. If a set contains almost all elements, any superset should also contain almost all elements. Of course, for different filters, "almost all" has different meanings. A principal filter, for instance, gives great preference to its generating element.

Definition 3.2.3. A filter $\mathcal{D}$ on I is an *ultrafilter* if for all $X \subseteq I$, $X \in \mathcal{D}$ or $I \setminus X \in \mathcal{D}$.

All principal filters are ultrafilters. The next lemma and theorem demonstrate that non-principal ultrafilters can be obtained by extending the Frechet filter on an infinite set.

Lemma 3.2.4. *Given a filter $\mathcal{D}$ on a set I and a subset $X \subset I$ such that $X \notin \mathcal{D}$, we can extend $\mathcal{D}$ to a filter $\mathcal{D}_X$ on I such that $\mathcal{D} \subseteq \mathcal{D}_X$ and $I \setminus X \in \mathcal{D}_X$.*

Proof. Let $\mathcal{D}_X = \{Y \subseteq I \,|\, \text{there exists } Z \in \mathcal{D} \text{ such that } Z \setminus X \subseteq Y\}$.

$\mathcal{D}_X$ is a filter on I:

- If $\emptyset \in \mathcal{D}_X$, then there exists $Z \in \mathcal{D}$ such that $Z \setminus X = \emptyset$, that is, $Z \subseteq X$. But then $X \in \mathcal{D}$, which contradicts our assumption. So $\emptyset \notin \mathcal{D}_X$. Also, for any $Z \in \mathcal{D}$, $Z \setminus X \subseteq I$, so $I \in \mathcal{D}_X$.
- For $A, B \in \mathcal{D}_X$, there exist sets $Z_A, Z_B \in \mathcal{D}$ such that $Z_A \setminus X \subseteq A$ and $Z_B \setminus X \subseteq B$. Then $(Z_A \cap Z_B) \setminus X \subseteq (A \cap B)$, so $A \cap B \in \mathcal{D}_X$.

- For $A \in \mathcal{D}_X$ and $A \subseteq B \subseteq I$, there exists $Z_A \in \mathcal{D}$ such that $Z_A \setminus X \subseteq A \subseteq B$, so $B \in \mathcal{D}_X$.

For any $Y \in \mathcal{D}$, take $Z = Y$. $Y \setminus X \subseteq Y$, so $Y \in \mathcal{D}_X$. Thus $\mathcal{D} \subseteq \mathcal{D}_X$.
To show that $I \setminus X \in \mathcal{D}_X$, take $Z = I$. $I \setminus X \subseteq I \setminus X$, so $I \setminus X \in \mathcal{D}_X$. □

Theorem 3.2.5 ([CK73, Proposition 4.1.3])**.** *For any filter $\mathcal{D}$ on I, there exists an ultrafilter $\mathcal{U}$ on I with $\mathcal{D} \subseteq \mathcal{U}$.*

Proof. Let $\mathcal{F}$ be the set of all filters on I extending $\mathcal{D}$, ordered by the subset relation. If $(C_\alpha : \alpha < \beta)$ is a chain in $\mathcal{F}$, then $\mathcal{C} = \bigcup_{\alpha<\beta} \mathcal{C}_\alpha$ is a filter on I:

- We have $\emptyset \notin \mathcal{C}$ since $\emptyset \notin \mathcal{C}_\alpha$ for all α, and $I \in \mathcal{C}$ since $I \in \mathcal{C}_\alpha$ for all α.
- If $A, B \in \mathcal{C}$, then $A \in \mathcal{C}_\alpha$ and $B \in \mathcal{C}_\beta$ for some α and β. Then $A, B \in \mathcal{C}_{\max\{\alpha,\beta\}}$, and $A \cap B \in \mathcal{C}_{\max\{\alpha,\beta\}} \subseteq \mathcal{C}$.
- If $A \in \mathcal{C}$ and $A \subseteq B \subseteq I$, then $A \in \mathcal{C}_\alpha$ for some α, so $B \in \mathcal{C}_\alpha \subseteq \mathcal{C}$.

$\mathcal{C}$ extends $\mathcal{D}$, since it is the union of a set of filters extending $\mathcal{D}$, so $\mathcal{C} \in \mathcal{F}$. Moreover, $\mathcal{C}$ is an upper bound for $(\mathcal{C} : \alpha < \beta)$, since $\mathcal{C}_\alpha \subseteq \mathcal{C}$ for all α. Applying Zorn's lemma, $\mathcal{F}$ has maximal elements. Let $\mathcal{U}$ be a maximal element. We claim that $\mathcal{U}$ is an ultrafilter.

Let $X \subseteq I$ be a subset such that $X \notin \mathcal{U}$. By Lemma 3.2.4 we can find a filter $\mathcal{U}_X$ on I such that $\mathcal{U} \subseteq \mathcal{U}_X$ and $I \setminus X \in \mathcal{U}_X$. Since $\mathcal{U}_X$ extends $\mathcal{U}$, it also extends $\mathcal{D}$. But $\mathcal{U}$ is a maximal element among the filters on I extending $\mathcal{D}$, and thus $\mathcal{U}_X = \mathcal{U}$. We have shown that for any $X \subset I$, if $X \notin \mathcal{U}$, then $I \setminus X \in \mathcal{U}$, so $\mathcal{U}$ is an ultrafilter. □

Note that no ultrafilter $\mathcal{U}$ on I extending the Frechet filter is principal, since for any $j \in I$, $I \setminus \{j\} \in \mathcal{D}_F \subset \mathcal{U}$. In fact, all nonprincipal ultrafilters are extensions of the Frechet filter.

Lemma 3.2.6. *The intersection $\bigcap \mathcal{D}$ of all nonprincipal ultrafilters $\mathcal{D}$ on an infinite set I is the Frechet filter $\mathcal{D}_F$ on I.*

Proof. Let $\mathcal{D}$ be a nonprincipal ultrafilter. For all $j \in I$, there exists $X_j \in \mathcal{D}$ such that $j \notin X_j$ (otherwise $\mathcal{D}$ would be principal generated by j). Then $Y_j = I \setminus \{j\} \in \mathcal{D}$, since $X_j \subseteq Y_j$. For all $A \in \mathcal{D}_F$, we have $A = I \setminus \{a_i\}_{i=1}^n$ for some finite set $\{a_i\}_{i=1}^n \subset I$. Then $A = \bigcap_{i=1}^n Y_{a_i} \in \mathcal{D}$. Thus $\mathcal{D}_F \subseteq \mathcal{D}$ for all nonprincipal ultrafilters $\mathcal{D}$, and $\mathcal{D}_F \subseteq \bigcap \mathcal{D}$.

Conversely, suppose $A \in \bigcap \mathcal{D}$, $A \notin \mathcal{D}_F$. Then by Lemma 3.2.4 we can extend $\mathcal{D}_F$ to a filter $\mathcal{D}_A$ containing $I \setminus A$. By Theorem 3.2.5 we can extend $\mathcal{D}_A$ to an ultrafilter $\mathcal{U}$ containing $I \setminus A$. This filter is nonprincipal, since it extends $\mathcal{D}_F$, and $A \notin \mathcal{U}$. But then $A \notin \bigcap \mathcal{D}$, contradicting our assumption. Thus $\bigcap \mathcal{D} \subseteq \mathcal{D}_F$, and we have shown that $\bigcap \mathcal{D} = \mathcal{D}_F$. □

The Ultraproduct Construction. Now that we have ultrafilters at our disposal, we are ready to introduce the ultraproduct construction.

Let $\{\mathcal{M}_i\}_{i\in I}$ be a collection of $\mathcal{L}$-structures indexed by an infinite set I, and let $\mathcal{D}$ be an ultrafilter on I. We will view the Cartesian product $\prod M_i$ of the domains of the $\mathcal{M}_i$ as the set of choice functions $\{f : I \to \bigcup M_i \,|\, \forall i \in I, f(i) \in M_i\}$. We define a relation $\sim_\mathcal{D}$ on $\prod M_i$ by $f \sim_\mathcal{D} g$ if and only if $\{i \in I \,|\, f(i) = g(i)\} \in \mathcal{D}$.

Proposition 3.2.7. *The relation $\sim_\mathcal{D}$ is an equivalence relation.*

Proof. Let $f, g, h \in \prod M_i$.

- The set $\{i \in I \mid f(i) = f(i)\} = I \in \mathcal{D}$, so $f \sim_{\mathcal{D}} f$, and $\sim_{\mathcal{D}}$ is reflexive.
- If $f \sim_{\mathcal{D}} g$, then $\{i \in I \mid g(i) = f(i)\} = \{i \in I \mid f(i) = g(i)\} \in \mathcal{D}$, so $g \sim_{\mathcal{D}} f$, and $\sim_{\mathcal{D}}$ is symmetric.
- Suppose $f \sim_{\mathcal{D}} g$ and $g \sim_{\mathcal{D}} h$. Let $A = \{i \in I \mid f(i) = g(i)\}$, $B = \{i \in I \mid g(i) = h(i)\}$, and $C = \{i \in I \mid f(i) = h(i)\}$. $A \in \mathcal{D}$ and $B \in \mathcal{D}$, so $A \cap B \in \mathcal{D}$. $A \cap B \subseteq C$, so $C \in \mathcal{D}$. Thus $f \sim_{\mathcal{D}} h$, and $\sim_{\mathcal{D}}$ is transitive.

Hence $\sim_{\mathcal{D}}$ is an equivalence relation. □

Definition 3.2.8. With $\{\mathcal{M}_i\}_{i\in I}$, $\mathcal{D}$, and $\sim_{\mathcal{D}}$ as above, the *ultraproduct* $\prod \mathcal{M}_i/\mathcal{D}$, is an $\mathcal{L}$-structure, defined as follows. Let $\mathcal{M} = \prod \mathcal{M}_i/\mathcal{D}$.

- The domain of $\mathcal{M}$, denoted $\prod M_i/\mathcal{D}$, is the set of equivalence classes of $\sim_{\mathcal{D}}$ in $\prod M_i$. We will denote the equivalence class of $f \in \prod M_i$ by $[f]$, or by $[f(i)|i \in I]$.
- For each function symbol $f \in \mathcal{L}$, we define the interpretation $f^{\mathcal{M}}$ by
$$f^{\mathcal{M}}([g_1], \ldots, [g_{n_f}]) = [f^{\mathcal{M}_i}(g_1(i), \ldots, g_{n_f}(i))|i \in I].$$
- For each relation symbol $R \in \mathcal{L}$, we define the interpretation $R^{\mathcal{M}}$ by
$$([g_1], \ldots, [g_{n_R}]) \in R^{\mathcal{M}} \text{ if and only if } \{i \in I \mid (g_1(i), \ldots, g_{n_R}(i)) \in R^{\mathcal{M}_i}\} \in \mathcal{D}.$$
- For each constant symbol $c \in \mathcal{L}$, we define the interpretation $c^{\mathcal{M}}$ by
$$c^{\mathcal{M}} = [c^{\mathcal{M}_i}|i \in I].$$

Using the "almost all" intuition for the ultrafilter $\mathcal{D}$, we can describe the elements of the domain of the ultraproduct as the classes of elements of the Cartesian product which are the same almost everywhere. A relation holds for elements of the ultraproduct if and only if the interpretation of the relation symbol holds for representatives of the element classes in almost all of the $\mathcal{M}_i$.

Functions are applied in the ultraproduct by applying the interpretation of the function symbol to representatives of the element classes in each $\mathcal{M}_i$. A constant in the ultraproduct is simply the equivalence class of the interpretation of the constant in each $\mathcal{M}_i$.

But there is something to check before we can accept this definition. In the definitions of the interpretations $f^{\mathcal{M}}$ and $R^{\mathcal{M}}$, we chose a representative g_i for each equivalence class $[g_i]$. We must show that the interpretations are independent of our choices of representatives.

Proposition 3.2.9. *Let $\mathcal{M} = \prod \mathcal{M}_i/\mathcal{D}$, defined as above. For all function symbols f and relation symbols R in $\mathcal{L}$, the interpretations $f^{\mathcal{M}}$ and $R^{\mathcal{M}}$ are well-defined.*

Proof. Let f be a function symbol. Suppose that for $1 \leq j \leq n_f$ we have $g_j, h_j \in \prod M_i$ with $g_j \sim_{\mathcal{D}} h_j$. If we define $g_f(i) = f^{\mathcal{M}_i}(g_1(i), \ldots, g_{n_f}(i))$ and $h_f(i) = f^{\mathcal{M}_i}(h_1(i), \ldots, h_{n_f}(i))$, we would like to show that $g_f \sim_{\mathcal{D}} h_f$. For all j, let $A_j = \{i \in I \mid g_j(i) = h_j(i)\}$. $A_j \in \mathcal{D}$ for all j, so $\bigcap_{j=1}^{n} A_j \in \mathcal{D}$. Now g_f and h_f certainly agree whenever all of the g_j and h_j agree, so $\bigcap_{j=1}^{n} A_j \subseteq A_f = \{i \in I \mid g_f(i) = h_f(i)\}$, hence $A_f \in \mathcal{D}$, and $g_f \sim_{\mathcal{D}} h_f$.

Let R be a relation symbol. Suppose that for $1 \leq j \leq n_R$ we have $g_j, h_j \in \prod M_i$ with $g_j \sim_{\mathcal{D}} h_j$. If we define $G = \{i \in I \mid (g_1(i), \ldots, g_{n_R}(i)) \in R^{\mathcal{M}_i}\}$ and $H = \{i \in I \mid (h_1(i), \ldots, h_{n_R}(i)) \in R^{\mathcal{M}_i}\}$, we would like to show that $G \in \mathcal{D}$ if and only if $H \in \mathcal{D}$. Again, we define $A_j = \{i \in I \mid g_j(i) = h_j(i)\} \in \mathcal{D}$ for all j. Suppose $G \in \mathcal{D}$. Then

$G \cap \bigcap_{j=1}^{n} A_j \in \mathcal{D}$. Now $(h_1(i), \ldots, h_{n_R}(i))$ is certainly in $R^{\mathcal{M}_i}$ whenever all of the g_j and h_j agree and $(g_1(i), \ldots, g_{n_R}(i))$ is in $R^{\mathcal{M}_i}$, so $G \cap \bigcap_{j=1}^{n} A_j \subseteq H$, and thus $H \in \mathcal{D}$. The converse follows by the same argument. □

The Fundamental Theorem of Ultraproducts states that the ultraproduct $\prod \mathcal{M}_i/\mathcal{D}$ of $\mathcal{L}$-structures satisfies the $\mathcal{L}$-formula ϕ if and only if almost all of the $\mathcal{M}_i$ satisfy ϕ.

Theorem 3.2.10 (Fundamental Theorem of Ultraproducts, [CK73, Theorem 4.1.9])**.** *Let $\mathcal{M}$ be the ultraproduct $\prod \mathcal{M}_i/\mathcal{D}$ of $\mathcal{L}$-structures with ultrafilter $\mathcal{D}$ on index set I. Let $\phi(\overline{v})$ be an $\mathcal{L}$-formula with free variables from $\overline{v} = (v_1, \ldots, v_n)$. Then for $\overline{[g]} = ([g_1], \ldots, [g_n]) \in (\prod M_i/\mathcal{D})^n$, $\prod \mathcal{M}_i/\mathcal{D} \models \phi(\overline{[g]})$ if and only if $\{i \in I \mid \mathcal{M}_i \models \phi(\overline{g(i)})\} \in \mathcal{D}$.*

Proof. The proof is by induction on terms and formulas.

First, we claim that if $t(\overline{v})$ is an $\mathcal{L}$-term, then $t^{\mathcal{M}}(\overline{[g]}) = [t^{\mathcal{M}_i}(\overline{g(i)}) \mid i \in I]$.

If t is a constant symbol c, then the claim is true by definition: $c^{\mathcal{M}} = [c^{\mathcal{M}_i} \mid i \in I]$.

If t is a variable v_j, then $t^{\mathcal{M}}(\overline{[g]}) = [g_j] = [g_j(i) \mid i \in I] = [t^{\mathcal{M}_i}(\overline{g(i)}) \mid i \in I]$.

If t is $f(t_1(\overline{v}), \ldots, t_{n_f}(\overline{v}))$, where f is a function symbol and $t_1, \ldots, t_{n_f}$ are $\mathcal{L}$-terms for which the claim is true, then

$$\begin{aligned}
t^{\mathcal{M}}(\overline{[g]}) &= f^{\mathcal{M}}(t_1^{\mathcal{M}}(\overline{[g]}), \ldots, t_{n_f}^{\mathcal{M}}(\overline{[g]})) \\
&= f^{\mathcal{M}}([t_1^{\mathcal{M}_i}(\overline{g(i)}) \mid i \in I], \ldots, [t_{n_f}^{\mathcal{M}_i}(\overline{g(i)}) \mid i \in I]) \text{ by induction} \\
&= [f^{\mathcal{M}_i}(t_1^{\mathcal{M}_i}(\overline{g(i)}), \ldots, t_{n_f}^{\mathcal{M}_i}(\overline{g(i)})) \mid i \in I] \text{ interpretation of } f \\
&= [t^{\mathcal{M}_i}(\overline{g(i)}) \mid i \in I].
\end{aligned}$$

Having established that terms behave as expected under interpretation, we may move on to proving the theorem. We begin with atomic formulas.

If $\phi(\overline{v})$ is $t_1(\overline{v}) = t_2(\overline{v})$, where t_1 and t_2 are $\mathcal{L}$-terms, then

$$\begin{aligned}
\mathcal{M} \models \phi(\overline{[g]}) \quad &\text{iff} \quad t_1^{\mathcal{M}}(\overline{[g]}) = t_2^{\mathcal{M}}(\overline{[g]}) \\
&\text{iff} \quad [t_1^{\mathcal{M}_i}(\overline{g(i)}) \mid i \in I] = [t_2^{\mathcal{M}_i}(\overline{g(i)}) \mid i \in I] \text{ applying the claim} \\
&\text{iff} \quad \{i \in I \mid t_1^{\mathcal{M}_i}(\overline{g(i)}) = t_2^{\mathcal{M}_i}(\overline{g(i)})\} \in \mathcal{D} \\
&\text{iff} \quad \{i \in I \mid \mathcal{M}_i \models \phi(\overline{g(i)})\} \in \mathcal{D}.
\end{aligned}$$

If $\phi(\overline{v})$ is $R(t_1(\overline{v}), \ldots, t_{n_R}(\overline{v}))$, where R is a relation symbol and $t_1, \ldots, t_{n_R}$ are $\mathcal{L}$-terms, then

$$\begin{aligned}
\mathcal{M} \models \phi(\overline{[g]}) \quad &\text{iff} \quad (t_1^{\mathcal{M}}(\overline{[g]}), \ldots, t_{n_R}^{\mathcal{M}}(\overline{[g]})) \in R^{\mathcal{M}} \\
&\text{iff} \quad ([t_1^{\mathcal{M}_i}(\overline{g(i)}) \mid i \in I], \ldots, [t_{n_R}^{\mathcal{M}_i}(\overline{g(i)}) \mid i \in I]) \in R^{\mathcal{M}} \text{ applying the claim} \\
&\text{iff} \quad \{i \in I \mid (t_1^{\mathcal{M}_i}(\overline{g(i)}), \ldots, t_{n_R}^{\mathcal{M}_i}(\overline{g(i)})) \in R^{\mathcal{M}_i}\} \in \mathcal{D} \text{ interpretation of } R \\
&\text{iff} \quad \{i \in I \mid \mathcal{M}_i \models \phi(\overline{g(i)})\} \in \mathcal{D}.
\end{aligned}$$

If $\phi(\overline{v})$ is $\neg\psi(\overline{v})$, where ψ is an $\mathcal{L}$-formula for which our assertion is true, then

$$\begin{aligned}\mathcal{M} \models \phi(\overline{[g]}) \quad &\text{iff} \quad \mathcal{M} \not\models \psi(\overline{[g]}) \\ &\text{iff} \quad \{i \in I \,|\, \mathcal{M}_i \models \psi(\overline{g(i)})\} \notin \mathcal{D} \text{ by induction} \\ &\text{iff} \quad I \setminus \{i \in I \,|\, \mathcal{M}_i \models \psi(\overline{g(i)})\} \in \mathcal{D} \text{ since } \mathcal{D} \text{ is an ultrafilter} \\ &\text{iff} \quad \{i \in I \,|\, \mathcal{M}_i \not\models \psi(\overline{g(i)})\} \in \mathcal{D} \\ &\text{iff} \quad \{i \in I \,|\, \mathcal{M}_i \models \phi(\overline{g(i)})\} \in \mathcal{D}.\end{aligned}$$

If $\phi(\overline{v})$ is $\psi(\overline{v}) \wedge \theta(\overline{v})$, where ψ and θ are $\mathcal{L}$-formulas for which our assertion is true, then

$$\begin{aligned}\mathcal{M} \models \phi(\overline{[g]}) \quad &\text{iff} \quad \mathcal{M} \models \psi(\overline{[g]}) \text{ and } \mathcal{M} \models \theta(\overline{[g]}) \\ &\text{iff} \quad \{i \in I \,|\, \mathcal{M}_i \models \psi(\overline{g(i)})\} \in \mathcal{D} \text{ and } \{i \in I \,|\, \mathcal{M}_i \models \theta(\overline{g(i)})\} \in \mathcal{D},\end{aligned}$$

by induction. Let $A = \{i \in I \,|\, \mathcal{M}_i \models \psi(\overline{g(i)})\}$ and $B = \{i \in I \,|\, \mathcal{M}_i \models \theta(\overline{g(i)})\}$. If $A \in \mathcal{D}$ and $B \in \mathcal{D}$, then $A \cap B \in \mathcal{D}$. Conversely, $A \cap B \subseteq A$ and $A \cap B \subseteq B$, so if $A \cap B \in \mathcal{D}$, then $A \in \mathcal{D}$ and $B \in \mathcal{D}$. Now,

$$\begin{aligned}A \cap B \in \mathcal{D} \quad &\text{iff} \quad \{i \in I \,|\, \mathcal{M}_i \models \psi(\overline{g(i)}) \text{ and } \mathcal{M}_i \models \theta(\overline{g(i)})\} \in \mathcal{D} \\ &\text{iff} \quad \{i \in I \,|\, \mathcal{M}_i \models \phi(\overline{g(i)})\} \in \mathcal{D}.\end{aligned}$$

If $\phi(\overline{v})$ is $\exists w\, \psi(\overline{v}, w)$, where w is a variable and ψ is an $\mathcal{L}$-formula for which our assertion is true, then $\mathcal{M} \models \phi(\overline{[g]})$ if and only if there exists $[h] \in \prod \mathcal{M}_i/\mathcal{D}$ such that $\mathcal{M} \models \psi(\overline{[g]}, [h])$, if and only if (by induction) there exists $[h]$ such that $\{i \in I \,|\, \mathcal{M}_i \models \psi(\overline{g(i)}, h(i))\} \in \mathcal{D}$. Let $A_{[h]} = \{i \in I \,|\, \mathcal{M}_i \models \psi(\overline{g(i)}, h(i))\}$ and let $B = \{i \in I \,|\, \mathcal{M}_i \models \phi(\overline{g(i)})\}$. We would like to show that there exists $[h]$ such that $A_{[h]} \in \mathcal{D}$ if and only if $B \in \mathcal{D}$.

Suppose there exists such an $[h]$. Then $A_{[h]} \subseteq B$, since for $i \in A$, $\mathcal{M}_i \models \psi(\overline{g(i)}, h(i))$, so $\mathcal{M}_i \models \phi(\overline{g(i)})$, since $h(i)$ satisfies the existential quantifier in $\mathcal{M}_i$. So $B \in \mathcal{D}$. Conversely, suppose $B \in \mathcal{D}$. Let $h \in \prod \mathcal{M}$ be such that for all $i \in B$, $h(i) \in \mathcal{M}_i$ is an element which satisfies the existential quantifier, and for $i \notin B$, $h(i)$ is an arbitrary element. Then $B \subseteq A_{[h]}$, so $A_{[h]} \in \mathcal{D}$.

This completes the proof by induction on formulas. We do not need to consider formulas constructed using $\vee$ or $\forall$, since these can be re-written to use only $\neg$, $\wedge$, and $\exists$. □

Applications of Ultraproducts. The Fundamental Theorem of Ultraproducts has a number of elegant consequences.

Corollary 3.2.11. *If $\mathcal{K}$ is an elementary class of $\mathcal{L}$-structures and $\{\mathcal{M}_i\}_{i \in I}$ is a collection of $\mathcal{L}$-structures in $\mathcal{K}$, indexed by an infinite set I, then for any ultrafilter $\mathcal{D}$ on I, $\prod \mathcal{M}_i/\mathcal{D}$ is in $\mathcal{K}$.*

Proof. Let T be a set of axioms for $\mathcal{K}$. For any $\phi \in T$, $\{i \in I \,|\, \mathcal{M}_i \models \phi\} = I$, since $\mathcal{M}_i \in \mathcal{K}$ for all $i \in I$. Now $I \in \mathcal{D}$, so $\prod \mathcal{M}_i/\mathcal{D} \models \phi$ by Theorem 3.2.10. Thus $\prod \mathcal{M}_i/\mathcal{D} \models T$, and $\prod \mathcal{M}_i/\mathcal{D}$ is in the class $\mathcal{K}$. □

Example 3.2.12. Let $\mathcal{D}$ be a nonprincipal ultrafilter on the set of primes, P. By Corollary 3.2.11, $\prod \mathbb{F}_p((t))/\mathcal{D}$ and $\prod \mathbb{Q}_p/\mathcal{D}$ are valued fields with cross section, since we saw in Example

3.1.12 that the class of valued fields with cross section is elementary. Also, $\prod \mathbb{Q}_p/\mathcal{D}$ has characteristic zero, since $\mathbb{Q}_p \models Char_0$ for all primes p.

Consider the characteristic of $\prod \mathbb{F}_p((t))/\mathcal{D}$. For any prime p, $\mathbb{F}_q((t)) \models Char_p$ if and only if $q = p$, so $\{q \in P \mid \mathbb{F}_q((t)) \models Char_p\} = \{p\}$. This is a finite set, so its complement is in the Frechet filter $\mathcal{D}_F$ on P. But $\mathcal{D}$ contains $\mathcal{D}_F$ by Lemma 3.2.6, so $P \setminus \{p\} \in \mathcal{D}$, and thus $\{p\} \notin \mathcal{D}$. Hence $\prod \mathbb{F}_p((t))/\mathcal{D} \models \neg Char_p$ for all $p \in P$. This means that $\prod \mathbb{F}_p((t))/\mathcal{D}$ has characteristic zero.

This result suggests that the ultraproducts $\prod \mathbb{F}_p((t))/\mathcal{D}$ and $\prod \mathbb{Q}_p/\mathcal{D}$ are very similar structures. The next lemma supports this intuition and sheds some light on the structure of the ultraproducts as valued fields.

Lemma 3.2.13. *For any nonprincipal ultrafilter $\mathcal{D}$, the residue class fields and value groups of $\prod \mathbb{Q}_p/\mathcal{D}$ and $\prod \mathbb{F}_p((t))/\mathcal{D}$ are isomorphic. Specifically,*

- $\mathfrak{v}((\prod \mathbb{Q}_p/\mathcal{D})^*) \cong \prod \mathbb{Z}/\mathcal{D} \cong \mathfrak{v}((\prod \mathbb{F}_p((t))/\mathcal{D})^*)$*, isomorphic as groups,*
- $\overline{\prod \mathbb{Q}_p/\mathcal{D}} \cong \prod \mathbb{F}_p/\mathcal{D} \cong \overline{\prod \mathbb{F}_p((t))/\mathcal{D}}$*, isomorphic as fields, and*
- $char(\prod \mathbb{F}_p/\mathcal{D}) = 0$*, so the residue class fields have characteristic zero.*

Proof. Let $\mathcal{M} = \prod \mathbb{Q}_p/\mathcal{D}$ and take $[f] \in M$, the domain of $\mathcal{M}$, with $[f] \neq [0]$. Then $f \not\sim_{\mathcal{D}} 0$, so $\{p \in P \mid f(p) \neq 0\} \in \mathcal{D}$, and we can find $g \sim_{\mathcal{D}} f$ such that $g(p) \neq 0$ for all p. We will use g as our representative element for the equivalence class $[f]$. By definition, $\mathfrak{v}^{\mathcal{M}}([f]) = [\mathfrak{v}^{\mathbb{Q}_p}(g(p)) \mid p \in P] = [p^{n_p} \mid p \in P]$, for some integers n_p. The value group consists of all equivalence classes of this form. Let ϕ be the function which maps $[p^{n_p} \mid p \in P]$ to $[n_p \mid p \in P] \in \prod \mathbb{Z}/\mathcal{D}$. It is easy to see that ϕ is a bijection, and $\phi([p^{n_p} \mid p \in P] \cdot [p^{m_p} \mid p \in P]) = \phi([p^{n_p+m_p} \mid p \in P]) = [n_p + m_p \mid p \in P] = [n_p \mid p \in P] + [m_p \mid p \in P] = \phi([p^{n_p} \mid p \in P])) + \phi([p^{m_p} \mid p \in P])$, so ϕ is an isomorphism of groups.

The same argument shows that $\mathfrak{v}(\prod \mathbb{F}_p((t))/\mathcal{D}) \cong \prod \mathbb{Z}/\mathcal{D}$, with the minor difference that elements of the value group are of the form $[t^{n_p} \mid p \in P]$, where the n_p are integers, since we choose t instead of p as a prime element for $F_p((t))$.

Thus, $\mathfrak{v}((\prod \mathbb{Q}_p/\mathcal{D})^*) \cong \mathfrak{v}((\prod \mathbb{F}_p((t))/\mathcal{D})^*)$.

Now consider $\overline{\mathcal{M}} = \mathcal{O}/\{[f] \in \mathcal{O} \mid \mathcal{M} \models 1 < \mathfrak{v}([f])\}$ (here we use $t_1 < t_2$ as an abbreviation for $(t_1 \leq t_2) \wedge \neg(t_1 = t_2)$). This is $\mathcal{O}/\{[f] \in \mathcal{M} \mid \{p \mid \mathbb{Q}_p \models 1 < \mathfrak{v}(f(p))\} \in \mathcal{D}\}$. The ideal in the denominator is the maximal idea I_1.

Consider the set R of distinct $[g] \in \mathcal{O}$ such that $g(p) \in \{0, \ldots, p-1\}$ for all $p \in P$. We would like to show that this set is a complete set of representatives for the residue class field.

Take $[g], [h] \in R$, $[g] \neq [h]$. Suppose $[g] \equiv [h] \pmod{I_1}$. Then we have $[g] + [f] = [h]$ for some $[f] \in I_1$. Now if $[f] \in I_1$, then $\{p \mid \mathbb{Q}_p \models 1 < \mathfrak{v}(f(p))\} \in \mathcal{D}$, and since $[g] \neq [h]$, $\{p \mid g(p) \neq h(p)\} \in \mathcal{D}$. But for any p for which both $g(p) \neq h(p)$ and $1 < \mathfrak{v}(f(p))$, we have $g(p) + f(p) \neq h(p)$, since $g(p)$ and $h(p)$ are taken from $\{0, \ldots, p-1\}$, and their difference cannot be divisible by p.

So $\{p \mid \mathbb{Q}_p \models 1 < \mathfrak{v}(f(p))\} \cap \{p \mid g(p) \neq h(p)\} \subseteq \{p \mid g(p) + f(p) \neq h(p)\} \in \mathcal{D}$, and hence $g + f \not\sim_D h$, which is a contradiction. Thus distinct elements of R are representatives for distinct equivalence classes mod I_1.

But every equivalence class mod I_1 has a representative in R, since we can take any representative f of the class and reduce each coordinate $f(p)$ mod p to an element of $\{0, \ldots, p-1\}$.

Letting $[g]$ be the resulting element of R, the difference $[f] - [g] = [f - g]$ is an element of I_1, since each coordinate of $f - g$is divisible by p, so $[g]$ is a representative for the class.

Thus R is a complete set of representatives for the classes mod I_1. Interpreting each $[g] \in R$ as an element of $\prod \mathbb{F}_p/\mathcal{D}$ in the obvious way, it is easy to verify that the resulting map from $\overline{\mathcal{M}}$ to $\prod \mathbb{F}_p/\mathcal{D}$ is an isomorphism of fields.

The same argument holds for $\prod \mathbb{F}_p((t))/\mathcal{D}$, the only difference being that we note that the difference between distinct $g(p)$ and $h(p)$ from $\{0, \ldots, p-1\}$ cannot be divisible by t.

Thus, $\overline{\prod \mathbb{Q}_p/\mathcal{D}} \cong \overline{\prod \mathbb{F}_p((t))/\mathcal{D}}$.

Finally, the residue class fields have characteristic zero by the same argument given in Example 3.2.12. That is, $\mathbb{F}_q \models Char_p$ only when $q = p$, so $\prod \mathbb{F}_p/\mathcal{D} \models \neg Char_p$ for all p, and thus $char(\prod \mathbb{F}_p/\mathcal{D}) = 0$. □

Next we will prove a theorem which is a significant step toward a proof of the Ax-Kochen Principle.

Theorem 3.2.14. *Let $\{\mathcal{M}_i\}_{i\in I}$ and $\{\mathcal{N}_i\}_{i\in I}$ be collections of $\mathcal{L}$-structures, indexed by the same infinite set I. Suppose that for all nonprincipal ultrafilters $\mathcal{D}$, $\prod \mathcal{M}_i/\mathcal{D} \equiv \prod \mathcal{N}_i/\mathcal{D}$. Then for any $\mathcal{L}$-sentence ϕ, $\mathcal{M}_i \models \phi$ for all but finitely many i if and only if $\mathcal{N}_i \models \phi$ for all but finitely many i.*

Proof. Suppose that $\mathcal{M}_i \models \phi$ for all but finitely many i. Then $A = \{i \in I \mid \mathcal{M}_i \models \phi\} \in \mathcal{D}_F$, the Frechet filter on I. By Lemma 3.2.6, all nonprincipal ultrafilters extend the Frechet filter, so $A \in \mathcal{D}$ for all nonprincipal ultrafilters $\mathcal{D}$. By Theorem 3.2.10, $\prod \mathcal{M}_i/\mathcal{D} \models \phi$, and by elementarily equivalence, $\prod \mathcal{N}_i/\mathcal{D} \models \phi$. Again by Theorem 3.2.10, $B = \{i \in I \mid \mathcal{N}_i \models \phi\} \in \mathcal{D}$ for all nonprincipal ultrafilters $\mathcal{D}$, so $B \in \bigcap \mathcal{D} = \mathcal{D}_F$ by Lemma 3.2.6. Thus $\mathcal{N}_i \models \phi$ for all but finitely many i.

The converse follows symmetrically. □

Now that we have established Theorem 3.2.14, the Ax-Kochen Principle will be proven if we can demonstrate the elementary equivalence of the ultraproducts $\prod \mathbb{F}_p((t))/\mathcal{D}$ and $\prod \mathbb{Q}_p/\mathcal{D}$ for all nonprincipal ultrafilters $\mathcal{D}$ on the set P of all primes. But to do so, we will need to appeal to more powerful techniques from model theory and dig deeper into the algebraic properties of $\mathbb{F}_p((t))$ and $\mathbb{Q}_p$.

3.3. **Types and Saturated Models.** This section is largely concerned with a useful technique for proving that two structures are isomorphic, the back-and-forth argument. We will begin with a demonstration of a simple back-and-forth argument in the case of countable dense linear orders without endpoints. We will then generalize the property of these structures which makes the back-and-forth argument possible by discussing types and saturated models.

We will often refer to ordinal and cardinal numbers, and we will assume some knowledge of transfinite arithmetic and transfinite induction. For more information, see Appendix A.

Let $\mathcal{L}_O = \{<\}$, where $<$ is a binary relation symbol, and let DLO be the $\mathcal{L}_O$-theory of dense linear orders without endpoints:

(1) $\forall x \neg(x < x)$
(2) $\forall x \forall y \forall z \, ((x < y \wedge y < z) \rightarrow x < z)$

(3) $\forall x \forall y \, (x < y \vee x = y \vee y < x)$
(4) $\forall x \forall y \, (x < y \to \exists z \, (x < z \wedge z < y))$
(5) $\forall x \exists y \exists z \, (y < x \wedge x < z)$

One example of a model for DLO is the $\mathcal{L}_O$-structure $\langle \mathbb{Q}, < \rangle$, where the interpretation of the symbol $<$ is the usual ordering of $\mathbb{Q}$. It is well known that $\mathbb{Q}$ is countable. We will use a back-and-forth argument to show that up to isomorphism, $\langle \mathbb{Q}, < \rangle$ is the only countable model for DLO. Note that this means up to $\mathcal{L}_O$-isomorphism as an $\mathcal{L}_O$-structure; we view $\mathbb{Q}$ only as an ordered set, not as a field.

Theorem 3.3.1 ([Mar02, Theorem 2.4.1])**.** *Let $\mathcal{A}$ and $\mathcal{B}$ be $\mathcal{L}_O$-structures with domains A and B such that $\mathcal{A} \models$ DLO, $\mathcal{B} \models$ DLO, and $|A| = |B| = \aleph_0$. Then $\mathcal{A} \cong \mathcal{B}$.*

Proof. Since A and B are countable, we can choose enumerations $a_0, a_1, \ldots$ and $b_0, b_1, \ldots$ of A and B. We will inductively construct a sequence of functions $f_i : A_i \to B_i$, between finite subsets $A_i \subset A$ and $B_i \subset B$, satisfying the following properties for each $i \geq 0$:

(1) For all $j < i$, $A_j \subseteq A_i$, $B_j \subseteq B_i$, and $f_j \subseteq f_i$. That is, if $a \in A_j$, then $a \in A_i$, $f_j(a) \in B_i$, and $f_i(a) = f_j(a)$.
(2) If $a < a'$, with $a, a' \in A_i$, then $f_i(a) < f_i(a')$. That is, each f_i is an $\mathcal{L}_O$-homomorphism between the substructures of $\mathcal{A}$ and $\mathcal{B}$ with domains A_i and B_i.
(3) We have $\{a_0, \ldots, a_{i-1}\} \subseteq A_i$ and $\{b_0, \ldots, b_{i-1}\} \subseteq B_i$.
(4) The function f_i is a bijection between A_i and B_i.

Given such a sequence of functions, we let $f = \bigcup_{i=0}^{\infty} f_i : \bigcup_{i=0}^{\infty} A_i \to \bigcup_{i=0}^{\infty} B_i$. Property 1 ensures that f is well-defined, for if $a \in \bigcup_{i=0}^{\infty} A_i$, there is some $N \geq 0$ such that $a \in A_j$ for all $j \geq N$, and the f_j agree on a for all $j \geq N$. Property 2 ensures that f is an $\mathcal{L}_O$-homomorphism. Property 3 ensures that $\bigcup_{i=0}^{\infty} A_i = A$ and $\bigcup_{i=0}^{\infty} B_i = B$, since the a_i and b_i enumerate all of A and B. Property 4 ensures that f is a bijection $A \to B$, and thus an isomorphism between $\mathcal{A}$ and $\mathcal{B}$.

For the base case, let $A_0 = B_0 = f_0 = \emptyset$. The four properties are trivially satisfied.

Given f_i satisfying the four properties, we first extend f_i to a function $g_{i+1} : A'_{i+1} \to B'_{i+1}$ (with $A_i \subseteq A'_{i+1}$ and $B_i \subseteq B'_{i+1}$) by ensuring that a_i is in the domain. Then we extend g_{i+1} to the next function $f_{i+1} : A_{i+1} \to B_{i+1}$ in the sequence by ensuring that b_i is in the range. Going "back and forth" in this way, we will ensure that f_{i+1} satisfies property 3.

If $a_{i+1} \in A_i$, then we simply let $A'_{i+1} = A_i$, $B'_{i+1} = B_i$, and $g_{i+1} = f_i$. Otherwise, we must pick an element $b^* \in B$ onto which to map a_{i+1}. Property 4 requires that $b^* \notin B_i$ (otherwise g_{i+1} would not be injective), and property 2 requires that for all $a \in A_i$, $a < a_{i+1}$ if and only if $f_i(a) < b^*$.

Exactly one of the following holds:

(1) a_{i+1} is less than every element of A_i, or
(2) a_{i+1} is greater than every element of A_i, or
(3) since A_i is finite, there exists a greatest element less than a_{i+1}, α, and a least element greater than a_{i+1}, β, such that $\alpha < \beta$.

In the first case, since $\mathcal{B} \models$ DLO, it has no greatest element, so there is some $b^* \in B \setminus B_i$ greater than every element of B_i. In the second case, $\mathcal{B}$ has no least element, so there is some $b^* \in B \setminus B_i$ less than every element of B_i. In the third case, $\mathcal{B}$ is dense, so we can find

$b^* \in B \setminus B_i$ such that $f_i(\alpha) < b^* < f_i(\beta)$. In any case, we have satisfied $b^* \notin B_i$ and $a < a_{i+1}$ if and only if $f_i(a) < b^*$.

Now define $A'_{i+1} = A_i \cup \{a_{i+1}\}$, $B'_{i+1} = B_i \cup \{b^*\}$, and $g_{i+1} : A'_{i+1} \to B'_{i+1}$ such that $g_{i+1}(a_{i+1}) = b^*$ and for all $a \in A_i$, $g_{i+1}(a) = f_i(a)$.

The other direction is symmetric. If $b_{i+1} \in B_i$, then we simply let $A_{i+1} = A'_{i+1}$, $B_{i+1} = B'_{i+1}$, and $f_{i+1} = g_{i+1}$. Otherwise, we must pick an element $a^* \in A$ to map onto b_{i+1}. We must have $a^* \notin A_i$ (otherwise f_{i+1} would not be well-defined), and property 2 requires that for all $a \in A_i$, $a < a^*$ if and only if $g_i(a) < b_{i+1}$. Using the fact that $\mathcal{A} \models \mathrm{DLO}$, we can pick such an a^* in the same way we picked b^*.

Now define $A_{i+1} = A'_{i+1} \cup \{a^*\}$, $B_{i+1} = B'_{i+1} \cup \{b_{i+1}\}$, and f_{i+1} such that $f_{i+1}(a^*) = b_{i+1}$ and for all $a \in A_i$, $f_{i+1}(a) = g_{i+1}(a)$. Note that we have maintained properties 1-4, and by induction we can construct the required sequence of functions. □

The key property of models of DLO which allows the back-and-forth argument to work is this: given a finite subset of the domain, if we specify a place in the ordering relative to the elements of the subset where we would like to find some new element, we are guaranteed to be able to find such an element, provided that its existence would not break the linearity of the ordering.

Types. To our toolbox of formulas, sentences, and theories, we add a new way of expressing first-order properties, types. Types will allow us to express properties of elements of a structure relative to a distinguished set of other elements. More precisely, given a set A of elements of a structure $\mathcal{M}$, a type captures the relationships that other elements could have relative to the elements of A.

Let $\mathcal{M}$ be an $\mathcal{L}$-structure with domain M. Given $A \subseteq M$, let $\mathcal{L}_A$ be the language $\mathcal{L} \cup \{c_a \,|\, a \in A\}$ where each c_a is a new distinct constant symbol. We view $\mathcal{M}$ is an $\mathcal{L}_A$-structure by interpreting $c_a^{\mathcal{M}} = a$ for each new constant symbol.

When we are working over $\mathcal{L}$, sentences and formulas may only refer explicitly to the elements of M which are interpretations of the constant symbols of $\mathcal{L}$. By expanding the language, we are allowing sentences to refer to the elements of A. Let $Th_A(\mathcal{M}) = \{\phi \,|\, \phi \text{ is an } \mathcal{L}_A\text{-sentence, and } \mathcal{M} \models \phi\}$. This theory is an extension of $Th(\mathcal{M})$, consisting of all sentences which are true in $\mathcal{M}$, when we are allowed to explicitly refer to the elements of A.

Earlier (Definition 3.1.16), we defined satisfiability of an $\mathcal{L}$-theory. There is also a concept of satisfiability of a set of $\mathcal{L}$-formulas.

Definition 3.3.2. A set of $\mathcal{L}$-formulas S with free variables $v_1, \ldots, v_n$ is called *satisfiable* if there is an $\mathcal{L}$-structure $\mathcal{M}$ with domain M and elements $a_1, \ldots, a_n \in M$ such that $\mathcal{M} \models \phi(a_1, \ldots, a_n)$ for all formulas $\phi \in S$. Note that the elements of the domain are fixed. The same value must be substituted for the same variable across all formulas.

Definition 3.3.3. Given an $\mathcal{L}$-structure $\mathcal{M}$ with domain M and a subset $A \subseteq M$, an *n-type* over A is a set P of $\mathcal{L}_A$ formulas in free variables $v_1, \ldots, v_n$, such that $Th_A(\mathcal{M}) \cup P$ is satisfiable. If for all $\mathcal{L}_A$-formulas ϕ in free variables $v_1, \ldots, v_n$, either $\phi \in P$ or $\neg\phi \in P$, then P is called *complete*.

The satisfiability condition here means that there is some $\mathcal{L}_A$-structure $\mathcal{N}$ with domain N which is a model for $Th_A(\mathcal{M})$, and that there are elements $b_1, \ldots, b_n \in N$ such that $\mathcal{N} \models \phi(b_1, \ldots, b_n)$ for all $\phi \in P$. We say that the elements $b_1, \ldots, b_n$ realize P in $\mathcal{N}$. If P is not realized in $\mathcal{M}$ we say that $\mathcal{M}$ omits the type P.

Example 3.3.4. We return to the example of $\langle \mathbb{Q}, < \rangle$ as an $\mathcal{L}_O$-structure. Let $A = \mathbb{N} \subset \mathbb{Q}$. We will define two types over A.

Let $P = \{c_0 < v_1, c_1 < v_1, c_2 < v_1, \ldots\}$. Note that the elements of $\mathbb{N}$ appear (represented by constant symbols) in the formulas of P. In order to show that P is a 1-type over A, we must show that $Th_A(\langle \mathbb{Q}, < \rangle) \cup P$ is satisfiable. Let $\Delta \subset Th_A(\langle \mathbb{Q}, < \rangle) \cup P$ be a finite subset. Only finitely many of the formulas in P appear in Δ, so we let i be the maximum integer such that $i < v_1 \in \Delta$. Then for all $\phi(v_1) = j < v_1 \in \Delta \cap P$, $\langle \mathbb{Q}, < \rangle \models \phi(i+1)$. The other sentences in Δ are true in $\langle \mathbb{Q}, < \rangle$ by definition, so $\langle \mathbb{Q}, < \rangle \models \Delta$. Thus $Th_A(\langle \mathbb{Q}, < \rangle) \cup P$ is finitely satisfiable, and therefore satisfiable by Compactness.

So P is a 1-type over A, but $\langle \mathbb{Q}, < \rangle$ omits P, since there is no rational number which is greater than every natural number.

Let $Q = \{\phi(v_1) \,|\, \langle \mathbb{Q}, < \rangle \models \phi(\frac{1}{2})\}$. The element $\frac{1}{2}$ realizes Q in $\langle \mathbb{Q}, < \rangle$, so $Q \cup Th_{\mathbb{Q}}(\langle \mathbb{Q}, < \rangle)$ is clearly satisfiable, and Q is a 1-type over A. Moreover, for any $\mathcal{L}_A$-formula ψ in one free variable, $\langle \mathbb{Q}, < \rangle \models \psi(\frac{1}{2})$ or $\langle \mathbb{Q}, < \rangle \models \neg\psi(\frac{1}{2})$, so either $\psi \in Q$ or $\neg\psi \in Q$. Thus Q is complete.

There is a useful generalization of the type Q in Example 3.3.4. For any $\mathcal{L}$-structure $\mathcal{M}$ with domain M, $A \subseteq M$, and elements $m_1, \ldots, m_n \in M$, we define $tp^{\mathcal{M}}(m_1, \ldots, m_n/A) = \{\phi(v_1, \ldots, v_n) \,|\, \mathcal{M} \models \phi(m_1, \ldots, m_n)\}$. By the argument given in the example, this type, called the complete type of $a_1, \ldots, a_n$ over A, is a complete n-type which is realized in $\mathcal{M}$.

Saturated Models. A κ-saturated model realizes all types over sets of cardinality less than κ.

Definition 3.3.5. Let T be a complete theory with infinite models in a countable language $\mathcal{L}$. Let κ be an infinite cardinal. A model $\mathcal{M} \models T$ with domain M is called *κ-saturated* if for all $A \subset M$ with $|A| < \kappa$ every type over A is realized in $\mathcal{M}$.

In Theorem 3.3.1, we constructed an isomorphism between any two countable models for DLO. Using a similar argument, we can construct partial elementary bijections between subsets of the domains of κ-saturated models.

Theorem 3.3.6 ([CK73, Lemma 5.1.11])**.** *Let κ be an infinite cardinal, and let $\mathcal{M}$ and $\mathcal{N}$ be κ-saturated models of a complete theory T with domains M and N respectively. Let $A \subseteq M$ and $B \subseteq N$ be subsets such that $|A| = |B| = \kappa$. Then there is a partial elementary bijection $f : \widetilde{A} \to \widetilde{B}$, where $A \subseteq \widetilde{A} \subseteq M$ and $B \subseteq \widetilde{B} \subseteq N$, and $|\widetilde{A}| = |\widetilde{B}| = \kappa$. By a partial elementary bijection, we mean that $\mathcal{M} \models \phi(x_1, \ldots, x_n)$ for ϕ an $\mathcal{L}$-formula and $x_1, \ldots, x_n \in \widetilde{A}$ if and only if $\mathcal{N} \models \phi(f(x_1), \ldots, f(x_n))$.*

Proof. Let $(a_\alpha : \alpha < \kappa)$ and $(b_\alpha : \alpha < \kappa)$ be enumerations of A and B respectively.

We will inductively construct a sequence of functions $(f_\alpha : \alpha < \kappa)$ such that each f_α is a partial elementary bijection with domain $A_\alpha \subset M$ and range $B_\alpha \subset N$. We require f_α to satisfy the following properties for all $\alpha < \kappa$:

(1) For all $\beta < \alpha$, $A_\beta \subseteq A_\alpha$, $B_\beta \subseteq B_\alpha$, and $f_\beta \subseteq f_\alpha$.

(2) We have $a_\alpha \in A_{\alpha+1}$ and $b_\alpha \in B_{\alpha+1}$.
(3) The function f_α is a bijection between A_α and B_α.
(4) We have $|A_\alpha| \leq |2\alpha| < \kappa$ and $|B_\alpha| \leq |2\alpha| < \kappa$.

Given such a sequence of functions, we let $\widetilde{A} = \bigcup_{\alpha<\kappa} A_\alpha$, $\widetilde{B} = \bigcup_{\alpha<\kappa} B_\alpha$, and $f = \bigcup_{\alpha<\kappa} f_\alpha$. Properties 1 and 3 guarantee that f is a well-defined elementary bijection $\widetilde{A} \to \widetilde{B}$, property 2 guarantees that $A \subseteq \widetilde{A}$ and $B \subseteq \widetilde{B}$, and property 4 guarantees that $|\widetilde{A}| = |\widetilde{B}| = \kappa$.

For the base case, let $A_0 = B_0 = f_0 = \emptyset$. The properties are trivially satisfied.

If α is a limit ordinal, we define $A_\alpha = \bigcup_{\beta<\alpha} A_\beta$, $B_\alpha = \bigcup_{\beta<\alpha} B_\beta$, and $f_\alpha = \bigcup_{\beta<\alpha} f_\beta$. Property 1 is clearly satisfied. Property 2 only requires certain elements to be in the domain and range of successor ordinals, so it is trivially satisfied. Property 3 is satisfied, since every element of B_α is in the range of some f_β for $\beta < \alpha$, since the f_β are surjective, and any two elements in A_α are in some A_β for $\beta < \alpha$, so they are sent to distinct elements, since the f_β are injective. Property 4 is true by transfinite arithmetic: α is the limit of the $\beta < \alpha$, so $|2\alpha|$ is the limit of $|2\beta|$ for $\beta < \alpha$, which bounds the cardinality of $\bigcup_{\beta<\alpha} A_\beta$ above. The argument for the cardinality of B_α is the same. Finally, f_α is elementary, since the same holds for all f_β, $\beta < \alpha$.

In the successor case, given f_α satisfying the properties, we first extend f_α to a function $g_{\alpha+1} : A'_{\alpha+1} \to B'_{\alpha+1}$ by ensuring that a_α is in the domain. Then we extend $g_{\alpha+1}$ to the next function $f_{\alpha+1} : A_{\alpha+1} \to B_{\alpha+1}$ in the sequence by ensuring that b_α is in the range.

If $a_\alpha \in A_\alpha$, then we simply let $g_{\alpha+1} = f_\alpha$. Otherwise, we must pick an element $n \in N$ onto which to map a_α. Consider the language $\mathcal{L}_{A_\alpha}$, which is $\mathcal{L}$ extended with a new constant symbol for each element in A_α. We may consider $\mathcal{N}$ as a $\mathcal{L}_{A_\alpha}$-structure by interpreting the constant symbol c_a (representing $a \in A_\alpha$) as $f_\alpha(a) \in B_\alpha$. We will identify the languages $\mathcal{L}_{A_\alpha}$ and $\mathcal{L}_{B_\alpha}$ by choosing the same constant symbol, c_a, to represent $a \in A_\alpha$ and $f_\alpha(a) \in B_\alpha$.

Now for any $\mathcal{L}_{A_\alpha}$-sentence ϕ, enumerate the new constant symbols which appear in ϕ, $c_{m_1}, \ldots, c_{m_t}$, and let $m_1, \ldots, m_t$ and $n_1, \ldots, n_t$ be their interpretations in $\mathcal{M}$ and $\mathcal{N}$, respectively. Form an $\mathcal{L}$-sentence ψ by replacing each new constant symbol c_{m_i} with a distinct variable v_i. Then $\mathcal{M} \models \phi$ if and only if $\mathcal{M} \models \psi(m_1, \ldots, m_t)$. Since f_α is elementary and $f_\alpha(m_i) = n_i$ for all i, $\mathcal{M} \models \psi(m_1, \ldots, m_t)$ if and only if $\mathcal{N} \models \psi(n_1, \ldots, n_t)$, if and only if $\mathcal{N} \models \phi$. Thus $Th_{A_\alpha}(\mathcal{M}) = Th_{B_\alpha}(\mathcal{N})$.

We can easily show by the method in Example 3.3.4 that $tp^{\mathcal{M}}(a_\alpha/A_\alpha)$ is a complete type realized in $\mathcal{M}$. Thus $Th_{A_\alpha}(\mathcal{M}) \cup tp^{\mathcal{M}}(a_\alpha/A_\alpha)$ is satisfiable, so $Th_{B_\alpha}(\mathcal{N}) \cup tp^{\mathcal{M}}(a_\alpha/A_\alpha)$ is satisfiable, and $tp^{\mathcal{M}}(a_\alpha/A_\alpha)$ is a complete type over B_α in $\mathcal{N}$, since we have identified the languages $\mathcal{L}_{A_\alpha}$ and $\mathcal{L}_{B_\alpha}$. Since $|B_\alpha| < \kappa$, and $\mathcal{N}$ is κ-saturated, $tp^{\mathcal{M}}(a_\alpha/A_\alpha)$ is realized in $\mathcal{N}$. Let n be an element of N which realizes this type.

Note that $n \notin B_\alpha$. For otherwise, letting ϕ be the $\mathcal{L}_{A_\alpha}$-formula $v_1 = c_n$, $\mathcal{N} \models \phi(n)$, so $\mathcal{M} \models \phi(a_\alpha)$, and $a_\alpha = m$, where m is the interpretation of c_n in A_α, and hence $a_\alpha \in A_\alpha$. But this contradicts the assumption that $a_\alpha \notin A_\alpha$.

Now define $g_{\alpha+1}$ by extending f_α such that $g_{\alpha+1}(a_\alpha) = n$. It is evident that $g_{\alpha+1}$ is elementary, for $\mathcal{M} \models \phi(a_\alpha, m_1, \ldots, m_t)$ with $m_1, \ldots, m_t \in A_\alpha$ if and only if $\phi(v_1, c_{m_1}, \ldots, c_{m_t}) \in tp^{\mathcal{M}}(a_\alpha/A_\alpha)$, if and only if $\mathcal{N} \models \phi(n, f_\alpha(m_1), \ldots, f_\alpha(m_t))$.

The other direction is symmetric. If $b_\alpha \in B_\alpha$, then we simply let $f_{\alpha+1} = g_{\alpha+1}$. Otherwise, we may pick an element $m \in M$ to map onto b_α such that m realizes $tp^{\mathcal{N}}(b_\alpha/B_\alpha)$. We define $f_{\alpha+1}$ by extending $g_{\alpha+1}$ such that $f_{\alpha+1}(m) = b_\alpha$. By the same argument, $f_{\alpha+1}$ is elementary.

Properties 1 and 2 are clearly satisfied. It is also clear that $f_{\alpha+1}$ is a surjection, since for every element we have added to the range, we have added an element to the domain mapping to it. By the observations that $m \notin A_\alpha$ and $n \notin B_\alpha$, $f_{\alpha+1}$ is well-defined and injective, so Property 3 is satisfied. Finally, we have added at most two elements to the domain and range, so since $|A_\alpha| \leq |2\alpha|$, $|A_{\alpha+1}| \leq |2(\alpha+1)|$. The same argument holds for the cardinality of $B_{\alpha+1}$, so Property 4 is satisfied.

Thus by transfinite induction we are able to construct the required sequence of functions. This completes the proof. □

Note that if a model $\mathcal{M}$ with domain M is κ-saturated, then we must have $\kappa \leq |M|$. For otherwise, if $|M| < \kappa$, then $\mathcal{M}$ would realize every type over M, the entire domain. In particular, it would realize the type $\{\neg(v_1 = c_m) \,|\, m \in M\}$. Any element $x \in M$ which realizes this type is not equal to any element of M, which is a contradiction.

If $\mathcal{M}$ is $|M|$-saturated, that is, as saturated as possible, then we simply say that $\mathcal{M}$ is saturated. As a corollary to the last theorem, saturated models of a given cardinality are unique up to isomorphism.

Corollary 3.3.7. *If $\mathcal{M}$ and $\mathcal{N}$ are saturated models of a complete theory, T, and they have the same cardinality κ, then $\mathcal{M} \cong \mathcal{N}$.*

Proof. We apply Theorem 3.3.6, taking as our subsets the entire domains M and N of $\mathcal{M}$ and $\mathcal{N}$, since $|M| = |N| = \kappa$. Then there is a function $f : M \to N$ which is an elementary bijection, and thus an isomorphism, between $\mathcal{M}$ and $\mathcal{N}$. □

Existence of Saturated Models. Now that the utility of saturated models for demonstrating isomorphism is clear, we will take up the issue of their existence. The following lemma shows that if a type P is omitted, we can always find an elementary extension in which P is realized.

Lemma 3.3.8. *Let $\mathcal{M}$ be an infinite $\mathcal{L}$-structure with domain M, $A \subseteq M$, and P an n-type over A. Then there exists an elementary extension of $\mathcal{M}$, $\mathcal{N}$, such that P is realized in $\mathcal{N}$. If $\mathcal{L}$ is countable, we can take $\mathcal{N}$ to have the same cardinality as $\mathcal{M}$.*

Proof. Since P is a type over A, $P \cup Th_A(\mathcal{M})$ is satisfiable. Let $\mathcal{N}_0$ be an $\mathcal{L}_A$-structure which satisfies $P \cup Th_A(\mathcal{M})$, and let $x_1, \ldots, x_n$ be the elements realizing P in $\mathcal{N}_0$.

Let $\Gamma = P \cup Diag_{el}(\mathcal{M})$. We will apply Compactness to prove that Γ is satisfiable by showing that $\mathcal{N}_0$ satisfies every finite subset of Γ.

Note that P consists of $\mathcal{L}_A$-formulas (with constant symbols for each element of A) and $Diag_{el}(\mathcal{M})$ consists of $\mathcal{L}_\mathcal{M}$-sentences (with constant symbols for each element of M). Since $\mathcal{L}_A \subseteq \mathcal{L}_\mathcal{M}$, we can consider the formulas in P as $\mathcal{L}_\mathcal{M}$-formulas.

Let Δ be a finite subset of Γ. There are finitely many $\mathcal{L}_\mathcal{M}$-formulas $\phi_1, \ldots, \phi_s \in P \cap \Delta$ and finitely many $\mathcal{L}_\mathcal{M}$-sentences $\psi_1, \ldots, \psi_t \in Diag_{el}(\mathcal{M}) \cap \Delta$. Let $\Phi(v_1, \ldots, v_n)$ be the $\mathcal{L}_\mathcal{M}$-formula $\phi_1 \wedge \phi_2 \wedge \ldots \wedge \phi_s$, and let Ψ be the $\mathcal{L}_\mathcal{M}$-sentence $\psi_1 \wedge \psi_2 \wedge \ldots \wedge \psi_t$.

Let $c_{a_1}, \ldots, c_{a_j}, c_{b_1}, \ldots, c_{b_k}$ be the new constant symbols of $\mathcal{L}_{\mathcal{M}}$ which appear in the formulas and sentences of Δ. The symbol c_{a_i} corresponds to the element $a_i \in A$ and the symbol c_{b_i} corresponds to the element $b_i \in M \setminus A$.

Now to show that the $\mathcal{L}_A$-structure $\mathcal{N}_0$ satisfies Δ, we must turn it into an $\mathcal{L}_{\mathcal{M}}$ structure by assigning interpretations to the constant symbols c_m for all $m \in M$. But since $c_{b_1} \ldots, c_{b_k}$ are the only symbols appearing in Δ, all other c_m may be assigned interpretations arbitrarily. Then it will suffice to show that $\mathcal{N}_0 \models \Phi(x_1, \ldots, x_n) \wedge \Psi$.

Let $\Psi'(w_1, \ldots, w_k)$ be the $\mathcal{L}_A$-formula formed by replacing each constant symbol c_{b_i} in Ψ with the variable w_i. Let θ be the $\mathcal{L}_A$-sentence $\exists w_1 \ldots \exists w_k \, \Psi'(w_1, \ldots, w_k)$. Now $\mathcal{M} \models \Psi'(b_1, \ldots, b_j)$, so $\mathcal{M} \models \theta$, and thus $\theta \in Th_A(\mathcal{M})$. But $\mathcal{N}_0 \models Th_A(\mathcal{M})$, so $\mathcal{N}_0 \models \theta$.

Thus there are elements $y_1, \ldots, y_k$ in the domain of $\mathcal{N}_0$ such that $\mathcal{N}_0 \models \Psi'(y_1, \ldots, y_k)$. Interpreting the constant c_{b_i} as y_i for each i, we see that $\mathcal{N}_0 \models \Psi$ as an $\mathcal{L}_{\mathcal{M}}$-structure. But also $\mathcal{N}_0 \models \Phi(x_1, \ldots, x_n)$ since $x_1, \ldots, x_n$ realize P in $\mathcal{N}_0$, so $\mathcal{N}_0 \models \Phi(x_1, \ldots, x_n) \wedge \Psi$, and thus $\mathcal{N}_0$ satisfies Δ.

Hence Γ is finitely satisfiable, and there is a model $\mathcal{N}$ for Γ of cardinality $|M|$ by Theorem 3.1.19. Now P is realized in $\mathcal{N}$, and moreover $\mathcal{N} \models Diag_{el}(\mathcal{M})$, so there is an elementary embedding of $\mathcal{M}$ into $\mathcal{N}$ by Lemma 3.1.29. Identifying $\mathcal{M}$ with $j(\mathcal{M})$, we can view $\mathcal{N}$ as an elementary extension of $\mathcal{M}$. □

Now that we can add elements to realize types, we can construct saturated models for certain cardinalities.

Theorem 3.3.9 ([Mar02, Theorem 4.3.12])**.** *Let $\mathcal{L}$ be a countable language, and let $\mathcal{M}$ be an infinite $\mathcal{L}$-structure with domain M. Let κ be an infinite cardinal. Then there is an elementary extension $\mathcal{N}$ of $\mathcal{M}$ with domain N such that $|N| \leq |M|^\kappa$, and $\mathcal{N}$ is κ^+-saturated.*

Proof. First we will note that in order to prove that a model is κ-saturated, it suffices to prove that the model realizes all 1-types. The general case follows by induction: If P is an n-type over A, let $Q = \{\phi(v_1, \ldots, v_{n-1}) \,|\, \phi \in P\}$, the $(n-1)$-type consisting of all formulas in P which do not include the last variable v_n. By induction, Q is realized by some $a_1, \ldots, a_{n-1} \in M$. Let $R = \{\psi(c_{a_1}, \ldots, c_{a_{n-1}}, v_n) \,|\, \psi(v_1, \ldots, v_n) \in P\}$, which is a 1-type (in the free variable v_n) over $A \cup \{a_1, \ldots, a_{n-1}\}$. Since adding finitely many elements does not increase the cardinality of A, R is realized in $\mathcal{M}$ by the base case. Suppose b realizes R. Then $a_1, \ldots, a_{n-1}, b$ realizes P.

Returning to the proof of the theorem, we will first apply Lemma 3.3.8 repeatedly to build a chain of elementary extensions of $\mathcal{M}$, each of which satisfies a particular type.

Our claim is that there exists an elementary extension $\mathcal{M}'$ of $\mathcal{M}$ with $|M'| \leq |M|^\kappa$ such that for all $A \subseteq M$ with $|A| \leq \kappa$, every 1-type over A is realized in $\mathcal{M}'$.

We need to pin down how many types we may need to satisfy. The number of subsets of M with cardinality less than or equal to κ is less than the number of functions, $\kappa \to M$, since each such subset is the range of one of these functions. This set of functions has cardinality $|M|^\kappa$.

Now given a subset $A \subseteq M$ with $|A| \leq \kappa$, the language $\mathcal{L}_A$ has cardinality at most κ, since $\mathcal{L}$ is countable. The set of $\mathcal{L}_A$-formulas is a subset of the set of finite strings with symbols from $\mathcal{L}_A$ plus our finite set of boolean connectors, quantifiers, etc. The cardinality of the set of finite strings of any given length l is the cardinality of the set of functions,

$\{1, \ldots, l\} \to (\mathcal{L}_A \cup \{v_1, \wedge, \vee, \ldots\})$, and this set of functions has cardinality $\kappa^l = \kappa$. Now there are $\aleph_0$ values for l, so the set of $\mathcal{L}_A$-formulas has cardinality $\aleph_0 \kappa = \kappa$, since $\aleph_0 \leq \kappa$.

Now types are elements of the power set of the set of $\mathcal{L}_A$-formulas, so the cardinality of the set of $\mathcal{L}_A$-formulas is at most the cardinality of the power set, that is, 2^κ.

Hence the total number of 1-types over all subsets A, with $|A| \leq \kappa$, is bounded above by $|M|^\kappa 2^\kappa = |M|^\kappa$, since $2 < |M|$.

Let $(P_\alpha : \alpha < |M|^\kappa)$ be an enumeration of all such types. We will build an elementary chain $(\mathcal{M}_\alpha : \alpha < |M|^\kappa)$ as follows:

(1) Let $\mathcal{M}_0 = \mathcal{M}$.
(2) For α a limit ordinal, let $\mathcal{M}_\alpha = \bigcup_{\beta<\alpha} \mathcal{M}_\beta$.
(3) For all α, apply Lemma 3.3.8 to find an elementary extension $\mathcal{M}_{\alpha+1}$ of $\mathcal{M}_\alpha$ such that $|M_{\alpha+1}| = |M_\alpha|$ and $\mathcal{M}_{\alpha+1}$ realizes P_α.

Now let $\mathcal{M}' = \bigcup_{\alpha<|M|^\kappa} \mathcal{M}_\alpha$. Since every type P_α is realized in some $\mathcal{M}_\alpha$, $\mathcal{M}'$ realizes every such type. It remains to show that $|M'| \leq |M|^\kappa$.

We will show by induction that for all $\alpha < |M|^\kappa$, $|M_\alpha| \leq |M|^\kappa$. In the base case, $|M_0| = |M| \leq |M|^\kappa$. If $|M_\alpha| \leq |M|^\kappa$, then $|M_{\alpha+1}| = |M_\alpha| \leq |M|^\kappa$. If α is a limit ordinal, then $|M_\alpha|$ is the union of a chain of sets of cardinality at most $|M|^\kappa$, so its cardinality is the limit of the cardinalities of these sets, which is bounded above by $|M|^\kappa$. So $|M_\alpha| \leq |M|^\kappa$.

Now $\mathcal{M}'$ is the union of a chain of sets of cardinalities at most $|M|^\kappa$, so it has cardinality at most $|M|^\kappa$. This completes the proof of the claim.

We have that $\mathcal{M}'$ realizes every 1-type over every subset $A \subset M$ with $|A| \leq \kappa$, but we do not yet have the $\mathcal{M}'$ is κ^+-saturated, since its domain is larger than that of $\mathcal{M}$, and thus there are additional types to realize.

We build another elementary chain $(\mathcal{N}_\alpha : \alpha < \kappa^+)$ as follows:

(1) Let $\mathcal{N}_0 = \mathcal{M}$.
(2) For α a limit ordinal, let $\mathcal{N}_\alpha = \bigcup_{\beta<\alpha} \mathcal{N}_\beta$.
(3) For all α, apply the previous claim to find an elementary extension $\mathcal{N}_{\alpha+1}$ of $\mathcal{N}_\alpha$ such that for all $A \subseteq N_\alpha$ (where N_α is the domain of $\mathcal{N}_\alpha$) with $|A| \leq \kappa$, every 1-type over A is realized in $\mathcal{N}_{\alpha+1}$. By the claim, $|N_{\alpha+1}| \leq |N_\alpha|^\kappa$.

Now let $\mathcal{N} = \bigcup_{\alpha<\kappa^+} \mathcal{N}_\alpha$. Let N be the domain of $\mathcal{N}$. For every $A \subseteq N$ such that $|A| < \kappa^+$, A is contained in the domain of some $\mathcal{N}_\alpha$, and every 1-type over A is realized in $\mathcal{N}_{\alpha+1}$, and therefore in $\mathcal{N}$. Thus $\mathcal{N}$ is κ^+-saturated. It remains to show that $|N| \leq |M|^\kappa$.

We will show by induction that for all $\alpha < \kappa^+$, $|N_\alpha| \leq |M|^\kappa$. In the base case, $|N_0| = |M| \leq |M|^\kappa$. If $|N_\alpha| \leq |M|^\kappa$, then $|N_{\alpha+1}| \leq |N_\alpha|^\kappa \leq (|M|^\kappa)^\kappa = |M|^\kappa$. If α is a limit ordinal, then $|N_\alpha|$ is the union of a chain of sets of cardinality at most $|M|^\kappa$, so $|N_\alpha| \leq |M|^\kappa$.

Now $\mathcal{N}$ is the union of a chain of sets of cardinalities at most $|M|^\kappa$ so it has cardinality at most $|M|^\kappa$. This completes the proof. □

Corollary 3.3.10. *If we assume the Continuum Hypothesis, there is a saturated model of $Th(\mathcal{M})$ with cardinality $\aleph_1$.*

Proof. By the Löwenheim-Skolem theorem (Theorem 3.1.25), there is a model $\mathcal{M}' \models Th(\mathcal{M})$ of cardinality $\aleph_0$. Applying Theorem 3.3.9, there is a $\aleph_1$-saturated elementary extension $\mathcal{N}$ of $\mathcal{M}'$ with domain N of cardinality at most $\aleph_0^{\aleph_0}$. If we assume the Continuum Hypothesis,

$\aleph_0^{\aleph_0} = \aleph_1$. Since an $\aleph_1$-saturated model must have cardinality at least $\aleph_1$, $|N| = \aleph_1$, and hence $\mathcal{N}$ is a saturated model for $Th(\mathcal{M})$. □

The existence of saturated models makes many results in model theory easier to prove, including the Ax-Kochen Principle. There are methods to eliminate the Continuum Hypothesis from some proofs which use saturated models, one of which is to employ a generalization of saturated models, called special models. For more information, see Appendix B. We will assume the existence of saturated models in order to simplify our arguments.

4. The Ax-Kochen Principle

4.1. **Hensel's Lemma.** One of our key tools in establishing the Ax-Kochen Principle will be Hensel's lemma. In valued fields in which Hensel's lemma holds, one can lift information about polynomials over the residue class field to polynomials over the valuation ring.

Definition 4.1.1. Let F be a valued field. We say that F is *Henselian* if F has the following property, which is one of the formulations of Hensel's lemma:

Let $f, g_0, h_0 \in \mathcal{O}[x]$ be monic polynomials. If the images of g_0 and h_0 in $\overline{F}[x]$, $\overline{g_0}$ and $\overline{h_0}$, are relatively prime, and if $\overline{g_0}\overline{h_0} = \overline{f}$, then there exist $g, h \in \mathcal{O}[x]$ such that $\overline{g} = \overline{g_0}$, $\overline{h} = \overline{h_0}$, and $f = gh$.

In this section, we will show that all complete discrete valued fields, and in particular the fields $\mathbb{Q}_p$ and $\mathbb{F}_p((t))$, are Henselian.

Throughout this section, we will assume that all valued fields have cross section, and we will write our value groups multiplicatively. This is a change from the notation in Section 2.4. In particular, we will take as the value group of a discrete valued field the group $\{\pi^n \mid n \in \mathbb{Z}\}$, where π is a prime element. We must specify what is meant by homomorphism of valued fields and valued subfield in this context.

Definition 4.1.2. Let F and F' be valued fields (with cross section), and let $\mathfrak{v}_F : F \to F$ and $\mathfrak{v}_{F'} : F' \to F'$ be their valuations. We say that a function $f : F \to F'$ is a *homomorphism of valued fields* if it is a field homomorphism which preserves valuations, that is, if $f \circ \mathfrak{v}_F = \mathfrak{v}_{F'} \circ f$. We say that a subfield $K \subseteq F$ is a *valued subfield* if $\mathfrak{v}_F(K) \subseteq K$. In this case, K is a valued field with cross section whose valuation is the restriction of $\mathfrak{v}_F$ to K.

We will assume familiarity with the resultant, an algebraic construction which gives information about the common roots of polynomials. The necessary facts about the resultant are developed in Appendix C. We will use the following results:

Theorem C.2. *Let R be a ring. If $f, g \in R[x]$ are relatively prime, $Res(f, g) \neq 0$.*

Lemma C.3. *Let R be a subring of a field K, and let $g, h \in R[x]$ with $deg(g) = m$, $deg(h) = n$. If $\rho = Res(g, h) \neq 0$, then for all $l \in R[x]$ such that $deg(l) \leq m + n - 1$, there exist $\phi, \psi \in R[x]$ with $deg(\phi) \leq n - 1$, $deg(\psi) \leq m - 1$ such that $g\phi + h\psi = \rho l$.*

Most of the work of proving Hensel's lemma in complete discrete valued fields is done in the following more general lemma.

Lemma 4.1.3 ([BS66, Theorem 4.3.1])**.** *Let F be a complete discrete valued field with valuation $\mathfrak{v}$ and prime element π. Let $f \in \mathcal{O}[x]$ be a polynomial of degree $m+n$. Suppose there are polynomials $g_0, h_0 \in \mathcal{O}[x]$ of degrees m and n respectively such that*

1. *f and $g_0 h_0$ have the same leading coefficient,*
2. *$Res(g_0, h_0) \neq 0$, and*
3. *letting $\pi^r = \mathfrak{v}(Res(g_0, h_0))$, we have $f \equiv g_0 h_0 \pmod{\pi^{2r+1}}$.*

Then there exist polynomials $g, h \in \mathcal{O}[x]$ of degrees m and n respectively such that

1. *$f = gh$,*
2. *$g \equiv g_0$ and $h \equiv h_0 \pmod{\pi^{r+1}}$, and*
3. *g and g_0 have the same leading coefficient, as do h and h_0.*

Proof. For all $k \geq 0$, we will define polynomials g_k and h_k so that $deg(g_k) = deg(g_0)$, $deg(h_k) = deg(h_0)$, and $f \equiv g_k h_k \pmod{\pi^{2r+k+1}}$.

We will accomplish this by defining $\phi_k, \psi_k \in \mathcal{O}[x]$ for all $k \geq 1$ and letting $g_k = g_0 + \pi^{r+1}\phi_1 + \ldots + \pi^{r+k}\phi_k$ and $h_k = h_0 + \pi^{r+1}\psi_1 + \ldots + \pi^{r+k}\psi_k$. If we require that $deg(\phi_k) \leq m-1$ and $deg(\psi_k) \leq n-1$, then we will have $deg(g_k) = deg(g_0)$ and $deg(h_k) = deg(h_0)$.

In the base case, we simply have $f \equiv g_0 h_0 \pmod{\pi^{2r+1}}$ by assumption.

Suppose we have $f \equiv g_{k-1} h_{k-1} \pmod{\pi^{2r+k}}$. Then $f = g_{k-1}h_{k-1} + \pi^{2r+k} l$, for some $l \in \mathcal{O}[x]$.

By inductive assumption, g_0 and g_{k-1} have the same leading coefficient, as do h_0 and h_{k-1}. The leading coefficient of f is the same as the product of the leading coefficients of g_0 and h_0 by assumption, so it is the same as the product of the leading coefficients of g_{k-1} and h_{k-1}. Thus l must have degree less than $m+n$.

Also, $g_{k-1} \equiv g_0$, $h_{k-1} \equiv h_0 \pmod{\pi^{r+1}}$ by construction, so $Res(g_{k-1}, h_{k-1}) \equiv Res(g_0, h_0) \pmod{\pi^{r+1}}$. But $\mathfrak{v}(Res(g_0, h_0)) = \pi^r$ by assumption, so $\mathfrak{v}(Res(g_{k-1}, h_{k-1})) = \pi^r$, and $Res(g_{k-1}, h_{k-1}) = \pi^r u$ for some unit u.

Applying Lemma C.3, there exist polynomials $\phi_k, \psi_k \in \mathcal{O}[x]$ such that $g_{k-1}\psi_k + h_{k-1}\phi_k = (\pi^r u)(u^{-1} l) = \pi^r l$ with $deg(\phi_k) \leq m-1$ and $deg(\psi_k) \leq n-1$.

Now we define $g_k = g_0 + \pi^{r+1}\phi_1 + \ldots + \pi^{r+k}\phi_k$ and $h_k = h_0 + \pi^{r+1}\psi_1 + \ldots + \pi^{r+k}\psi_k$. We need to show that $f \equiv g_k h_k \pmod{\pi^{2r+k+1}}$.

Expanding,

$$\begin{aligned} g_k h_k &= (g_{k-1} + \pi^{r+k}\phi_k)(h_{k-1} + \pi^{r+k}\psi_k) \\ &= g_{k-1}h_{k-1} + \pi^{r+k}(g_{k-1}\psi_k + h_{k-1}\phi_k) + \pi^{2r+2k}\phi_k\psi_k \\ &= (f - \pi^{2r+k} l) + \pi^{2r+k} l + \pi^{2r+2k}\phi_k\psi_k \\ &= f + \pi^{2r+2k}\phi_k\psi_k. \end{aligned}$$

So $f \equiv g_k h_k \pmod{\pi^{2r+k+1}}$, as was to be shown, since $2k \geq k+1$ when $k \geq 1$.

Having established the claim by induction, we let $g = g_0 + \pi^{r+1}\phi_1 + \pi^{r+2}\phi_2 + \ldots$, which is an element of $\mathcal{O}[x]$, since F is complete. Similarly, we let $h = h_0 + \pi^{r+1}\psi_1 + \pi^{r+2}\psi_2 + \ldots \in \mathcal{O}[x]$.

Since the degrees of the ϕ_i are all less than m and the degrees of the ψ_i are all less than n, the leading coefficient of g is the same as that of g_0, and the leading coefficient of h is the same as that of h_0. We also clearly have $g \equiv g_0$ and $h \equiv h_0 \pmod{\pi^{r+1}}$.

Finally, $f \equiv g_k h_k \pmod{\pi^{2r+k+1}}$ for all $k \geq 1$, and $g_k h_k \equiv gh \pmod{\pi^{r+k+1}}$ for all $k \geq 1$, so $f \equiv gh \pmod{\pi^{r+k+1}}$ for all $k \geq 1$. Letting k to go infinity, we have $f = gh$ in $\mathcal{O}$. □

Theorem 4.1.4. *All complete discrete valued fields are Henselian.*

Proof. Let F be a complete discrete valued field, and let $f, g_0, h_0 \in \mathcal{O}[x]$ be monic polynomials, such that the images of g_0 and h_0 in $\overline{F}[x]$, $\overline{g_0}$ and $\overline{h_0}$, are relatively prime, and $\overline{g_0}\overline{h_0} = \overline{f}$. We would like to show that there exist $g, h \in \mathcal{O}[x]$ such that $\overline{g} = \overline{g_0}$, $\overline{h} = \overline{h_0}$, and $gh = f$.

Let $m = deg(g_0)$ and $n = deg(h_0)$. Since f, g_0, and h_0 are monic, $\overline{f}$, $\overline{g_0}$, and $\overline{h_0}$ are also monic. Then $deg(\overline{g_0}) = m$, $deg(\overline{h_0}) = n$, and since $\overline{f} = \overline{g_0}\overline{h_0}$, $deg(f) = m + n$. Also, f and $g_0 h_0$ have the same leading coefficient, 1. This satisfies condition 1 of Lemma 4.1.3.

Let $\rho = Res(g_0, h_0)$. Since ρ is computed from the coefficients of g_0 and h_0 by addition and multiplication, we can mod out by π, and $\overline{\rho} = Res(\overline{g_0}, \overline{h_0})$. Now $\overline{g_0}$ and $\overline{h_0}$ are relatively prime in $\overline{F}$, so $Res(\overline{g_0}, \overline{h_0}) \neq 0$. Thus $\rho \not\equiv 0 \pmod{\pi}$, and in particular $\rho \neq 0$, satisfying condition 2. Hence $\mathfrak{v}(\rho) = \pi^0$. Finally, $\overline{f} = \overline{g_0}\overline{h_0}$, so $f \equiv g_0 h_0 \pmod{\pi}$, satisfying condition 3 of Lemma 4.1.3 with $r = 0$.

The lemma gives us polynomials $g, h \in \mathcal{O}[x]$ such that $g \equiv g_0 \pmod{\pi}$, $h \equiv h_0 \pmod{\pi}$, and $f = gh$, as required by Hensel's lemma. □

Consequences of Hensel's Lemma. Henselian valued fields have a number of properties which will be useful in the proof of the Ax-Kochen Principal. The most important is that the residue class field of a Henselian valued field can be embedded as a subfield in the case of characteristic zero.

Lemma 4.1.5 ([CK73, Lemma 5.4.13 (ii)])**.** *Let F be a Henselian valued field with valuation $\mathfrak{v}$ such that $char(\overline{F}) = 0$. Then there exists a subfield $F_0 \subseteq \mathcal{O}$ such that $F_0 \cong \overline{F}$, where the isomorphism is given by $\phi(x) = \overline{x}$.*

Proof. Since $char(\overline{F}) = 0$, then we must also have $char(F) = 0$, for if $1 + \ldots + 1 = 0$, then $\overline{1} + \ldots + \overline{1} = \overline{0}$. Thus there is a natural embedding of the rationals into F. Our first step will be to show that rationals in F are contained in $\mathcal{O}$.

Since $\mathcal{O}$ is a ring, all integers in F are elements of $\mathcal{O}$. Recall that $\mathfrak{v}(1) = 1$, since $\mathfrak{v}$ is a multiplicative group homomorphism. Since $char(\overline{F}) = 0$, if n is a positive integer, $\overline{n} = \overline{1} + \ldots + \overline{1} \neq 0$, so n is not in the maximal ideal I_1, and we do not have $\mathfrak{v}(n) > 1$. But $n \in \mathcal{O}$, so $\mathfrak{v}(n) = 1$. By Lemma 2.4.6, $\mathfrak{v}(-n) = \mathfrak{v}(n) = 1$, so all integers have valuation 0.

Now if m/n is a rational in F, $\mathfrak{v}(m/n) = \mathfrak{v}(m)\mathfrak{v}(n)^{-1} = 1$. So the rationals are a subfield of the valuation ring $\mathcal{O}$.

By a simple application of Zorn's lemma, the rationals are contained in a maximal subfield F_0 of $\mathcal{O}$. More explicitly, the union of a chain of fields is a field, so every chain of subfields of $\mathcal{O}$ has an upper bound, and by Zorn's lemma, the set of subfields of $\mathcal{O}$ has maximal elements.

Since only elements with valuation 1 have inverses in $\mathcal{O}$, $\mathfrak{v}(x) = 1$ for all nonzero $x \in F_0$. Thus 0 is the only element of F_0 with valuation greater than 1, and the kernel of the residue map ϕ is trivial, so ϕ maps F_0 isomorphically onto a subfield G_0 of $\overline{F}$. We will use Hensel's lemma to show that every element of $\overline{F}$ must be in G_0, proving their equality.

Suppose that $a \in \overline{F}$ and a is algebraic over G_0. Then there is a monic irreducible polynomial $p_0(x) \in G_0[x]$ such that $p_0(a) = 0$. Choosing preimages for all coefficients of p_0

under ϕ, we obtain a polynomial $p \in F_0[x]$ such that $\overline{p} = p_0$. Now $\overline{p}$ factors in $\overline{F}[x]$ as $\overline{p}(x) = q_0(x)(x-a)$, where $q_0(x)$ and $(x-a)$ are relatively prime since $char(\overline{F}) = 0$. Applying Hensel's lemma, there are polynomials $q, r \in \mathcal{O}[x]$ such that $\overline{q} = q_0$, $\overline{r} = x - a$, and $p = qr$.

Let c and d be the leading coefficients of q and r respectively. Since p is monic, $cd = 1$, so $\mathfrak{v}(c)\mathfrak{v}(d) = 1$, but $c, d \in \mathcal{O}$, so $\mathfrak{v}(c) = \mathfrak{v}(d) = 1$. Thus neither have residue 0, and we have $deg(q) = deg(q_0)$ and $deg(r) = deg(x-a) = 1$.

Let $r = b_1 x + b_0$. Then $y = -b_0/b_1$ is a root of r and therefore a root of p. Now F_0 is isomorphic to G_0, and $\overline{p}$ is irreducible in G_0, so p is irreducible in F_0, and hence $y \notin F_0$.

We showed that b_1 has valuation 1. Also $b_0 \in \mathcal{O}$, so $\mathfrak{v}(b_0) \geq 1$, and so $\mathfrak{v}(y) = \mathfrak{v}(b_0)\mathfrak{v}(b_1)^{-1} \geq 1$, and $y \in \mathcal{O}$. Since all the generators of $F_0[y]$ are in the ring $\mathcal{O}$, $F_0[y] \subseteq \mathcal{O}$. This contradicts the maximality of F_0 as a subfield of $\mathcal{O}$. Hence there are no elements of $\overline{F}$ algebraic over G_0.

Now suppose that $a \in \overline{F}$ and a is transcendental over G_0. We will use the same contradiction strategy. Pick some $y \in \mathcal{O}$ such that $\overline{y} = a$. Now for any nonzero $p(x) \in F_0[x]$, $\overline{p(y)} = \overline{p}(a) \neq 0$, since a is transcendental over G_0. In particular, $p(y) \neq 0$, so y is transcendental over F_0. Also, $p(y)$ does not have residue 0, but $p(y) \in \mathcal{O}$, so $\mathfrak{v}(p(y)) = 1$. Thus for any $q, r \in F_0[x]$, $\mathfrak{v}(q(y)/r(y)) = 1$, so $q(y)/r(y) \in \mathcal{O}$. All elements of the transcendental extension have this form, so $F_0(y)$ is contained in $\mathcal{O}$, once again contradicting the maximality of F_0.

Since there are no elements of $\overline{F}$ algebraic or transcendental over G_0, $G_0 = \overline{F}$, and thus ϕ is an isomorphism onto $\overline{F}$, and $F_0 \cong \overline{F}$. □

We will now introduce the Henselization of a valued field, which is similar in concept to algebraic closure. Intuitively, the Henselization of a valued field is the minimal Henselian valued field containing it. The Henselization is defined by means of a universal property.

Definition 4.1.6. Let G be a valued field. A Henselian valued field K is said to be a *Henselization* of G if

(1) the field G is a valued subfield of K, and
(2) if F is a Henselian valued field and $\mu : G \to F$ is an embedding of valued fields, then μ extends uniquely to an embedding $\lambda : K \to F$ of valued fields.

The universal property guarantees that the Henselization is unique up to unique isomorphism. For suppose that K and K' are both Henselizations of the valued field G. Then G certainly embeds into both K and K', so there exist unique embeddings of valued fields $\lambda : K \to K'$ and $\lambda' : K' \to K$. Then $\lambda' \circ \lambda : K \to K$ is an embedding of valued fields. But by the universal property, there is a unique embedding of valued fields $K \to K$, and this must be the identity. So $\lambda' \circ \lambda = id_K$, and λ is an isomorphism.

The Henselization of any valued field F may be constructed as follows. Let F^s be the separable closure of F in its algebraic closure (in characteristic 0, this is just the algebraic closure). Extend the valuation $\mathfrak{v}$ to a valuation $\mathfrak{v}^s$ on F^s, and let L be the subgroup of the Galois group of F^s over F consisting of all automorphisms which are also valued field homomorphisms, that is, they preserve $\mathfrak{v}^s$. The fixed field of L is the Henselization of F. In the special case that F has a completion, $\widehat{F}$, this construction corresponds to the separable

closure (or relative algebraic closure in the case of characteristic 0) of F in $\widehat{F}$. We have the following lemma, which we will state here without proof.

Lemma 4.1.7 ([Rib99, Theorem 5.2])**.** *Every valued field F (with valuation $\mathfrak{v}_F$) has a Henselization K (with valuation $\mathfrak{v}_K$), which is unique up to unique isomorphism of valued fields. The value groups and residue class fields of K and F are isomorphic as groups and fields respectively.*

The next lemmas consist of a number of facts about Henselizations and extensions of Henselian fields which will be necessary for the back-and-forth argument in Section 4.2. The proofs of these facts are technical, and they rely on too much of the theory of valued fields to develop in this thesis. They can be found in full in Ribenboim [Rib99] and Chang and Keisler [CK73].

Lemma 4.1.8 ([Rib99, Theorem 5.1])**.** *The valuation on a Henselian valued field extends in a unique way to an algebraic extension. Equivalently, let F_0 and G_0 be Henselian valued fields with valuations $\mathfrak{v}_{F_0}$ and $\mathfrak{v}_{G_0}$ which are isomorphic as valued fields. Let F and G be algebraic extensions of F_0 and G_0 respectively which are isomorphic as fields, and let ϕ be the isomorphism. Given extensions $\mathfrak{v}_F$ and $\mathfrak{v}_G$, of $\mathfrak{v}_{F_0}$ and $\mathfrak{v}_{G_0}$ which are valuations on F and G respectively, ϕ is an isomorphism of valued fields.*

Lemma 4.1.9 ([CK73, Lemma 5.4.13 (vi)])**.** *Let F be a valued field with valuation $\mathfrak{v}$, F_0 a valued subfield, and $\widetilde{F_0}$ the relative algebraic closure of F_0 in F, that is, the set of all elements in F algebraic over F_0. Then $\mathfrak{v}(\widetilde{F_0}^*) = \{x \in \mathfrak{v}(F^*) \mid x^n \in \mathfrak{v}(F_0^*), n \in \mathbb{Z}\}$, the closure under roots of $\mathfrak{v}(F_0^*)$ in $\mathfrak{v}(F^*)$. If F is a Henselian valued field such that $\overline{F} = \overline{F_0}$, $char(\overline{F}) = 0$, and $\mathfrak{v}(\widetilde{F_0^*}) = \mathfrak{v}(F_0^*)$ (that is, $\mathfrak{v}(F_0^*)$ is already closed under roots), then $\widetilde{F_0}$ is a Henselization of F_0.*

Lemma 4.1.10 ([CK73, Lemma 5.4.13 (vii)])**.** *Let F and G be Henselian valued fields (with valuations $\mathfrak{v}_F$ and $\mathfrak{v}_G$) with Henselian valued subfields F_0 and G_0 respectively such that $F_0 \cong G_0$ by an isomorphism f. Suppose $x \in F$ is transcendental over F_0 and $y \in G$ is transcendental over G_0. Suppose further that $\mathfrak{v}(F_0(x)^*) = \mathfrak{v}(F_0^*)$, $\overline{F_0(x)} = \overline{F_0}$, and for all $a \in F_0$, $f(\mathfrak{v}_F(x-a)) = \mathfrak{v}_G(y - f(a))$. Then $\mathfrak{v}_G(G_0(y)^*) = \mathfrak{v}_G(G_0^*)$, $\overline{G_0(y)} = \overline{G_0}$, and f can be extended to an isomorphism $F_0(x) \cong G_0(y)$.*

Lemma 4.1.11 ([CK73, Lemma 5.4.13 (viii)])**.** *Let F be a Henselian valued field with a Henselian valued subfield F_0, and suppose $x \in F$ is transcendental over F_0. If $\overline{F_0(x)} \cong \overline{F_0}$ and $\mathfrak{v}(F_0^*)$ is nontrivial, then adjoining x does not increase the cardinality of the value group: $|\mathfrak{v}(F_0(x)^*)| = |\mathfrak{v}(F_0^*)|$.*

4.2. **Establishing Elementary Equivalence.** We are now in a position to prove the Ax-Kochen Principle. We saw in Theorem 3.2.14 that it suffices to prove the elementary equivalence of the ultraproducts $\prod \mathbb{Q}_p/\mathcal{D}$ and $\prod \mathbb{F}_p((t))/\mathcal{D}$ for all nonprincipal ultrafilters $\mathcal{D}$ on the set of all primes P.

In Example 3.2.12, we applied Corollary 3.2.11 to show that for any nonprincipal ultrafilter $\mathcal{D}$, $\prod \mathbb{Q}_p/\mathcal{D}$ and $\prod \mathbb{F}_p((t))/\mathcal{D}$ are valued fields with cross section. We would like to show that these valued fields are Henselian.

Lemma 4.2.1. *The class of Henselian valued fields with cross section is elementary.*

Proof. We have already exhibited a set of first-order axioms in the language $\mathcal{L}_{VF}$ for the class of valued fields with cross section (see Example 3.1.12), so it remains to extend these axioms to include Hensel's lemma.

In order to express Hensel's lemma, we must be able to make statements about polynomials whose coefficients lie in the domain of our structure. We will represent a polynomial of degree at most n by an $(n+1)$-tuple of coefficients and make appropriate statements about these coefficients.

We first note we can quantify over polynomials by using one quantifier for each coefficient. For a polynomial of degree at most n we will write $\exists_n f(x)$ and $\forall_n f(x)$ as abbreviations for $\exists a_0 \ldots \exists a_n$ and $\forall a_0 \ldots \forall a_n$. The $(n+1)$-tuple $(a_0, \ldots, a_n)$ will then represent $f(x)$. Here we are using the variables a_i for clarity. Formally, they are choices of v_j from our infinite set of variables $\mathcal{V} = \{v_1, v_2, \ldots\}$.

For the following abbreviations, suppose $f(x)$ is represented by $(a_0, \ldots, a_n)$, and $g(x)$ is represented by $(b_0, \ldots, b_m)$, with $m \geq n$.

We can express the statement that a polynomial has coefficients in the valuation ring $\mathcal{O}$ by requiring that all of its coefficients have valuation at least 1. We will write $f \in \mathcal{O}[x]$ as an abbreviation for

$$(1 \leq \mathfrak{v}(a_0)) \wedge \ldots \wedge (1 \leq \mathfrak{v}(a_n).$$

We can easily express the statement that a polynomial is monic. We will write $Monic(f)$ as an abbreviation for

$$a_n = 1.$$

We can express equality of polynomials by stating the equality of the coefficients. We will write $f = g$ as an abbreviation for

$$(a_0 = b_0) \wedge \ldots \wedge (a_n = b_n) \wedge (0 = b_{n+1}) \wedge \ldots \wedge (0 = b_m).$$

We can form new polynomials by addition, subtraction, and multiplication. We will write $f + g$ to mean the polynomial which is represented by $(a_0 + b_0, \ldots, a_n + b_n, b_{n+1}, \ldots, b_m)$, and subtraction is just addition with an application of the additive inverse function, $-$, to each coefficient of g. We will write fg to mean the polynomial which is represented by the $(nm+1)$-tuple $(a_0b_0, a_0b_1 + a_1b_0, \ldots, a_nb_m)$.

We will also need to work with the images of polynomials in the residue class field. To express equality of the images of two polynomials in the residue class field, we will state that their difference is a polynomial whose coefficients all have valuation greater than 1 (and thus is the zero polynomial in the residue class field). Using the standard abbreviation $(x < y)$ for $(x \leq y) \wedge \neg(x = y)$, we will write $ResEq(f, g)$ as an abbreviation for

$$(1 < \mathfrak{v}(a_0 + -(b_0)) \wedge \ldots \wedge (1 < \mathfrak{v}(a_n + -(b_n))) \wedge (1 < \mathfrak{v}(-(b_{n+1}))) \wedge \ldots \wedge (1 < \mathfrak{v}(-(b_m)))).$$

Finally, Hensel's lemma includes the statement that the images of two polynomials in the residue class field are relatively prime. We will write $ResRelPrime(f, g)$ as an abbreviation for

$$\neg(\exists_m p(x) \exists_m q(x) \exists_m r(x)\, (p \in \mathcal{O}[x]) \wedge (q \in \mathcal{O}[x]) \wedge (r \in \mathcal{O}[x]) \wedge ResEq(pq, f) \wedge ResEq(pr, g)).$$

Using these abbreviations, the following first-order sentence, $Hensel_n$, expresses Hensel's lemma for polynomials of degree at most n.

$$\begin{aligned}\forall_n f(x)\forall_n g_0(x)\forall_n h_0(x) \quad & ((f \in \mathcal{O}[x]) \wedge (g_0 \in \mathcal{O}[x]) \wedge (h_0 \in \mathcal{O}[x]) \\ & \wedge Monic(f) \wedge Monic(g_0) \wedge Monic(h_0) \\ & \wedge ResRelPrime(g_0, h_0) \wedge ResEq(g_0h_0, f)) \rightarrow \\ (\exists_n g(x)\exists_n h(x) \quad & (g \in \mathcal{O}[x]) \wedge (h \in \mathcal{O}[x]) \\ & \wedge ResEq(g, g_0) \wedge ResEq(h, h_0) \wedge (f = gh))\end{aligned}$$

Appending the infinite set of sentences $\{Hensel_n \mid n \in \mathbb{N}\}$ to the axioms for the theory of valued fields with cross section, we obtain a set of axioms for the theory of Henselian valued fields with cross section. □

Together with Corollary 3.2.11, Lemma 4.2.1 shows that for any ultrafilter $\mathcal{D}$, $\prod \mathbb{Q}_p/\mathcal{D}$ and $\prod \mathbb{F}_p((t))/\mathcal{D}$ are Henselian. In the next theorem, we will work with the value groups of these valued fields as $\mathcal{L}_G$-structures, where $\mathcal{L}_G = \{\cdot, 1\}$ is the language of groups, with symbols interpreted in the natural way. Similarly, we will work with the residue class fields as $\mathcal{L}_F$-structures, where $\mathcal{L}_F = \{+, \cdot, -, 0, 1\}$ is the language of fields, with symbols interpreted in the natural way.

In Lemma 3.2.13, we showed that the value groups and residue class fields of the ultraproducts are isomorphic. The following general theorem shows that this is enough to prove that the ultraproducts themselves are elementarily equivalent. The proof, which is quite lengthy, uses a back-and-forth argument, properties of valuations, types and saturation, and the lemmas at the end of Section 4.1.

Theorem 4.2.2 ([CK73, Theorem 5.4.12])**.** *Suppose that F and G are Henselian valued fields with valuations $\mathfrak{v}_F$ and $\mathfrak{v}_G$ respectively such that $\mathfrak{v}_F(F^*) \equiv \mathfrak{v}_G(G^*)$ (as $\mathcal{L}_G$-structures), $\overline{F} \equiv \overline{G}$ (as $\mathcal{L}_F$-structures), and $char(\overline{F}) = char(\overline{G}) = 0$. Then $F \equiv G$.*

Proof. Suppose F_1 and G_1 are valued subfields of F and G. We will write $f_1 : F_1 \leftrightarrow G_1$ if and only if f_1 is an isomorphism between F_1 and G_1, and f_1 restricted to the value group of F_1 is a partial elementary bijection between $\mathfrak{v}_F(F_1^*)$ and $\mathfrak{v}_G(G_1^*)$ as subsets of $\mathfrak{v}_F(F^*)$ and $\mathfrak{v}_G(G^*)$.

The plan for the proof is as follows:

(1) We will show that we can reduce to the case in which F and G are saturated models of cardinality $\aleph_1$.
(2) We will show that the residue class fields $\overline{F}$ and $\overline{G}$ are relatively algebraically closed subfields of F and G respectively, and there exists $f_0 : \overline{F} \leftrightarrow \overline{G}$.
(3) We will show that given $f_1 : F_1 \leftrightarrow G_1$ between relatively algebraically closed valued subfields of F and G, such that $\overline{F} \subseteq F_1$, $\overline{G} \subseteq G_1$, $f_0 \subseteq f_1$, and $\mathfrak{v}_F(F_1^*)$ and $\mathfrak{v}_G(G_1^*)$ are countable, then given any element $x \in F$, we can extend f_1 to $f_2 : F_2 \leftrightarrow G_2$ such that $x \in F_2$, $F_1 \subseteq F_2$, $G_1 \subseteq G_2$, $f_1 \subseteq f_2$, and $\mathfrak{v}_F(F_2)$ and $\mathfrak{v}_G(G_2)$ are countable.
(4) We will show that the same extension result holds if we exchange the roles of F and G.

(5) We will use these results and a back-and-forth argument to construct an isomorphism between F and G.

(1) *Reducing to the saturated case.* Assuming the Continuum Hypothesis, there exist saturated models $\mathcal{F} \models Th(F)$ and $\mathcal{G} \models Th(G)$ of cardinality $\aleph_1$, by Theorem 3.3.10. Once again, we stress that the Continuum Hypothesis merely simplifies our arguments, and there are methods for eliminating it from the proof (see Appendix B).

We would like to show that $\mathcal{F}$ and $\mathcal{G}$ satisfy the conditions of the theorem. Since the class of Henselian valued fields is elementary, the axioms for the class are a subset of both $Th(F)$ and $Th(G)$, so $\mathcal{F}$ and $\mathcal{G}$ are Henselian valued fields.

For any $\mathcal{L}_F$-formula ϕ about the residue class field, we can transform ϕ into an $\mathcal{L}_{VF}$-formula ϕ' about the valued field as follows. First, we will restrict all variables to the valuation ring by replacing every instance of a quantifier $\exists v_i\, \psi$ or $\forall v_i\, \psi$ (where ψ is some formula) with $\exists v_i\, (1 \leq \mathfrak{v}(v_i)) \wedge \psi$ or $\forall v_i\, (1 \leq \mathfrak{v}(v_i)) \rightarrow \psi$. Additionally, if $v_1, \ldots, v_j$ are free variables in ϕ, we restrict these to the valuation ring as well by adding to ϕ: $(1 \leq \mathfrak{v}(v_1)) \wedge \ldots \wedge (1 \leq \mathfrak{v}(v_j) \wedge \phi$. Now we will replace equality by congruence modulo the maximal ideal I_1 by replacing every instance of $t_1 = t_2$ (where t_1 and t_2 are terms) with $1 < \mathfrak{v}(t_1 + -(t_2))$.

It should be easy to convince yourself that for every $\mathcal{L}_F$-sentence ϕ, $\overline{\mathcal{F}} \models \phi$ if and only if $\mathcal{F} \models \phi'$. But $\mathcal{F} \equiv F$, so $\mathcal{F} \models \phi'$ if and only if $F \models \phi'$ if and only if $\overline{F} \models \phi$. By the same argument, $\overline{\mathcal{G}} \models \phi$ if and only if $\overline{G} \models \phi$. But $\overline{F} \equiv \overline{G}$, so $\overline{\mathcal{F}} \models \phi$ if and only if $\overline{\mathcal{G}} \models \phi$, and thus $\overline{\mathcal{F}} \equiv \overline{\mathcal{G}}$.

In particular, the $\mathcal{L}_F$-sentences $\neg Char_p$ for each prime p in the theory $Char_0$ can each be transformed by this method. Call the theory made up of these transformed sentences $Char_0'$. Since the residue class field of F has characteristic zero, $\overline{F} \models Char_0$, so $F \models Char_0'$, and since $F \equiv \mathcal{F}$, $\mathcal{F} \models Char_0'$, and thus $\overline{\mathcal{F}} \models Char_0$. Since $\overline{\mathcal{G}} \equiv \overline{\mathcal{F}}$, both residue class fields have characteristic zero.

Similarly, for any $\mathcal{L}_G$-formula ϕ about the value group, we can transform ϕ into an $\mathcal{L}_{VF}$-formula ϕ' about the valued field by restricting all variables to the value group. We replace every instance of a quantifier $\exists v_i\, \psi$ or $\forall v_i\, \psi$ (where ψ is some formula) with $\exists v_i\, V(v_i) \wedge \psi$ or $\forall v_i\, V(v_i) \rightarrow \psi$. Additionally, if $v_1, \ldots, v_j$ are free variables in ϕ, we restrict these to the value group as well by adding to ϕ: $V(v_1) \wedge \ldots \wedge V(v_j) \wedge \phi$. Again, $\mathfrak{v}_F(\mathcal{F}^*) \models \phi$ if and only if $\mathcal{F} \models \phi'$, so $\mathfrak{v}_F(\mathcal{F}^*) \equiv \mathfrak{v}_G(\mathcal{G}^*)$ by the same argument.

Thus $\mathcal{F}$ and $\mathcal{G}$ satisfy the conditions of the theorem. It suffices to prove the theorem in the special case of saturated models of cardinality $\aleph_1$, since then we will have $\mathcal{F} \equiv \mathcal{G}$. But $F \equiv \mathcal{F}$ and $G \equiv \mathcal{G}$, so we will have $F \equiv G$, completing the proof in general.

For the remainder of the proof, we will assume that F and G are saturated models of cardinality $\aleph_1$. In order to prove that $F \equiv G$, we will prove the stronger condition (by Theorem 3.1.30) that $F \cong G$. Keep in mind that we are only working with the saturated case. We do not claim that the valued fields are isomorphic in the general case, just elementarily equivalent.

Note that we cannot simply apply Corollary 3.3.7 to show that $F \cong G$, even though these models are saturated. The corollary requires the additional assumption that $F \equiv G$, which is exactly what we are trying to prove. However, we will use this corollary to show that $\overline{F} \cong \overline{G}$.

As an additional consequence of the method of transforming formulas about the residue class field and value group into formulas about the valued field, we will show that $\overline{F}$ and $\overline{G}$ are saturated models, and except in the trivial case $\mathfrak{v}_F(F^*) = \mathfrak{v}_G(G^*) = \{1\}$, $\mathfrak{v}_F(F^*)$ and $\mathfrak{v}_G(G^*)$ are saturated models.

Note that the residue class fields are infinite by characteristic zero, and except in the trivial case, the value groups are infinite by the linear order properties. Any type in the residue class field or value group of F over a set of cardinality at most $\aleph_1$ can be transformed into a type in F over a set of cardinality at most $\aleph_1$, and since all such types are realized in F, this type is realized in the residue class field or value group, and hence these models are $\aleph_1$-saturated. The same argument holds for the residue class field and value group of G.

Now any $\aleph_1$-saturated model has cardinality at least $\aleph_1$. But the value group of F is a subset of F, so it has cardinality at most $\aleph_1$. And the residue class field of F embeds into F by Lemma 4.1.5, so it has cardinality at most $\aleph_1$. The same argument holds in G, so $\overline{F}$, $\overline{G}$, $\mathfrak{v}_F(F^*)$, and $\mathfrak{v}_G(G^*)$ are all saturated, and $|\overline{F}| = |\overline{G}| = |\mathfrak{v}_F(F^*)| = |\mathfrak{v}_G(G^*)| = \aleph_1$.

(2) *The base case: residue class fields.* By Lemma 4.1.5, $\overline{F}$ and $\overline{G}$ are isomorphic to subfields of F and G. We will identify the residue class fields with these subfields. Since $\overline{F}$ and $\overline{G}$ are saturated, and by assumption $\overline{F} \equiv \overline{G}$, $\overline{F} \cong \overline{G}$ by Corollary 3.3.7. Call the isomorphism f_0. We have $f_0 : \overline{F} \leftrightarrow \overline{G}$, since f_0 is an elementary bijection between $\mathfrak{v}_F(\overline{F}^*)$ and $\mathfrak{v}_G(\overline{G}^*)$, as these value groups only contain a singe element, 1, which is already a constant symbol in the language.

Now we can dispense with the trivial case $\mathfrak{v}_F(F^*) = \mathfrak{v}_G(G^*) = \{1\}$, for in this case, $\overline{F} = F$ and $\overline{G} = G$, so f_0 provides the desired isomorphism $F \cong G$. We will assume for the remainder of the proof that we are not in the trivial case, and thus the value groups of F and G are saturated of cardinality $\aleph_1$.

It remains to show that $\overline{F}$ and $\overline{G}$ are relatively algebraically closed in F and G. Let $p \in \overline{F}[t]$. Since $\overline{F} \subseteq \mathcal{O}_F$ by Lemma 4.1.5, we can consider p as an element of $\mathcal{O}_F[t]$. Let $x \in F$ such that $\mathfrak{v}_F(x) < 1$. Substituting x for t, $p(x)$ is a sum of terms of the form $a_m x^m$, where $\mathfrak{v}_F(a_m) = 1$ if $a_m \neq 0$. Thus $\mathfrak{v}_F(a_m x^m) = \mathfrak{v}_F(x)^m < 1$. Each of the powers of $\mathfrak{v}_F(x)$ is distinct, so $p(x)$ is a sum of terms with distinct values, and by Lemma 2.4.6 (4), $\mathfrak{v}_F(p(x))$ is the minimum of these, which is less than 1. But $\mathfrak{v}_F(0) > 1$, so x is not a root of p.

Thus every root of p in F is an element of $\mathcal{O}_F$. Let x be one such root. Then $\overline{x}$ is defined, and since $\overline{p} = p$, $p(\overline{x}) = \overline{p(x)} = \overline{0} = 0$. So p already has the root $\overline{x} \in \overline{F}$, and we can factor p in $\overline{F}$ as $(t - \overline{x})q$, where q is a polynomial of lower degree. Now if $x \neq \overline{x}$, x is still a root of q, and we can apply the same argument again to factor q. Repeating this process until we reach a polynomial of degree 1, we see that we must have $x = \overline{x}$. So every root of p in F is in $\overline{F}$, and $\overline{F}$ is relatively algebraically closed.

The same argument shows that $\overline{G}$ is relatively algebraically closed in G.

(3-4) *The inductive step: extending isomorphisms.* We have $f_1 : F_1 \leftrightarrow G_1$ between relatively algebraically closed subfields of F and G, such that $\overline{F} \subseteq F_1$, $\overline{G} \subseteq G_1$, $f_0 \subseteq f_1$, and $\mathfrak{v}_F(F_1^*)$ and $\mathfrak{v}_G(G_1^*)$ are countable. Suppose $x \in F_1$. Then x is already in the domain, so we can easily satisfy (3) by simply taking $f_2 = f_1$.

Otherwise, suppose $x \notin F_1$. Since F_1 is relatively algebraically closed in F, x is transcendental over F_1. We will first prove (3) in two special cases, then prove the general case.

Case 1: Adjoining x to F_1 does not change the value group of F_1. That is, $\mathfrak{v}_F(F_1(x)^*) = \mathfrak{v}_F(F_1^*)$.

First, note that since $\overline{F} \subseteq F_1 \subset F_1(x)$, and the residue class field of $\overline{F}$ is already the whole residue class field of F, we have $\overline{F_1(x)} = \overline{F} = \overline{F_1}$.

Since F_1 is relatively algebraically closed, $\widetilde{F_1} = F_1$, where $\widetilde{F_1}$ is the relative algebraic closure of F_1 in F. Moreover, F is Henselian, $\overline{F_1} = \overline{F}$, $char(\overline{F}) = 0$, and trivially $\mathfrak{v}_F(\widetilde{F_1^*}) = \mathfrak{v}_F(F_1^*)$, so we can apply Lemma 4.1.9 to show that $\widetilde{F_1}$ is a Henselization of F_1. In particular, F_1 is already Henselian. The same argument applied to G_1 shows that G_1 is Henselian.

We have established that F_1 and G_1 are Henselian, that x is transcendental over F_1, and that adjoining x does not change the value group or residue class field of F_1. In order to apply Lemma 4.1.10, it remains to find $y \in G$ transcendental over G_1 such that for all $a \in F_1$, $f_1(\mathfrak{v}_F(x-a)) = \mathfrak{v}_G(y - f_1(a))$.

We will find our y by using the fact that G is saturated. That is, we will express the valuation condition required on y as a 1-type, which must be realized in G.

Since $\mathfrak{v}_F(F_1(x))$ is countable, the set $\{\mathfrak{v}_F(x-b) \,|\, b \in F_1\}$ is countable, and thus there is a countable subset $A_1 \subset F_1$ such that for all $b \in F_1$, there exists $a \in A_1$ with $\mathfrak{v}_F(x-a) = \mathfrak{v}_F(x-b)$.

Let $S = f(A_1) \cup \mathfrak{v}_G(G_1) \subset G_1$, and let $\mathcal{L}_S$ be $\mathcal{L}_{VF} \cup \{c_s \,|\, s \in S\}$, the language of valued fields extended with a new constant symbol for each element of $f_1(A_1)$ and for each element of the value group $\mathfrak{v}_G(G_1)$. For all $a \in A_1$, let $\phi_a(v_1)$ be the $\mathcal{L}_S$-formula

$$c_{f_1(\mathfrak{v}(x-a))} = \mathfrak{v}(v_1 - c_{f_1(a)}).$$

Note that $\mathfrak{v}_F(x-a) \in \mathfrak{v}_F(F_1(x)) = \mathfrak{v}_F(F_1)$, so $f_1(\mathfrak{v}(x-a)) \in \mathfrak{v}_G(G_1)$, and $c_{f_1(\mathfrak{v}(x-a))}$ is a constant symbol in $\mathcal{L}_S$.

Let $P = \{\phi_a \,|\, a \in A_1\}$. We would like to show that P is a 1-type over S, so we must show that $P \cup Th_A(G)$ is satisfiable. We will show that every finite subset of $P \cup Th_A(G)$ is satisfiable, then apply Compactness.

Claim: For every finite set $A \subset A_1$, there is $y_A \in G$ such that for all $a \in A$, $f_1(\mathfrak{v}_F(x-a)) = \mathfrak{v}_G(y_A - f_1(a))$.

Choose $b \in A$ such that $w = \mathfrak{v}_F(x-b)$ takes on its maximum value. We have $\mathfrak{v}_F(F_1) = \mathfrak{v}_F(F_1(x))$, so $w \in \mathfrak{v}_F(F_1)$. Now for each positive integer n, we have seen that $\mathfrak{v}_F(n) = 1$, so $\mathfrak{v}_F(nw) = 1 \cdot w = w$. Thus for all n and all $a \in A$,

$$\begin{aligned} \mathfrak{v}_F(b - nw - a) &\geq \min(\mathfrak{v}_F(b-x), \mathfrak{v}_F(nw), \mathfrak{v}_F(x-a)) \\ &\geq \min(w, w, \mathfrak{v}_F(x-a)) \\ &\geq \mathfrak{v}_F(x-a). \end{aligned}$$

Now by Lemma 2.4.6, equality holds above whenever $\mathfrak{v}_F(x-a) < w$. We claim that this is the case for all but at most one n. For suppose we have $m < n$ with $\mathfrak{v}_F(b-mw-a) > \mathfrak{v}_F(x-a)$

and $\mathfrak{v}_F(b-nw-a) > \mathfrak{v}_F(x-a)$. Then

$$\begin{aligned} w &= \mathfrak{v}_F((n-m)w) \\ &\geq \min(\mathfrak{v}_F(b-mw-a), \mathfrak{v}_F(-b+nw+a)) \\ &> \mathfrak{v}_F(x-a), \end{aligned}$$

in which case equality holds above and $\mathfrak{v}_F(b-nw-a) = \mathfrak{v}_F(x-a)$, a contradiction.

Since A is finite, and for each a there is at most one positive integer n such that $\mathfrak{v}_F(b-nw-a) \neq \mathfrak{v}_F(x-a)$, we can choose n such that from all $a \in A$, $\mathfrak{v}_F(b-nw-a) = \mathfrak{v}_F(x-a)$.

Let $y_A = f_1(b-nw)$. Then for all $a \in A$,

$$\begin{aligned} f_1(\mathfrak{v}_F(x-a)) &= f_1(\mathfrak{v}_F(b-nw-a)) \\ &= \mathfrak{v}_G(y_A - f_1(a)), \end{aligned}$$

since $f_1 \circ \mathfrak{v}_F = \mathfrak{v}_G \circ f_1$. This completes the proof of the claim.

Let Δ be any finite subset of $P \cup Th_S(G)$. Let A be the subset of A_1 consisting of all a such that $\phi_a(v_1) \in \Delta$. Applying the claim, there is $y_A \in G$ such that for all $a \in A$, $f_1(\mathfrak{v}_F(x-a)) = \mathfrak{v}_G(y_A - f_1(a))$. That is, $G \models \phi_a(y_A)$ for all $a \in A$. Clearly, G also satisfies all $\mathcal{L}_S$-sentences of $Th_S(G)$ in Δ, so Δ is satisfiable. By Compactness (Theorem 3.1.18), $P \cup Th_S(G)$ is satisfiable.

Hence P is a 1-type over S. Since $f_1(A_1)$ is countable, and $\mathfrak{v}_G(G_1)$ is countable by assumption, S is countable. Now G is $\aleph_1$-saturated, so P is realized by an element $y \in G$.

Now we have that for all $a \in A_1$, $f_1(\mathfrak{v}_F(x-a)) = \mathfrak{v}_G(y - f_1(a))$. We must show that the same is true for all $b \in F_1$.

Let $\mathfrak{v}_F(x-b) = d$ (the valuation of $x-b$ cannot be 0, since then we would have $x = b$, but $x \notin F_1$). Since $\overline{F_1(x)} = \overline{F_1}$, there exists $b' \in F_1$ such that $\overline{b'} = \overline{(x-b)d^{-1}}$, that is, $\mathfrak{v}_F((x-b)d^{-1} - b') > 1$. Multiplying both sides by $d = \mathfrak{v}_F(d)$ (by the cross section property),

$$\begin{aligned} \mathfrak{v}_F((x-b)d^{-1} - b')\mathfrak{v}_F(d) &> d \\ \mathfrak{v}_F(x-b-b'd) &> \mathfrak{v}_F(x-b). \end{aligned}$$

By the definition of A_1, there exists $a \in A_1$ with $\mathfrak{v}_F(x-(b+b'd)) = \mathfrak{v}_F(x-a)$, and thus $\mathfrak{v}_F(x-a) = \mathfrak{v}_F(x-(b+b'd)) > \mathfrak{v}_F(x-b)$. By Lemma 2.4.6 (3) and (4),

$$\begin{aligned} \mathfrak{v}_F(a-b) &= \mathfrak{v}_F((x-b)-(x-a)) \\ &= \min(\mathfrak{v}_F(x-b), \mathfrak{v}_F(-(x-a))) \\ &= \min(\mathfrak{v}_F(x-b), \mathfrak{v}_F(x-a)) \\ &= \mathfrak{v}_F(x-b) \\ &< \mathfrak{v}_F(x-a). \end{aligned}$$

Applying f_1, $\mathfrak{v}_G(f_1(a) - f_1(b)) = f_1(\mathfrak{v}_F(a-b)) < f_1(\mathfrak{v}_F(x-a)) = \mathfrak{v}_G(y - f_1(a))$, since $a \in A_1$. Hence,

$$\begin{aligned}
\mathfrak{v}_G(y - f_1(b)) &= \mathfrak{v}_G((y - f_1(a)) + (f_1(a) - f_1(b))) \\
&= \min(\mathfrak{v}_G(y - f_1(a)), \mathfrak{v}_G(f_1(a) - f_1(b))) \\
&= \mathfrak{v}_G(f_1(a) - f_1(b)) \\
&= f_1(\mathfrak{v}_F(a-b)) \\
&= f_1(\mathfrak{v}_F(x-b),
\end{aligned}$$

as was to be shown.

Finally, we conclude that $y \notin G_1$, for if $y \in G_1$, then $f_1^{-1}(y) \in F_1$, so $0 = \mathfrak{v}_G(y-y) = f_1(\mathfrak{v}_F(x - f_1^{-1}(y)))$, so $x = f_1^{-1}(y)$, and $x \in F_1$, contradicting our choice of x.

We have satisfied all of the hypotheses of Lemma 4.1.10. The lemma tells us that $\mathfrak{v}_G(G_1(y)*) = \mathfrak{v}_G(G_1)$, $\overline{G_1(y)} = \overline{G_1}$, and f_1 can be extended to an isomorphism $g_1 : F_1(x) \cong G_1(y)$.

We have not yet finished satisfying the conditions of (3). In particular, $F_1(x)$ and $G_1(y)$ are not necessarily relatively algebraically closed. By assumption, F_1 is relatively algebraically closed, so by Lemma 4.1.9, $\mathfrak{v}(F_1^*)$ is closed under roots in $\mathfrak{v}(F^*)$. But $\mathfrak{v}(F_1(x)^*) = \mathfrak{v}(F_1^*)$, so $\mathfrak{v}(F_1(x)^*)$ is closed under roots. Again by Lemma 4.1.9, the relative algebraic closure of $F_1(x)$ in F is a Henselization of $F_1(x)$. Call this field F_2. The same argument shows that G_2, the relative algebraic closure of $G_1(x)$ in G, is a Henselization of $G_1(x)$.

By Lemma 4.1.8, g_1 can be extended to an isomorphism $f_2 : F_2 \cong G_2$. By Lemma 4.1.7, $\mathfrak{v}_F(F_2^*) = \mathfrak{v}_F(F_1(x)^*) = \mathfrak{v}_F(F_1^*)$, and $\mathfrak{v}_G(G_2^*) = \mathfrak{v}_G(G_1(y)^*) = \mathfrak{v}_G(G_1^*)$. In particular, $\mathfrak{v}_F(F_2)$ remains countable, and f_2 remains a partial elementary bijection between $\mathfrak{v}_F(F_2^*)$ and $\mathfrak{v}_G(G_2^*)$. This completes the proof of the first special case.

Case 2: The element x is in the value group of F. That is, $x \in \mathfrak{v}(F^*)$.

The function f_1 restricted to $\mathfrak{v}_F(F_1^*)$ is a partial elementary bijection onto $\mathfrak{v}_G(G_1^*)$. Since $\mathfrak{v}_G(G^*)$ is $\aleph_1$-saturated and $\mathfrak{v}_F(F_1^*)$ is countable, we can choose an element $y \in \mathfrak{v}_G(G^*)$ which realizes the complete type of x in $\mathfrak{v}_F(F)$ over $\mathfrak{v}_F(F_1^*)$, where we interpret the constant symbol corresponding to an element of $\mathfrak{v}_F(F_1^*)$ in $\mathfrak{v}_G(G^*)$ by its image under f_1.

Thus, letting V be the subgroup of $\mathfrak{v}_F(F^*)$ generated by $\mathfrak{v}_F(F_1^*)$ and x, and letting W be the subgroup of $\mathfrak{v}_G(G^*)$ generated by $\mathfrak{v}_G(G_1^*)$ and y, the restriction of f_1 extends to a partial elementary bijection between V and W by mapping x to y. Since we have only added one generator to a countable group in each case, V and W are countable.

Define an extension of f_1, $g_1 : F_1(x) \to G_1(y)$, by

$$g_1\left(\frac{d_0 + \ldots + d_m x^m}{e_0 + \ldots + e_n x^n}\right) = \frac{f_1(d_0) + \ldots + f_1(d_m) y^m}{f_1(e_0) + \ldots + f_1(e_n) y^n},$$

with all coefficients d_i and e_j in F_1. Since f_1 is a field isomorphism, it is easy to check that g_1 is a field isomorphism. Checking that it is an isomorphism of valued fields takes a little more work.

Let $p(x) = e_0 + \ldots + e_n x^n$ with coefficients in F_1. Suppose that for some indices r, s with $r < s$, $e_r \neq 0$, and $e_s \neq 0$, we have $\mathfrak{v}_F(e_r x^r) = \mathfrak{v}_F(e_s x^s)$. Then $x^r \mathfrak{v}_F(e_r) = x^s \mathfrak{v}_F(e_s)$, since $\mathfrak{v}_F(x) = x$ by the cross section property, and $x^{s-r} = \mathfrak{v}_F(e_s)(\mathfrak{v}_F(e_r))^{-1} \in \mathfrak{v}_F(F_1)$. But F_1 is

relatively algebraically closed, so by Lemma 4.1.9, $\mathfrak{v}_F(F_1)$ is closed under roots, and thus $x \in \mathfrak{v}_F(F_1) \subseteq F_1$, contradicting our choice of x.

Thus for all distinct nonzero coefficients e_r, e_s, $\mathfrak{v}_F(e_r x^r) \neq \mathfrak{v}_F(e_s x^s)$, and there is a term $e_q x^q$ of least valuation. By Lemma 2.4.6 (4), $\mathfrak{v}_F(p(x)) = \mathfrak{v}_F(e_q)x^q \in V$. Since the valuation of any polynomial is in V, the valuation of any rational function must also be in V, so $\mathfrak{v}_F(F_1(x)^*) = V$. The same argument shows that $\mathfrak{v}_G(G_1(y)^*) = W$. We have established that $F_1(x)$ and $G_1(y)$ are valued subfields of F and G respectively, since $\mathfrak{v}_F(F_1(x)^*) = V \subseteq F_1(x)$ and similarly for $G_1(y)$.

We have $g_1(p(x)) = f_1(e_0) + \ldots + f_1(e_n)y^n$, and the same argument as above shows that $\mathfrak{v}_G(f_1(e_0) + \ldots + f_1(e_n)y^n) = \mathfrak{v}_G(f_1(e_q))y^q$. Further, since $f_1 \circ \mathfrak{v}_F = \mathfrak{v}_G \circ f_1$,

$$\begin{aligned} g_1(\mathfrak{v}_F(p(x))) &= g_1(\mathfrak{v}_F(e_q)x^q) \\ &= f_1(\mathfrak{v}_F(e_q))y^q \\ &= \mathfrak{v}_G(f_1(e_q))y^q \\ &= \mathfrak{v}_G(g_1(p(x))), \end{aligned}$$

so $g_1 \circ \mathfrak{v}_F = \mathfrak{v}_G \circ g_1$, and g_1 is an isomorphism of valued fields.

We will now establish the conditions of (3) by first passing to Henselizations, then closing the value groups under roots, and finally taking relative algebraic closures, all aided by the lemmas of Section 4.1.

By Lemma 4.1.7, $F_1(x)$ and $G_1(y)$ have Henselizations F_3 and G_3. The fields F and G are Henselian, so by the definition of Henselization, F_3 and G_3 embed as valued subfields of F and G respectively. Since Henselizations are unique up to isomorphism, and $F_1(x) \cong G_1(y)$, there is an isomorphism of valued fields $g_3 : F_3 \cong G_3$. The lemma also tells us that $\mathfrak{v}_F(F_3^*) = V$ and $\mathfrak{v}_G(G_3^*) = W$.

Let $\widetilde{V}$ and $\widetilde{W}$ be the closures under roots of V and W in $\mathfrak{v}_F(F)$ and $\mathfrak{v}_G(G)$ respectively. Since V and W are countable, and in closing under roots we add at most one element for each natural number power and each element, $\widetilde{V}$ and $\widetilde{W}$ are countable. Moreover, for every element added to $\widetilde{V}$, there is a corresponding element of $\mathfrak{v}_G(G^*)$ added to $\widetilde{W}$, since $\mathfrak{v}_G(G^*)$ is saturated and g_1 restricted to V is a partial elementary bijection onto W. Thus the restriction of g_1 to $\widetilde{V}$ can be extended to a partial elementary bijection of $\widetilde{V}$ onto $\widetilde{W}$.

Let F_4 and G_4 be the subfields of F and G generated by $F_3 \cup \widetilde{V}$ and $G_3 \cup \widetilde{W}$ respectively. The field F_4 is algebraic over F_3, since every generator of F_4 not in F_4 is the root of some polynomial with coefficients in $\mathfrak{v}_F(F_3^*) \subset F_3$. Now letting $\widetilde{F_3}$ be the relative algebraic closure of F_3 in F, we have $F_4 \subseteq \widetilde{F_3}$, so $\mathfrak{v}_F(F_4^*) \subseteq \mathfrak{v}_F(\widetilde{F_3}^*) = \widetilde{V}$ by Lemma 4.1.9. But $\widetilde{V} \subseteq \mathfrak{v}_F(F_4^*)$, so $\mathfrak{v}_F(F_4^*) = \widetilde{V}$. The same argument shows that $\mathfrak{v}_G(G_4^*) = \widetilde{W}$.

Now the extension of g_1 on V to a partial elementary bijection between $\widetilde{V}$ and $\widetilde{W}$ together with the isomorphism $g_3 : F_3 \cong G_3$ defines a field isomorphism $g_4 : F_4 \cong G_4$. By Lemma 4.1.8, this field isomorphism is also be a valued field isomorphism.

The value groups of F_4 and G_4, $\widetilde{V}$ and $\widetilde{W}$, are closed under roots, so by Lemma 4.1.9, the relative algebraic closures of F_4 and G_4, $\widetilde{F_4}$ and $\widetilde{G_4}$, are Henselizations of F_4 and G_4 respectively. Once again, since Henselizations are unique up to isomorphism, there is an isomorphism of valued fields $f_2 : \widetilde{F_4} \to \widetilde{G_4}$ extending g_4.

Henselizations have the same value groups as their base fields by Lemma 4.1.7, so $\mathfrak{v}_F(\widetilde{F_4}^*) = \widetilde{V}$ and $\mathfrak{v}_G(\widetilde{G_4}^*) = \widetilde{W}$. We have already seen that $\widetilde{V}$ and $\widetilde{W}$ are countable and that the isomorphism restricts to a partial elementary bijection between them. Hence, taking $F_2 = \widetilde{F_4}$ and $G_2 = \widetilde{G_4}$, we have $f_2 : F_2 \leftrightarrow G_2$. This completes the proof of the second special case.

The general case. We have x transcendental over F_1, and we may assume that x is not in the value group of F and that the value group of $F_1(x)$ strictly contains the value group of F_1, since these cases have been dealt with. The idea now is to repeatedly apply the second special case to first adjoin each new element of the value group which would be added upon adjoining x. Then when we adjoin x to the result, no further elements are added to the value group, and we are done by the first special case.

We have already established that $\overline{F_1(x)} = \overline{F_1} = \overline{F}$, and we have dealt with the case in which the value group is trivial, so by Lemma 4.1.11, $\mathfrak{v}_F(F_1(x)^*)$ is countable.

Let $\{x_i \,|\, i \in \mathbb{N}\}$ be an enumeration of the elements of $\mathfrak{v}_F(F_1(x)^*)$ not in $\mathfrak{v}_F(F_1^*)$. Let $F_{x_0} = F_1$ and $G_{x_0} = G_1$. Applying the second special case, for each $i \in \mathbb{N}$, we can extend f_{x_i} to $f_{x_{i+1}} : F_{x_{i+1}} \leftrightarrow G_{x_{i+1}}$, with $x_i \in F_{x_{i+1}}$.

Let $F_2 = \bigcup_{i\in\mathbb{N}} F_{x_i}$, $G_2 = \bigcup_{i\in\mathbb{N}} G_{x_i}$, and $f_2 = \bigcup_{i\in\mathbb{N}} f_{x_i}$. Then $f_2 : F_2 \leftrightarrow G_2$, and $\mathfrak{v}_F(F_1(x)^*) \subseteq \mathfrak{v}_F(F_2^*)$.

But we are not quite done, because adjoining x to F_2 may add elements to the value group of F_2. So we repeat this argument, finding for each $i \geq 2$ an extension of f_{i-1}, $f_i : F_i \leftrightarrow G_i$ such that $\mathfrak{v}_F(F_{i-1}(x)^*) \subseteq \mathfrak{v}_F(F_i^*)$.

Let $F_\omega = \bigcup_{i\geq 1} F_i$, $G_\omega = \bigcup_{i\geq 1} G_i$, and $f_\omega = \bigcup_{i\geq 1} f_i$. Then $f_\omega : F_\omega \leftrightarrow G_\omega$. Consider $F_\omega(x)$. For element $x' \in F_\omega(x)$, $x' \in F_i(x)$ for some i, and thus $x' \in F_{i+1} \subseteq F_\omega$. So adjoining x to F_ω does not add any elements to the value group, and we have reduced to the first special case.

All arguments given above hold with the roles of F and G reversed, so we have also established (4).

(5) *The back-and-forth argument.* Let $(a_\alpha : \alpha < \aleph_1)$ and $(b_\alpha : \alpha < \aleph_1)$ be enumerations of F and G respectively. We start with the isomorphism $f_0 : \overline{F} \leftrightarrow \overline{G}$ established in (2) and inductively build a chain of isomorphisms $(f_\alpha : \alpha < \aleph_1)$ such that for each α, a_α is in the domain of $f_{\alpha+1}$ (using (3)) and b_α is in the range of $f_{\alpha+1}$ (using (4)). By the familiar back-and-forth argument, $f = \bigcup_{\alpha<\aleph_1} f_\alpha$ is an isomorphism $F \cong G$. This completes the proof. □

Now that the heavy lifting is done, what remains is putting together the pieces.

Theorem 4.2.3 (Ax-Kochen Principle)**.** *Let ϕ be an $\mathcal{L}_{VF}$-sentence. Then $\mathbb{Q}_p \models \phi$ for all but finitely many primes p if and only if $\mathbb{F}_p((t)) \models \phi$ for all but finitely many primes p.*

Proof. The class of Henselian valued fields is elementary by Lemma 4.2.1 and therefore closed under ultraproduct by Corollary 3.2.11. For all p, $\mathbb{Q}_p$ and $\mathbb{F}_p((t))$ are complete discrete valued fields, so by Theorem 4.1.4 they are Henselian valued fields. Thus for any ultrafilter $\mathcal{D}$ on the set of primes, $\prod \mathbb{Q}_p/\mathcal{D}$ and $\prod \mathbb{F}_p((t))/\mathcal{D}$ are Henselian valued fields.

In Lemma 3.2.13, we saw that for any nonprincipal ultrafilter $\mathcal{D}$, the residue fields $\overline{\prod \mathbb{Q}_p/\mathcal{D}}$ and $\overline{\prod \mathbb{F}_p((t))/\mathcal{D}}$ have characteristic zero. We also showed that $\mathfrak{v}(\prod \mathbb{Q}_p/\mathcal{D}) \cong \prod \mathbb{Z}/\mathcal{D} \cong$

$\mathfrak{v}(\prod \mathbb{F}_p((t))/\mathcal{D})$ and $\overline{\prod \mathbb{Q}_p/\mathcal{D}} \cong \prod \mathbb{F}_p/\mathcal{D} \cong \overline{\prod \mathbb{F}_p((t))/\mathcal{D}}$. By Theorem 3.1.30, isomorphism implies elementary equivalence, so these Henselian valued fields have elementarily equivalent value groups and residue class fields. These are the conditions of Theorem 4.2.2, so $\prod \mathbb{Q}_p/\mathcal{D} \equiv \prod \mathbb{F}_p((t))/\mathcal{D}$.

Since the elementary equivalence holds for any nonprincipal ultrafilter, applying Theorem 3.2.14 completes the proof. □

4.3. **The Ax-Kochen Theorem.** The Ax-Kochen Principle can be used to prove a whole family of theorems about the p-adic fields, but its most famous application is the Ax-Kochen Theorem, which addresses Artin's conjecture that $\mathbb{Q}_p$ is C_2 for all primes p. After traveling far afield, we finally return to nontrivial zeros of homogeneous polynomials.

An important subtlety to the Ax-Kochen Theorem arises from the fact that the property C_2 cannot be expressed as a first-order $\mathcal{L}_{VF}$-sentence, since we cannot quantify over polynomials of all degrees. Thus, we cannot apply the Ax-Kochen Principle to prove that $\mathbb{Q}_p$ is C_2 for all but finitely many p.

However, when we restrict our attention to polynomials of a fixed degree, we can express a property which is equivalent to $C_2(d)$ as an $\mathcal{L}_{VF}$-sentence, and we can apply the Ax-Kochen Principle to prove that $\mathbb{Q}_p$ is $C_2(d)$ for all but finitely many p. Note that the finite set of exceptional primes may be different for each degree d.

Theorem 4.3.1 (Ax-Kochen Theorem)**.** *For all degrees $d > 0$, there exists a finite set of primes $P(d)$ such that for all $p \notin P(d)$, if f is a homogeneous polynomial over $\mathbb{Q}_p$ of degree d in n variables such that $n > d^2$, then f has a nontrivial zero in $\mathbb{Q}_p^n$.*

Proof. Let $C_2(d)$ be the property that every homogeneous polynomial of degree d in n variables such that $n > d^2$ has a nontrivial zero.

We first show that $C_2(d)$ is equivalent to the property that every homogeneous polynomial of degree d in d^2+1 variables has a nontrivial zero. We will call this property ϕ_d. Clearly $C_2(d)$ implies ϕ_d. Conversely, if $f(x_1, \ldots, x_n)$ is a homogeneous polynomial of degree d with $n > d^2$, then setting the extra variables to 0, $g(x_1, \ldots, x_{d^2+1}) = f(x_1, \ldots, x_{d^2+1}, 0, \ldots, 0)$ is either the zero polynomial or a homogeneous polynomial of degree d in d^2+1 variables. In the first case, any nontrivial choice of values for the $x_1, \ldots, x_{d^2+1}$ is a nontrivial zero of f. In the second case, if ϕ_d holds, then g has a nontrivial zero $(\alpha_1, \ldots, \alpha_{d^2+1})$, so $(\alpha_1, \ldots, \alpha_{d^2+1}, 0, \ldots, 0)$ is a nontrivial zero of f.

We would like to express the property ϕ_d as an $\mathcal{L}_{VF}$-sentence in order to apply the Ax-Kochen Principle. To do this, we need to quantify over all possible homogeneous polynomials of degree d in $d^2 + 1$ variables.

Each monomial of such a polynomial has degree d, so it is a choice of $d^2 + 1$ exponents $n_1, \ldots, n_{d^2+1} \in \mathbb{N}$ for the $d^2 + 1$ variables, such that $\sum_{i=1}^{d^2+1} n_i = d$. Letting $\theta(d)$ be the total number of such choices, we can enumerate all possible monomials as $m_1, \ldots, m_{\theta(d)}$. Then each homogeneous polynomial of degree d in $d^2 + 1$ variables is uniquely determined by a choice of $\theta(d)$ coefficients $a_1, \ldots, a_{\theta(d)}$, one for each m_i, such that at least one of the coefficients is nonzero.

If m_i is the monomial $x_1^{n_1} \dots x_{d^2+1}^{n_{d^2+1}}$, then we define the $\mathcal{L}_{VF}$-term $m_i(x_1, \dots, x_{d^2+1})$ in free variables $(x_1, \dots, x_{d^2+1})$ to be

$$\underbrace{x_1 \cdot \ldots \cdot x_1}_{n_1 \text{ times}} \cdot \ldots \cdot \underbrace{x_{d^2+1} \cdot \ldots \cdot x_{d^2+1}}_{n_{d^2+1} \text{ times}}.$$

We will use the variables a_i and x_i for clarity. Formally, they are choices of v_j from our infinite set of variables $\mathcal{V} = \{v_1, v_2, \dots\}$. We can express the property ϕ_d with the following $\mathcal{L}_{VF}$-sentence:

$$\begin{aligned} \forall a_1 \dots \forall a_{\theta(d)} \quad & \neg((a_1 = 0) \wedge \dots \wedge (a_{\theta(d)} = 0)) \rightarrow \\ (\exists x_1 \dots \exists x_{d^2+1} \quad & \neg((x_1 = 0) \wedge \dots \wedge (x_{d^2+1} = 0)) \wedge \\ & (a_1 \cdot m_1(x_1, \dots, x_{d^2+1}) + \dots \\ & + a_{\theta(d)} \cdot m_{\theta(d)}(x_1, \dots, x_{d^2+1}) = 0)). \end{aligned}$$

Now for a valued field F, $F \models \phi_d$ if and only if F has the property $C_2(d)$. By Theorem 2.4.16, $\mathbb{F}_p((t))$ is C_2, and therefore has the property $C_2(d)$, for all primes p. Thus for all primes p, $\mathbb{F}_p((t)) \models \phi_d$.

Applying the Ax-Kochen Principle, $\mathbb{Q}_p \models \phi_d$ for all but finitely many p, and thus $\mathbb{Q}_p$ has the property $C_2(d)$ for all but finitely many p. Let $P(d)$ be this finite exceptional set.

Then for all $p \notin P(d)$, $C_2(d)$ says that if f is a homogeneous polynomial over $\mathbb{Q}_p$ of degree d in n variables such that $n > d^2$, then f has a nontrivial zero in $\mathbb{Q}_p^n$. □

Appendix A. Ordinals, Cardinals, and Transfinite Induction

This appendix gives a very brief and relatively informal overview of the transfinite numbers. The interested reader is encouraged to find a more thorough development, for instance in Jech's *Set Theory* [Jec03].

There are two types of transfinite numbers, ordinals and cardinals. Intuitively, ordinals generalize ordered numbers ("first", "second", "third"), while cardinals generalize amount, ("one", "two", "three"). Since the standard set theoretic construction defines cardinals as special types of ordinals, we will take up ordinals first.

Ordinals. Ordinals represent order relations which are linear and well founded; that is, there a least element, and every element has a unique element immediately following it in the order. In this way, they generalize the order properties of sets of natural numbers, and, as we will see, provide a structure upon which induction makes sense.

We will begin with an informal description of ordinals, and then present the set theoretic construction. We start with a canonical least ordinal, 0, which represents the ordering on the empty set. Aside from 0, there are two types of ordinals, successor ordinals and limit ordinals.

Given an ordinal α, there is a successor ordinal $\alpha + 1$ which represents the ordering of α with an additional element appended which is greater than all the elements in the ordering α.

Given an infinite set of ordinals, C, there is a limit ordinal which represents the ordering on all elements in all the orderings in C. The first limit ordinal (also the first infinite ordinal) is ω, which is the limit of the ordinals $\{0, 1, 2, \ldots\}$ and represents the ordering on the set of all natural numbers. Note that since the successor of a finite ordinal is still an ordering on finitely many elements, we cannot arrive at ω through the successor operation by appending elements one by one, only by the limit construction.

The table below demonstrates the order relations represented by a few ordinals. The circles are ordered left to right. The successor operation is indicated by adding a circle on the right, and the limit operation is represented by $(\ldots)$. The ordinal $\omega 2$ is the limit of the ordinals $\{0, 1, \ldots, \omega, \omega + 1, \ldots\}$.

0	
1	$\circ$
2	$\circ\,\circ$
$\vdots$	$\vdots$
ω	$\circ\,\circ\,\circ\,\circ\ldots$
$\omega + 1$	$\circ\,\circ\,\circ\,\circ\ldots\circ$
$\omega + 2$	$\circ\,\circ\,\circ\,\circ\ldots\circ\,\circ$
$\vdots$	$\vdots$
$\omega 2$	$\circ\,\circ\,\circ\,\circ\ldots\circ\,\circ\,\circ\,\circ\ldots$
$\vdots$	$\vdots$

Ordinals are quite useful for indexing infinite collections and performing induction in infinite settings. For example, if C is a chain of sets with a least element and order relation (defined by inclusion) corresponding to the ordinal β, we can index the elements of C by $(C_\alpha : \alpha < \beta)$.

If there is a proposition P_α for each ordinal α (collections of propositions like this often correspond to collections of objects indexed by ordinals), then we can prove that P_α is true for all α by a method similar to induction on the natural numbers. The main difference is that we must also deal with the limit case.

Theorem A.1 (Transfinite Induction, [Mar02, Theorem A.8])**.** *Suppose that P_α is a proposition for each ordinal α. Suppose that*

(1) *P_0 is true,*
(2) *if P_α is true, then $P_{\alpha+1}$ is true, and*
(3) *if α is a limit ordinal and P_β is true for all $\beta < \alpha$, then P_α is true.*

Then P_α is true for all ordinals α.

Examples of transfinite induction in this thesis can be found in the proofs of theorems requiring back and forth arguments, most explicitly in Theorem 3.3.6.

There is a very elegant set theoretic construction of the ordinals. Since the relation $\in$ is the primitive binary relation of set theory, we will construct our ordinals so that they are

ordered by $\in$. In particular, we will define an ordinal to be the set containing all ordinals less than it.

We define $0 = \emptyset$, since there are no ordinals less than 0.

Given an ordinal α, we define $\alpha + 1 = \alpha \cup \{\alpha\}$. Then $\alpha + 1$ is the set containing all the elements of α (all ordinals less than α) and α itself.

Given a set of ordinals C, we define the limit of C by $\delta = \bigcup_{\alpha \in C} \alpha$. Suppose that C is unbounded above, that is, for each $\alpha \in C$ there is a $\beta \in C$ such that $\alpha < \beta$. Then $\alpha \in \beta$, and hence $\alpha \in \delta$, so $\alpha < \delta$, and δ is greater than every element of C.

The table below demonstrates the set theoretic representations of a few ordinals.

$$\begin{array}{rl}
0 & \emptyset \\
1 & \{0\} = \{\emptyset\} \\
2 & \{0, 1\} = \{\emptyset, \{\emptyset\}\} \\
\vdots & \vdots \\
\omega & \{0, 1, 2, \ldots\} = \{\emptyset, \{\emptyset\}, \{\emptyset\{\emptyset\}\}, \ldots\} \\
\omega + 1 & \{0, 1, 2, \ldots, \omega\} = \{\emptyset, \{\emptyset\}, \{\emptyset\{\emptyset\}\}, \ldots, \{\emptyset, \{\emptyset\}, \{\emptyset\{\emptyset\}\}, \ldots\}\} \\
\vdots & \vdots
\end{array}$$

By repeatedly taking limits, we can construct larger and larger ordinals. We can construct $\omega 3$ as the limit of $\{\omega 2, \omega 2 + 1, \ldots\}$. The limit of $\{\omega, \omega 2, \omega 3, \ldots\}$ is $\omega\omega = \omega^2$. The limit of $\{\omega^2, \omega^2 2, \omega^2 3, \ldots\}$ is $\omega^2\omega = \omega^3$, and the limit of $\{\omega, \omega^2, \omega^3, \ldots\}$ is ω^ω. Continuing in this way, we can construct larger towers ω^{ω^ω}, $\omega^{\omega^{\omega^\omega}}$, and so forth. The limit of all these towers is yet another ordinal.

However, all the ordinals we have discussed so far are still relatively small. To say what we mean by small, we must introduce the notion of cardinality.

Cardinals. We say that two sets have the same cardinality if there is a bijection between them. Finite sets with different numbers of elements clearly have distinct cardinalities, since their elements cannot be put into 1-1 correspondence. With his famous diagonalization argument, Cantor showed that infinite sets can also have distinct cardinalities.

Formally, we define the cardinality of a set A to be the least ordinal α such that A can be put into bijection with α, and we denote this ordinal by $|A|$. A cardinal is an ordinal which is the cardinality of some set.

Note that as a consequence of this definition, we can describe the cardinals as those ordinals which cannot be put into bijection with any ordinals less than themselves, since for any such ordinal α, $|\alpha| = \alpha$.

All finite ordinals are cardinals. The first infinite cardinal is ω. When we are working with ω as a cardinal, we will denote it by $\aleph_0$. This is the cardinality of the set of natural numbers. If a set A has cardinality $\aleph_0$, we say that A is countable, since A can be "counted", that is, put into bijection with $\mathbb{N}$.

Cantor also showed that any countable union of countable sets is countable. All the ordinals described above can be constructed as limits of countable sequences of ordinals, so they are all countable.

There is a simple construction of the first uncountable ordinal. Take C to be the set of all countable ordinals. The limit of C is $\aleph_1 = \bigcup_{\alpha \in C} \alpha$. The limit $\aleph_1$ is strictly greater than every countable ordinal, so it must be uncountable. Moreover, it contains only countable ordinals, so it is the least uncountable ordinal, and thus is a cardinal, the smallest cardinal greater than $\aleph_0$.

Repeating this argument, we can construct the next cardinal $\aleph_2$ by taking the limit of all ordinals of cardinality $\aleph_1$. The limit of the cardinals $\{\aleph_0, \aleph_1, \aleph_2, \ldots\}$ is the limit cardinal $\aleph_\omega$, and further limits produce greater cardinals, indexed by the ordinals. If $\kappa = \aleph_\alpha$ for some ordinal α, we denote by κ^+ the next cardinal, $\aleph_{\alpha+1}$.

We can define addition, multiplication, and exponentiation of cardinals. For κ and λ cardinals and A and B disjoint sets with $|A| = \kappa$ and $|B| = \lambda$, we define $\kappa + \lambda = |A \cup B|$, the cardinality of the union of A and B, $\kappa\lambda = |A \times B|$, the cardinality of the cartesian product of A and B, and $\kappa^\lambda = |A^B|$, the cardinality of the set of functions from B to A.

The following facts are useful for determining the cardinalities of sets.

Theorem A.2 (Cardinal Arithmetic)**.** *Let κ and λ be cardinals. If both κ and λ are finite, then addition, multiplication, and exponentiation agree with the usual arithmetic of natural numbers. Otherwise,*

(1) $\kappa + \lambda = \kappa\lambda = \max(\kappa, \lambda)$,
(2) *if λ is infinite and $\kappa \leq \lambda$, then $\kappa^\lambda = 2^\lambda$, and*
(3) *if λ is finite and κ is infinite, then $\kappa^\lambda = \kappa$.*

Theorem A.3. *Let $(A_\alpha : \alpha < \beta)$ be a chain of sets indexed by the ordinal β, where $A_\alpha \subseteq A_{\alpha'}$ if $\alpha < \alpha'$. For all $\alpha < \beta$, let $\kappa_\alpha = |A_\alpha|$. Then if $A = \bigcup_{\alpha<\beta} A_\alpha$, $|A| = \bigcup_{\alpha<\beta} \kappa_\alpha$, the limit of the cardinals κ_α.*

The diagonalization argument provides a different way of constructing distinct infinite cardinals. Cantor showed that for any set A, its power set $\mathcal{P}(A)$ has strictly greater cardinality. The cardinality of $\mathcal{P}(A)$ is $2^{|A|}$, since the elements of the power set are in bijection with the functions $A \to \{0, 1\}$. A subset $B \subseteq A$ corresponds to the function f_B defined by $f_B(a) = 1$ if $a \in B$ and $f_B(a) = 0$ if $a \notin B$.

As a consequence of Cantor's Theorem, the sequence $\aleph_0, 2^{\aleph_0}, 2^{2^{\aleph_0}}, \ldots$ is an increasing sequence of distinct cardinals. The terms of this sequence are sometimes denoted $\beth_0, \beth_1, \beth_2, \ldots$, and by taking limits, $\beth_\alpha$ may be defined for any ordinal α. For all ordinals α, $\aleph_\alpha \leq \beth_\alpha$, but the question of whether the sequences $\aleph_0, \aleph_1, \ldots$ and $\beth_0, \beth_1, \ldots$ differ is independent from the usual axioms of set theory.

Continuum Hypothesis. *There are no cardinals between $\aleph_0$ and $2^{\aleph_0}$; that is, $\aleph_1 = 2^{\aleph_0}$.*

Generalized Continuum Hypothesis. *For all ordinals α, there are no cardinals between $\aleph_\alpha$ and $2^{\aleph_\alpha}$; that is, $\aleph_{\alpha+1} = 2^{\aleph_\alpha}$.*

Appendix B. Special Models

In our proof of the Ax-Kochen Theorem, we assumed the Continuum Hypothesis in order to use Theorem 3.3.10, that all complete theories have saturated models. The Continuum Hypothesis can be eliminated from the proof by replacing saturated models with special models.

Definition B.1. Let T be a complete theory with infinite models in a countable language $\mathcal{L}$. A model $\mathcal{M} \models T$ with domain M is called special if it is the union of an elementary chain of models $(\mathcal{M}_\beta : \beta < |M|, \beta \text{ an infinite cardinal})$ such that each $\mathcal{M}_\beta$ is β^+-saturated. The elementary chain is called a specializing chain of $\mathcal{M}$.

Note that in the definition, nothing is required about the cardinalities of the $\mathcal{M}_\beta$.

The analogue of Corollary 3.3.7 also holds for special models.

Theorem B.2 ([CK73, Theorem 5.1.17])**.** *If $\mathcal{M}$ and $\mathcal{N}$ are special models of a complete theory T of the same cardinality $\kappa > \aleph_0$, then $\mathcal{M} \cong \mathcal{N}$.*

The idea of the proof is to use a back and forth argument, where at each stage partial elementary bijections are constructed using Theorem 3.3.6 between subsets of the β^+-saturated submodels of $\mathcal{M}$ and $\mathcal{N}$.

The advantage of special models is that we can show that all complete theories have special models without appealing to the Continuum Hypothesis.

Theorem B.3 ([CK73, Proposition 5.1.8])**.** *For any $\mathcal{L}$-structure $\mathcal{M}$, there is a special elementary extension of $\mathcal{M}$.*

The idea of the proof is to construct a chain of κ-saturated models for increasing cardinalities κ, then take limits.

The proof of Theorem 4.2.2 can be altered to reduce to the special case instead of the saturated case. Unfortunately, this complicates the argument significantly, since the back and forth argument must take the specializing chains into account. Additionally, the special models guaranteed by Theorem B.3 may have cardinality larger than $\aleph_1$, so more complicated cardinality and enumeration arguments are required.

Appendix C. The Resultant

Definition C.1. Let R be a ring. The resultant, $Res : R[x] \times R[x] \to R$, is the function which maps two polynomials over R, $f = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_0$ and $g = b_m x^m + b_{m-1}x^{m-1} + \ldots + b_0$ of degrees n and m respectively, to the determinant of the following

$(m+n)\times(m+n)$ matrix:

$$\begin{array}{c} \\ m \\ \\ \\ \\ \\ n \\ \\ \\ \\ \\ \end{array}\!\!\left\{\begin{array}{c} \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \end{array}\right.\!\!\underbrace{\left(\begin{array}{ccccccc} a_n & \ldots\ldots & a_0 & 0 & \ldots\ldots & 0 \\ 0 & a_n & \ldots\ldots & a_0 & 0 & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ldots\ldots & 0 & a_n & \ldots\ldots & a_0 \\ b_m & \ldots\ldots\ldots\ldots & b_0 & 0 & \ldots & 0 \\ 0 & b_m & \ldots\ldots\ldots\ldots & b_0 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & b_m & \ldots\ldots\ldots\ldots & b_0 \end{array}\right)}_{m+n}.$$

The next theorem gives an alternate expression for the resultant. We will only use it for the implication that $Res(f,g)\neq 0$ if f and g are relatively prime.

Theorem C.2 ([Lan02, Proposition 8.3])**.** *Let R be a subring of a field K, and let $f,g\in R[x]$, with $f=a_nx^n+a_{n-1}x^{n-1}+\ldots+a_0$ and $g=b_mx^m+b_{m-1}x^{m-1}+\ldots+b_0$. Then $Res(f,g)=a_n^mb_m^n\prod_{i=1}^n\prod_{j=1}^m(\alpha_i-\beta_j)$, where the α_i and β_j are the roots of f and g in an algebraic closure of K. Thus $Res(f,g)=0$ if and only if f and g have a common root, and if f and g are relatively prime, $Res(f,g)\neq 0$.*

Proof. Consider the linear equations

$$\begin{aligned} x^{m-1}f(x) &= a_nx^{m+n-1}+a_{n-1}x^{m+n-2}+\ldots+a_0x^{m-1}\\ x^{m-2}f(x) &= a_nx^{m+n-2}+a_{n-1}x^{m+n-3}+\ldots+a_0x^{m-2}\\ &\vdots \quad \vdots\\ f(x) &= a_nx^n+a_{n-1}x^{n-1}+\ldots+a_0\\ x^{n-1}g(x) &= b_mx^{m+n-1}+b_{n-1}x^{m+n-2}+\ldots+b_0x^{n-1}\\ x^{n-2}g(x) &= b_mx^{m+n-2}+b_{n-1}x^{m+n-3}+\ldots+b_0x^{n-2}\\ &\vdots \quad \vdots\\ g(x) &= b_mx^m+b_{m-1}x^{m-1}+\ldots+b_0. \end{aligned}$$

Let C be the vector on the left, $(x^{m-1}f(x),x^{m-2}f(x),\ldots,g(x))$. Let $C_1,\ldots,C_{m+n}$ be the vectors of coefficients of x, with powers of x aligned. So $C_1=(a_n,0,\ldots,0,b_m,0,\ldots,0)$, the coefficients of x^{m+n-1}, $C_2=(a_{n-1},a_n,0,\ldots,0,b_{m-1},b_m,0,\ldots,0)$, the coefficients of x^{m+n-2}, etc. Note that these vectors are the columns of the resultant matrix.

The linear equations can then be expressed as $C=C_1x^{n+m-1}+\ldots+C_{m+n}x^0$, and the right side of this equality is just the resultant matrix multiplied by the column vector $(x^{n+m-1},\ldots,x^0)$.

If we replace the $(m+n)^{th}$ column of the resultant matrix with the column vector C, Cramer's rule tells us that

$$\frac{det(C_1, \ldots, C_{m+n-1}, C)}{det(C_1, \ldots, C_{m+n})} = x^0 = 1,$$

since x^0 is the $(m+n)^{th}$ entry of $(x^{n+m-1}, \ldots, x^0)$. By $det(v_1, \ldots, v_n)$, we mean the determinant of the matrix with columns $v_1, \ldots, v_n$.

So $Res(f,g) = det(C_0, \ldots, C_{m+n}) = det(C_0, \ldots, C_{m+n-1}, C)$. Computing this determinant, we find that every term contains a factor of $f(x)$ or $g(x)$ from the column C. Grouping the terms divisible by f and those divisible by g, we find that there are polynomials $p(x), q(x) \in R[x]$ such that $p(x)f(x) + q(x)g(x) = Res(f,g)$.

Suppose that f and g have a common root α in an algebraic closure of K. Substituting α for x in the equation above, we see that $Res(f,g) = 0$.

Now in the algebraic closure, we can factor f and g as $f = a_n \prod_{i=1}^{n}(x - \alpha_i)$ and $g = b_m \prod_{j=1}^{m}(x - \beta_j)$. Comparing the coefficients of powers of x, we obtain the following expressions for the coefficients:

$$\begin{aligned} a_n &= a_n \\ a_{n-1} &= -a_n(\alpha_1 + \ldots + \alpha_n) \\ \vdots &\quad \vdots \\ a_0 &= (-1)^n a_n(\alpha_1 \alpha_2 \ldots \alpha_n) \end{aligned}$$

and similarly for the b_j. In this way, we can view the coefficients a_i and b_j as symmetric polynomials $(-1)^i a_n S_i(\alpha_1, \ldots, \alpha_n)$ and $(-1)^j b_m T_j(\beta_1, \ldots, \beta_m)$, where $deg(S_i) = n - i$ and $deg(T_j) = m - j$.

Now computing the resultant, we see that

$$Res(f,g) = \begin{vmatrix} a_n S_n & \ldots & (-1)^n a_0 S_0 & 0 & \ldots & 0 \\ 0 & a_n S_n & \ldots & (-1)^n a_0 S_0 & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & a_n S_n & \ldots & (-1)^n a_0 S_0 \\ b_m T_m & \ldots & \ldots & (-1)^m b_0 T_0 & \ldots & 0 \\ 0 & b_m T_m & \ldots & \ldots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & b_m T_m & \ldots & (-1)^m b_0 T_0 \end{vmatrix}.$$

The first m rows have a factor of a_n, and the next n rows have a factor of b_m, so factoring them out,

$$Res(f,g) = a_n^m b_m^n \begin{vmatrix} S_n & \dots & (-1)^n S_0 & 0 & \dots & 0 \\ 0 & S_n & \dots & (-1)^n S_0 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & S_n & \dots & (-1)^n S_0 \\ T_m & \dots & & (-1)^m T_0 & 0 & \dots & 0 \\ 0 & T_m & \dots & & (-1)^m T_0 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & T_m & \dots & & (-1)^m T_0 \end{vmatrix}.$$

Computing this determinant as the sum of products of one element from each row and column, we see that as a polynomial in the α_i and β_j, $Res(f,g)$ has degree mn. Terms in the sum of degree mn come from, for instance, picking all of the S_0 and all of the T_m, or picking all of the T_0 and all of the S_n. No terms of greater degree can be produced.

But if $\alpha_i = \beta_j$ for any $0 \leq i \leq n$ and $0 \leq j \leq m$, $Res(f,g) = 0$, so $(\alpha_i - \beta_j)$ divides $Res(f,g)$. Thus $\prod_{i=1}^{n}\prod_{j=1}^{m}(\alpha_i - \beta_j)$ divides $Res(f,g)$, but both are polynomials of degree mn, so they differ only by a constant factor. By plugging in values for the α_i and β_j, it is easy to see that this constant factor is $a_n^m b_m^n$. This completes the proof. □

Our application of the resultant in the proof of Hensel's lemma uses the following result.

Lemma C.3 ([BS66, Ch. 4 Sec. 3 Lemma])**.** *Let R be a subring of a field K, and let $g, h \in R[x]$ with $deg(g) = m$, $deg(h) = n$. If $\rho = Res(g,h) \neq 0$, then for all $l \in R[x]$ such that $deg(l) \leq m+n-1$, there exist $\phi, \psi \in R[x]$ with $deg(\phi) \leq n-1$, $deg(\psi) \leq m-1$ such that $g\phi + h\psi = \rho l$.*

Proof. Let

$$\begin{aligned} g &= g_{m+n-1}x^{m+n-1} + \ldots + g_0, \\ h &= h_{m+n-1}x^{m+n-1} + \ldots + h_0, \\ \phi &= \phi_{m+n-1}x^{m+n-1} + \ldots + \phi_0, \\ \psi &= \psi_{m+n-1}x^{m+n-1} + \ldots + \psi_0, \text{and} \\ l &= l_{m+n-1}x^{m+n-1} + \ldots + l_0, \end{aligned}$$

where we set all excess coefficients to 0. The values of the g_j, h_j, and l_i are given. We must find values for the ϕ_k and ψ_k such that for all $0 \leq i \leq m+n-1$, $\sum_{j+k=i} g_j\phi_k + \sum_{j+k=i} h_j\psi_k = \rho l_i$, that is, $g\phi + h\psi = \rho l$.

This is a system of $m+n$ linear equations in $m+n$ variables, the ϕ_k and ψ_k. The corresponding matrix, M, is the transpose of the resultant matrix for g and h. The determinant of this matrix is $Res(g,h) = \rho \neq 0$, so this system has a solution.

Moreover, according to the cofactor formula for the inverse,

$$M^{-1} = \frac{1}{|M|}C^T = \frac{1}{\rho}C^T,$$

where C is the cofactor matrix of M. Solving $M \begin{pmatrix} \phi_k \\ \psi_k \end{pmatrix} = \rho(l_i)$ for the ϕ_k and ψ_k, we find $\begin{pmatrix} \phi_k \\ \psi_k \end{pmatrix} = M^{-1}\rho \begin{pmatrix} l_i \end{pmatrix} = C^T \begin{pmatrix} l_i \end{pmatrix} \in R^{m+n}$, so all the ϕ_k, ψ_k are elements of R, and thus $\phi, \psi \in R[x]$. □

References

[AK65] James Ax and Simon Kochen. Diophantine problems over local fields. I. *Amer. J. Math.*, 87:605–630, 1965.

[BS66] A. I. Borevich and I. R. Shafarevich. *Number theory*. Translated from the Russian by Newcomb Greenleaf. Pure and Applied Mathematics, Vol. 20. Academic Press, New York, 1966.

[CK73] C. C. Chang and H. J. Keisler. *Model theory*. North-Holland Publishing Co., Amsterdam, 1973. Studies in Logic and the Foundations of Mathematics, Vol. 73.

[Gre69] Marvin J. Greenberg. *Lectures on forms in many variables*. W. A. Benjamin, Inc., New York-Amsterdam, 1969.

[Har77] Robin Hartshorne. *Algebraic geometry*. Springer-Verlag, New York, 1977. Graduate Texts in Mathematics, No. 52.

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

[Lan02] Serge Lang. *Algebra*, volume 211 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, third edition, 2002.

[Mar02] David Marker. *Model theory*, volume 217 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2002. An introduction.

[Rib99] Paulo Ribenboim. *The theory of classical valuations*. Springer Monographs in Mathematics. Springer-Verlag, New York, 1999.