A MODEL-THEORETIC PROOF OF HILBERT'S NULLSTELLENSATZ

NICOLAS FORD

Abstract. The goal of this paper is to present a proof of the Nullstellensatz using tools from a branch of logic called model theory. In doing so, I hope to demonstrate how logical tools can be applied to branches of mathematics outside of logic itself. The reader should be reasonably acquainted with some basic concepts from algebra, but no familiarity with logic is assumed.

Contents

Introduction

Model theory is a branch of logic which studies the relationship between mathematical structures and the formulas that describe them. Model theorists use the logical properties of formal languages to get results about the objects which interpret them. In addition to being a fairly deep subject in its own right, model theory is also frequently applied to other areas of mathematics, most notably set theory, algebra, and analysis.

The goal of this document is to use model-theoretic techniques to prove David Hilbert's Nullstellensatz, a famous theorem in algebraic geometry. Throughout this paper, I assume familiarity with some basic notions from algebra, including fields, algebraic closure, prime and maximal ideals, quotient rings, polynomial rings, and the fraction field of an integral domain. On the other hand, I assume virtually no familiarity with logic or model theory, and I will be developing all of the model-theoretic machinery necessary for the proof in the intervening sections.

The proof presented herein is neither the original one nor the shortest or simplest that has ever been published. Nonetheless, it is my belief that it provides a good example of how the tools of logic may be used to prove theorems in other areas of mathematics, and it is for that reason that I present it here.

Notes on Notation

Before beginning the discussion, a couple of pieces of notation should be established. In this paper, a lower-case letter with a bar over it refers to a finite sequence. The terms in the sequence $\bar{a}$ are called $a_1, a_2, \ldots, a_n$, where n is the number of terms in the sequence. If R is a ring, then $R[\bar{x}]$ is the polynomial ring over R with n variables. Finally, the symbol $\subset$ is used only to refer to a proper subset relation, that is, $A \subset B$ iff $A \subseteq B$ and $A \neq B$.

1. Languages, Structures, and Terms

Unlike some higher-level branches of mathematics, logic is directly concerned with the strings of symbols that represent propositions and predicates. Therefore, in order to speak coherently about the relationships between two different structures, we must be sure that we are using the same symbols to make statements about both structures.

Definition 1.1. A *language* is a collection $\mathscr{L}$ of symbols[1], split into several disjoint subsets: for each $k \in \mathbb{N}$, a set R_k of k-place *relation symbols*; for each $m \in \mathbb{N}$, a set F_m of m-place *function symbols*; and a set C of *constant symbols*.

So, given such a set of symbols, we need to know what it means for two mathematical stuctures to share the same language. The definition is very straightforward.

Definition 1.2. Given a language $\mathscr{L}$, an *$\mathscr{L}$-structure* is a set A, called the *universe*, together with *interpretations* for each of the symbols in $\mathscr{L}$: for each constant symbol $c \in C$, there is a distinct element $c^A \in A$; for each relation symbol $R \in R_k$, there is a distinct subset $R^A \subseteq A^k$; and for each function symbol $f \in F_m$ there is a distinct function $f^A : A^m \to A$.

Example 1.3. Let $\mathscr{L}_R$ be the language with two constant symbols, 0 and 1, one 1-place function symbol, $-$, and two 2-place function symbols, $+$ and $\cdot$. Then, for example, $\mathbb{Z}$ is an $\mathscr{L}_R$-structure with the set $\mathbb{Z}$ itself as the universe, and the obvious interpretations of the symbols; that is, $0^{\mathbb{Z}} = 0$, $1^{\mathbb{Z}} = 1$, $-^{\mathbb{Z}}x = -x$, $x +^{\mathbb{Z}} y = x + y$, and $x \cdot^{\mathbb{Z}} y = xy$. In fact, any ring is an $\mathscr{L}_R$-structure with similar interpretations of the symbols.

Example 1.4. Let $\mathscr{L}_{DG}$ be the language with just one 2-place relation symbol, called $\sim$. Any directed graph G can be made into an $\mathscr{L}_{DG}$-structure by letting the universe be the set of vertices and saying that for any two vertices v and w, $v \sim^G w$ iff there is an edge from v to w.

Example 1.5. If R is a ring, we can construct a language for left R-modules. Let $\mathscr{L}_M(R)$ be the language with one constant symbol, 0, one 2-place function symbol, $+$, and a 1-place function symbol for each $r \in R$. Then if M is a left R-module, it is an $\mathscr{L}_M(R)$-structure in which $0^M = 0$, $a +^M b = a + b$, and $r^M(a) = ra$.

Using the symbols from a language, we can refer to specific elements of the corresponding structure. For example, in the language $\mathscr{L}_R$ from Example 1.3, we

[1]The symbols in a language can be any mathematical object at all. Sometimes we think of the symbols as being identical with the glyphs used to represent them on paper; however, it is frequently useful to use other objects, like the ring elements in Example 1.5.

may refer to the element $-((1+1)+1)$, and this makes sense in any $\mathscr{L}_R$ structure. The string of symbols $-((1+1)+1)$ is called a term of $\mathscr{L}_R$. Any ring R will have an unambiguous interpretation of the term in question, namely $-^R((1+^R 1)+^R 1)$, which is, by definition, some element of R.

Definition 1.6. Consider a language $\mathscr{L}$, together with a new set V of symbols, which will be used to represent variables. The set of *terms of* $\mathscr{L}$ is the smallest set $\mathscr{T}$ such that:

- Every variable symbol $x \in V$ is in $\mathscr{T}$.
- Every constant symbol c of $\mathscr{L}$ is in $\mathscr{T}$.
- For every m-place function symbol f of $\mathscr{L}$, and every $t_1, \ldots, t_m \in \mathscr{T}$, the string $f(t_1, \ldots, t_m)$ is in $\mathscr{T}$.

If a term has no free variables, it is said to be *closed*. If a term t has free variables $x_1, \ldots, x_n$, we will sometimes write $t(x_1, \ldots, x_n)$ or $t(\bar{x})$ to emphasize this fact. If the letter t is written alone, it may or may not be closed.

If $t(\bar{x})$ is a term, A is an $\mathscr{L}$-structure, and $\bar{a}$ is a sequence of elements of A, then $t^A(\bar{a})$ is the element of A that you get by plugging in a_i for x_i. More precisely, we can define it by induction on the length of the term as follows:

- If t is the variable symbol x_i, then $t^A(\bar{a}) = a_i$.
- If t is the constant symbol c, then $t^A(\bar{a}) = c^A$.
- If t is $f(t_1, \ldots, t_m)$ for some function symbol f and terms $t_1(\bar{x}), \ldots, t_m(\bar{x})$, then $t^A(\bar{a}) = f^A(t_1^A(\bar{a}), \ldots, t_m^A(\bar{a}))$.

2. Formulas and Theories

We can use these terms to form more complex strings which make statements about the structures which interpret a language. They will also have interpretations, this time as truth values rather than elements of a structure.

Definition 2.1. Given a language $\mathscr{L}$, an *atomic formula of* $\mathscr{L}$ is defined as follows:

- If t_1 and t_2 are terms, then the string $t_1 = t_2$ is an atomic formula.
- If R is a k-place relation symbol and $\bar{t}$ is a sequence of terms, then the string $R(t_1, \ldots, t_k)$ is an atomic formula.

Given this definition, we say that the set of *formulas of* $\mathscr{L}$ is the smallest set $\mathscr{F}$ such that:

- If ϕ is an atomic formula, then $\phi \in \mathscr{F}$.
- If $\psi \in \mathscr{F}$, then $\neg\psi \in \mathscr{F}$.
- If ψ_1 and ψ_2 are in $\mathscr{F}$, then $\psi_1 \wedge \psi_2$ and $\psi_1 \vee \psi_2$ are in $\mathscr{F}$.
- If $\psi \in \mathscr{F}$ and the variable x is among the free variables of ψ, then $\exists x\, \psi$ and $\forall x\, \psi$ are in $\mathscr{F}$.

A formula without free variables is called a *sentence*, and a set of sentences is called a *theory*. When writing formulas with free variables, we adopt a similar convention to the one we used for terms, that is, $\phi(\bar{x})$ denotes a formula ϕ in which the variables in $\bar{x}$ are free. When writing out formulas, it is common to adopt the following pieces of shorthand:

- $x \neq y$ instead of $\neg(x = y)$
- $\phi \to \phi'$ instead of $(\neg\phi) \vee \phi'$
- $\phi \leftrightarrow \phi'$ instead of $(\phi \to \phi') \wedge (\phi' \to \phi)$
- $\bigwedge_{i=1}^n \phi_i$ instead of $\phi_1 \wedge \cdots \wedge \phi_n$

- $\bigvee_{i=1}^{n} \phi_i$ instead of $\phi_1 \vee \cdots \vee \phi_n$

If $\phi(\bar{v})$ is a formula, A is an $\mathscr{L}$-structure, and $\bar{a}$ is a sequence of elements of A, we define the relation $A \models \phi(\bar{a})$ (read "$\bar{a}$ satisfies ϕ in A") as follows:

- If $\phi(\bar{v})$ is the atomic formula $t_1(\bar{v}) = t_2(\bar{v})$, then $A \models \phi(\bar{a})$ if and only if $t_1^A(\bar{a}) = t_2^A(\bar{a})$.
- If $\phi(\bar{v})$ is the atomic formula $R(t_1(\bar{v}), \ldots, t_k(\bar{v}))$, then $A \models \phi(\bar{a})$ if and only if $(t_1^A(\bar{a}), \ldots, t_k^A(\bar{a})) \in R^A$.
- If $\phi(\bar{v}) = \neg\psi(\bar{v})$, then $A \models \phi(\bar{a})$ if and only if $A \not\models \psi(\bar{a})$.
- If $\phi(\bar{v}) = \psi_1(\bar{v}) \wedge \psi_2(\bar{v})$, then $A \models \phi(\bar{a})$ if and only if $A \models \psi_1(\bar{a})$ and $A \models \psi_2(\bar{a})$.
- If $\phi(\bar{v}) = \psi_1(\bar{v}) \vee \psi_2(\bar{v})$, then $A \models \phi(\bar{a})$ if and only if $A \models \psi_1(\bar{a})$ or $A \models \psi_2(\bar{a})$.
- If $\phi(\bar{v}) = \exists x\, \psi(\bar{v}, x)$, then $A \models \phi(\bar{a})$ if and only if $A \models \psi(\bar{a}, b)$ for some $b \in A$.
- If $\phi(\bar{v}) = \forall x\, \psi(\bar{v}, x)$, then $A \models \phi(\bar{a})$ if and only if $A \models \psi(\bar{a}, b)$ for every $b \in A$.

We adopt a few simplifying conventions for the $\models$ symbol. If ϕ is a sentence, then we omit the parentheses and just write $A \models \phi$ (read "A models ϕ"). If T is a theory, we say that $A \models T$ if $A \models \phi$ for each $\phi \in T$. Finally, if $\phi(\bar{v})$ is a formula, then we say $T \models \phi(\bar{v})$ if for every model A of T, there is a sequence $\bar{a}$ of elements of A such that $A \models \phi(\bar{a})$.

Definition 2.2. If $\phi_1(\bar{v})$ and $\phi_2(\bar{v})$ are formulas of a language $\mathscr{L}$, then we say that ϕ_1 and ϕ_2 are *semantically equivalent* if $\{\phi_1\} \models \phi_2$ and $\{\phi_2\} \models \phi_1$.

The satisfaction relation is the basis for what model theorists mean when they refer to "truth." If ϕ is a formula and A is a structure, the statement $A \models \phi(\bar{a})$ means that the formula ϕ is true of $\bar{a}$ in A. In this way, sentences represent propositions about structures, and formulas with free variables represent predicates.

Example 2.3. Let $\mathscr{L}_{DG}$ be the language of directed graphs described in Example 1.4. Consider the theory T consisting of the following sentences:

- $\forall x \forall y (x \sim y \leftrightarrow y \sim x)$
- $\forall x (\neg(x \sim x))$

Then we can see that for any graph G, $G \models T$ if and only if G is an undirected graph with no self loops.

Example 2.4. Let $\mathscr{L}_R$ be the language of rings described in Example 1.3. One possibility for a theory of rings is the theory which contains the following sentences (writing ab instead of $a \cdot b$):

- $\forall x \forall y \forall z (x + (y + z) = (x + y) + z)$
- $\forall x [(x + 0 = x) \wedge (x + (-x) = 0)]$
- $\forall x \forall y (x + y = y + x)$
- $\forall x \forall y \forall z (x(yz) = (xy)z)$
- $\forall x [(x1 = x) \wedge (1x = x)]$
- $\forall x \forall y \forall z [(x(y + z) = xy + yz) \wedge ((x + y)z = xz + yz)]$

Adding the following sentences would make it into a theory of fields:

- $\forall x \forall y (xy = yx)$
- $\forall x [(x \neq 0) \rightarrow \exists y (xy = 1)]$

For each n, let p_n be the sentence $\forall a_0 \cdots \forall a_n[((a_1 \neq 0) \vee \cdots \vee (a_n \neq 0) \vee (a_0 = 0)) \rightarrow \exists x(a_0 + a_1 x + \cdots + a_n x^n = 0)]$. If we take our theory of fields and add the sentences p_n for each $n > 0$, we get a new theory called ACF. It is easy to see that $F \models ACF$ if and only if F is an algebraically closed field. This theory will figure heavily into our proof of the Nullstellensatz.

When proving facts about formulas, one frequently wants to use some sort of inductive argument. We need some variable on which to perform induction. A natural choice is the length of the formula, meaning the number of symbols (counting each occurrence of a symbol seperately) that comprise it. Using this method, it is only necessary to prove that your statement holds for atomic formulas, and that if your statement holds for ψ_1 and ψ_2, then it holds for $\neg\psi_1$, $\psi_1 \wedge \psi_2$, and $\exists x\, \psi_1$. The cases for $\vee$ and $\forall$ will then be handled automatically, because $(\psi_1 \vee \psi_2)$ is semantically equivalent to $\neg(\neg\psi_1 \wedge \neg\psi_2)$ and $\forall x\, \psi$ is semantically equivalent to $\neg\exists x(\neg\psi)$.

We conclude this section with the following basic fact about formulas, which will be necessary for a small part of the proof of the Nullstellensatz.

Definition 2.5. A formula is said to be *quantifier-free* if it contains neither $\exists$ nor $\forall$.

Proposition 2.6. *For every quantifier-free formula ϕ, there exists a quantifier-free formula ψ of the form $\bigvee_{i=1}^{m} \bigwedge_{j=1}^{n} \theta_{ij}$, with θ_{ij} atomic or negated atomic, such that ϕ is semantically equivalent to ψ. Such a ψ is said to be in* disjunctive normal form*.*

Proof. Let $\Theta = \{\theta_i\}_{i=1}^{m}$ be the set of all atomic formulas which appear in ϕ. If $s : \Theta \rightarrow \{0, 1\}$ is a function assigning truth values (0 is false, 1 is true) to every θ_i, then let S be the set of all such functions s for which, when each θ_i is given the truth value $s(\theta_i)$, ϕ is true. Let $\psi_i^s = \theta_i$ if $s(\theta_i) = 1$ and $\neg\theta_i$ if $s(\theta_i) = 0$. Then ϕ is semantically equivalent to $\bigvee_{s \in S} \bigwedge_{i=1}^{m} \psi_i^s$, which is in disjunctive normal form. □

3. Homomorphisms and Elementary Substructures

Just like homomorphisms of various structures in algebra, there is a notion of a structure homomorphism which preserves the relations and functions encoded by the symbols of the language

Definition 3.1. If A and B are $\mathscr{L}$-structures, then a function $h : A \rightarrow B$ is a *homomorphism* if for every sequence $\bar{a}$ of elements of A, we have (writing $h(\bar{a})$ instead of $(h(a_1), \ldots, h(a_n))$):

- For every constant symbol c in $\mathscr{L}$, $h(c^A) = c^B$.
- For every function symbol f in $\mathscr{L}$, $h(f^A(\bar{a})) = f^B(h(\bar{a}))$
- For every relation symbol R in $\mathscr{L}$, if $\bar{a} \in R^A$, then $h(\bar{a}) \in R^B$

An embedding is an injective homomorphism in which, for $\bar{a} \in A^k$ and R a relation symbol, $\bar{a} \in R^A$ *if and only if* $h(\bar{a}) \in R^B$. If an embedding is surjective, it is called an *isomorphism*.

This notion of homomorphism closely matches the analogous notions from algebra. Homomorphisms in the language of rings from Example 1.3 will be ring homomorphisms. Likewise for homomorphisms in the language of modules from Example 1.5. The definition of substructure is similarly related to the corresponding algebraic concept.

Definition 3.2. If A and B are $\mathscr{L}$-structures with $A \subseteq B$, then we say that A is a *substructure* of B if the inclusion map from A to B is an embedding.

A substructure in the language of rings is a subring, a substructure in the language of directed graphs is a subgraph, and so on, as expected.

It is easy to see that isomorphisms will preserve the truth of all formulas of $\mathscr{L}$. A natural question to ask, then, is whether there are other homomorphisms which have the same property. It is at least necessary that such a homomorphism be an embedding, because if $h(a) = h(b)$ but $a \neq b$, then if $\phi(x, y)$ is the formula $x = y$, we have that $A \not\models \phi(a, b)$ but $B \models \phi(h(a), h(b))$. Similarly, if for some relation symbol R and some $\bar{a}$, $h(\bar{a}) \in R^B$ but $\bar{a} \notin R^B$, we can find a similar formula which is satisfied in one structure but not the other.

It turns out that a homomorphism does not need to be an isomorphism in order to preserve the truth of formulas. Ones which do are the subject of the rest of this section.

Definitions 3.3. An embedding $h : A \to B$ is called *elementary* if, for all $\bar{a} \in A^n$ and all formulas $\phi(\bar{x})$, $A \models \phi(\bar{a})$ if and only if $B \models \phi(h(\bar{a}))$. If $A \subseteq B$ and the inclusion map is an elementary embedding, then we say A is an *elementary substructure* of B.

Let T be a theory. We say that T is *model-complete* if every embedding between models of T is elementary.

We conclude this section by introducing a property which is strictly stronger than model-completeness, and which will be used to prove the model-completeness of the theory of algebraically closed fields.

Definition 3.4. Let T be theory. If for every formula $\phi(\bar{v})$ there is a quantifier-free formula $\phi'(\bar{v})$ such that $T \models (\phi \leftrightarrow \phi')$, we say that T *has quantifier elimination.*

Theorem 3.5. *Let T be a theory. If T has quantifier elimination, then T is model-complete.*

Proof. Suppose $A \models T$ and $B \models T$, and let $h : A \to B$ be an embedding. Let $\phi(\bar{v})$ be a formula. Because T has quantifier elimination, there is some quantifier-free formula $\phi'(\bar{v})$ so that $T \models (\phi \leftrightarrow \phi')$. By definition, this means that, for all models M of T and all sequences $\bar{m}$ of elements of M, $M \models \phi(\bar{m})$ if and only if $M \models \phi'(\bar{m})$. Therefore, it is sufficient to show that h preserves the truth of all quantifier-free formulas.

We proceed by induction. If $\phi(\bar{v})$ is an atomic formula of the form $t_1(\bar{v}) = t_2(\bar{v})$, then the definition of homomorphism together with the injectivity of h gives us that $t_1^A(\bar{a}) = t_2^A(\bar{a})$ if and only if $t_1^B(h(\bar{a})) = t_2^B(h(\bar{a}))$. A similar argument shows that atomic formulas formed from relation symbols are preserved.

Suppose $\phi(\bar{v}) = \neg\psi(\bar{v})$ for some formula ψ whose truth is preserved by h. Then $A \models \phi(\bar{a}) \Leftrightarrow A \not\models \psi(\bar{a}) \Leftrightarrow B \not\models \psi(h(\bar{a})) \Leftrightarrow B \models \phi(h(\bar{a}))$. A similar argument will prove the theorem for formulas of the form $\psi_1 \wedge \psi_2$. □

4. The Model-Completeness of ACF

The proof of the Nullstellensatz will make use of the model-completeness of the theory of algebraically closed fields. While the proof is a bit long, the argument is quite straightforward. (The essential parts of the proof come from lecture notes by Joe Mileti, which appear to no longer be available.)

The following two lemmas will be necessary.

Lemma 4.1. *All algebraically closed fields are infinite.*

Proof. If $K = \{a_1, \ldots, a_m\}$, then $(x - a_1) \cdots (x - a_m) + 1$ has no roots. □

Lemma 4.2. *If K is an algebraically closed field, and $p, q \in K[x]$, then $p|q^n$, where $n = \deg p$, if and only if every root of p is also a root of q.*

Proof. Because K is algebraically closed, p and q split completely, say $p = (x - a_1) \cdots (x - a_k)$ and $q = (x - b_1) \cdots (x - b_m)$. Suppose $p|q^n$ and c is a root of p. Then c is one of the a_i's, so c is also one of the b_i's. Suppose every root of p is also a root of q. Then let c be a root of p and let r be its multiplicity. Then c is a root of q, so $(x - c)|q$, so $(x - c)^r|(x - c)^n|q^n$. □

Theorem 4.3. *The theory of algebraically closed fields has quantifier elimination.*

Proof. Given a formula $\phi(\bar{v})$, we want to find a quantifier-free formula $\phi'(\bar{v})$ for which $ACF \models (\phi(\bar{v}) \leftrightarrow \phi'(\bar{v}))$. We can prove this by induction on the length of the formula. If ϕ is atomic, or if $\phi = \neg\psi_1$ or $\phi = \psi_1 \wedge \psi_2$, with ψ_i equivalent to a quantifier-free formula, the statement is trivial.

So suppose $\phi(\bar{v}) = \exists x\, \psi(\bar{v}, x)$, with ψ equivalent to a quantifier-free formula. Then by Proposition 2.6, ψ may be put into disjunctive normal form. Note that any atomic formula $\theta(\bar{w})$ is equivalent to $p(\bar{w}) = 0$, where p is a polynomial. So, because $ACF \models (\exists x(a \vee b) \leftrightarrow \exists x\, a \vee \exists x\, b)$, it is sufficient to find a quantifier-free equivalent to formulas of the form $\exists x[\bigwedge_{i=1}^{m}(p(\bar{v}, x) = 0) \wedge \bigwedge_{i=1}^{k}(q(\bar{v}, x) \neq 0)]$. Note that if $q = \prod q_i$, then $ACF \models ((\bigwedge_{i=1}^{k}[q_i(\bar{v}, x) \neq 0]) \leftrightarrow q(\bar{v}, x) \neq 0)$, so we may assume that $k = 1$.

We can also reduce m to 1. We may think of $p_i(\bar{v}, x)$ as a polynomial over x whose coefficients are terms involving $\bar{v}$. Say $a_i(\bar{v}) \cdot x^{d_i}$ is the leading term of p_i, and suppose that $1 \leq d_1 \leq d_2$. Then we can find a set of polynomials $\{p_i{}'\}$ which has the same set of simultaneous roots as $\{p_i\}$, but for which $\sum_{i=1}^{m} \deg p_i' < \sum_{i=1}^{m} \deg p_i$. If $a_1(\bar{v}) = 0$, we can set $p_1{}' = p_1 - a_1(\bar{v})x^{d_1}$ and $p_i{}' = p_i$ for $i \neq 1$; otherwise, set $p_2{}'' = a_1(\bar{v})p_2 - a_2(\bar{v})x^{d_2 - d_1}p_1$ and $p_i{}'' = p_i$ for $i \neq 2$. Then $\bigwedge_{i=1}^{m}(p_i(\bar{v}, x) = 0)$ is equivalent to

$$[a_1(\bar{v}) = 0 \wedge \bigwedge_{i=1}^{m}(p_i{}'(\bar{v}, x) = 0)] \vee [a_1(\bar{v}) \neq 0 \wedge \bigwedge_{i=1}^{m}(p_i{}''(\bar{v}, x) = 0)].$$

So, whenever we have at least two polynomials of degree greater than 0 over x, we may reduce the total degree of all of our polynomials. (The $\vee$ allows us to split the existential quantifier like before.) Repeat this process until only one of the polynomials has degree greater than 0 over x. We may pull all of the constant polynomials out of the existential quantfier, because $ACF \models [\exists x(a(\bar{v}) \wedge b(\bar{v}, x)) \leftrightarrow a(\bar{v}) \wedge \exists x\, b(\bar{v}, x)]$.

We now only have to deal with formulas of the form $\exists x(p(\bar{v}, x) = 0 \wedge q(\bar{v}, x) \neq 0)$, in which either p or q might not actually be present. If p is not present, then note that polynomials have finitely many solutions, but algebraically closed fields are infinite, so there exists an x which is not a root of q if and only if $q \not\equiv 0$. If q is not present, then note that p has a solution if and only if all the coefficients of p are zero or at least one of the non-constant coefficients is nonzero.

Otherwise, by Lemma 4.2, there exists such an x if and only if $\neg(p|q^{\deg p})$. And we may express $f|g$ as a quantifier-free formula by use of the Euclidean algorithm. There are some polynomials a and b such that $g = af + b$, and a and b can be found using polynomial long division, so their coefficients will be terms in $\bar{v}$. So $f|g$ if and only if every coefficient of b is zero. □

Corollary 4.4. *The theory of algebraically closed fields is model-complete.*

5. Algebraic Preliminaries

The proof of the Nullstellensatz will require familiarity with two ancillary results from commutative algebra, which we present here.

Proposition 5.1. *If F is a field, then every ideal of $F[\bar{x}]$ is finitely generated.*

Proof. We prove that if every ideal of R is finitely generated, then every ideal of $R[y]$ is finitely generated. Take I an ideal in $R[y]$ and take f_0 of minimal degree in I. Then, for each $i > 0$, take f_i of minimal degree in $I \setminus (f_0, \ldots, f_{i-1})$. Let a_i be the initial coefficient of f_i. Then the ideal in R generated by all of the a_i will be equal to $(a_0, \ldots, a_N)$ for some N. Then we claim that $I = (f_0, \ldots, f_N)$. Suppose not. Then $f_{N+1} \in I \setminus (f_0, \ldots, f_N)$. But there exist λ_i such that $a_{N+1} = \sum_{i=0}^{N} \lambda_i a_i$. Then consider $g = \sum_{i=0}^{N} \lambda_i f_i x^{k_i}$, where $k_i = \deg(f_{N+1}) - \deg(f_i)$. Then $f_{N+1} - g \in I \setminus (f_0, \ldots, f_N)$, and it has smaller degree than f_{N+1}, contradicting the minimality assumption. □

Definition 5.2. If I is an ideal in a ring R, then the *radical of I* is the set $\sqrt{I} = \{r \in R : r^n \in I \text{ for some } n\}$. If $I = \sqrt{I}$, then we say that I is a *radical ideal*.

Lemma 5.3. *Let I be a radical ideal in an integral domain R. Then for every $x \notin I$, there is a prime ideal $P \subseteq R$ such that $I \subseteq P$ and $x \notin I$.*

Proof. Take some $x \notin I$. Let $X = \{x^n : n \in \mathbb{N}\}$. Then if K is the fraction field of R, consider the ring $X^{-1}R = \{r/s \in K : r \in R,\ s \in X\}$, and define $X^{-1}I \subseteq X^{-1}R$ similarly. Let ϕ be the natural map from R to $(X^{-1}R/X^{-1}I)$. Because $x \notin I$, we know that $x^n \notin I$ for any n. So $1 = x^n/x^n \notin X^{-1}I$ for any n, so $X^{-1}I$ is a proper ideal of $X^{-1}R$. Therefore, we may take a maximal ideal P in $(X^{-1}R/X^{-1}I)$. The preimage $\phi^{-1}(P)$ is prime and contains I. Now, let $\psi : X^{-1}R \to (X^{-1}R/X^{-1}I)$ be the projection map. Suppose $x \in \phi^{-1}(P) \subseteq \psi^{-1}(P)$. Then $1 = x/x \in \psi^{-1}(P)$, meaning that $\psi^{-1}(P)$ is not a proper ideal of $X^{-1}R$, contradicting the surjectivity of ψ. So $x \notin \phi^{-1}(P)$. □

6. The Nullstellensatz

Now that the model-theoretic preliminaries have been established, we are ready to prove the Nullstellensatz itself. The proof is surprisingly short.

Theorem 6.1 (Nullstellensatz)**.** *For $S \subseteq K[\bar{x}]$, let $V(S) = \{\bar{a} \in K^n : f(\bar{a}) = 0 \text{ for all } f \in S\}$. Then if I and J are radical ideals in $K[\bar{x}]$ with $I \subset J$, then $V(J) \subset V(I)$.*

Proof. Clearly $V(J) \subseteq V(I)$, so we need to show that $V(J) \neq V(I)$. Take a polynomial $f \in J \setminus I$. By Lemma 5.3, there exists some prime ideal P such that $I \subseteq P$ and $f \notin P$. Let L be the algebraic closure of the fraction field of $K[\bar{x}]/P$,

and let ϕ be the natural map from $K[\bar{x}]$ into L. Then, letting $\bar{y} = \phi(\bar{x})$, observe that $f(\bar{y}) \neq 0$ and for all $g \in I$, $g(\bar{y}) = 0$.

By Proposition 5.1, I is finitely generated, so let $h_1, \ldots, h_m$ be its generators. Then, if $\psi(\bar{w})$ is the formula $(h_1(\bar{w}) = 0 \wedge \cdots \wedge h_m(\bar{w}) = 0 \wedge f(\bar{w}) \neq 0)$, we see that $\bar{y}$ satisfies ψ in L. Therefore, $L \models \exists \bar{w}\, \psi(\bar{w})$, so, because the theory of algebraically closed fields is model-complete, $K \models \exists \bar{w}\, \psi(\bar{w})$. Let $\bar{a}$ be an element of K^n which satisfies ψ. Then we see that $\bar{a} \in V(I) \setminus V(J)$, so $V(J) \neq V(I)$. □

Though the result just proved is frequently called the Nullstellensatz, many algebraists prefer an alternative version of the statement, whose proof we present here.

Theorem 6.2. *The only maximal ideals of $K[x_1, \ldots, x_n]$ are those of the form $(x_1 - a_1, \ldots, x_n - a_n)$ for $a_1, \ldots, a_n \in K$.*

Proof. $K[\bar{x}]$ is itself a radical ideal, so for any proper radical ideal $I \subset K[\bar{x}]$, $V(I)$ is nonempty. In particular, because every maximal ideal is radical, $V(M)$ is nonempty for every maximal ideal M.

Let M be a maximal ideal. We claim that $V(M)$ has exactly one element. To see this, suppose not, and take $\bar{a} \neq \bar{b}$ in M. Let $I = \{f : f(\bar{a}) = 0\}$. Observe that there is a polynomial $p \in K[\bar{x}]$ such that $p(\bar{a}) = 0$ and $p(\bar{b}) \neq 0$: if $a_i \neq b_i$, let $p(\bar{x}) = x_i - a_i$. Therefore, $M \subset I$. But $I \neq A$ (because, for instance, $1 \notin I$), which is a contradiction.

Let $\bar{a}$ be the unique element of $V(M)$. By the preceding argument, $M = \{f : f(\bar{a}) = 0\}$, so each polynomial $p_i(\bar{x}) = x_i - a_i$ is in M. Thus the ideal $P = (p_1, \ldots, p_n)$ is contained in M. Consider the projection map $\pi : K[\bar{x}] \to K[\bar{x}]/P$. For each i, $\pi(x_i) = \pi(a_i)$, so if $f(\bar{a}) = 0$, then $\pi(f) = 0$. Therefore, $M \subseteq P$. The result follows. □

References

[1] C. C. Chang, H. Jerome Keisler. Model Theory. Elsevier. 1990.
[2] Wilfrid Hodges. A Shorter Model Theory. Cambridge UP. 1997.
[3] David Marker. Model Theory: An Introduction. Springer. 2002.