

DEFINABLE SETS OVER FINITE FIELDS

by

Kadir Güray Güner

B.S., Mathematics, Boğaziçi University, 2010

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Master of Science

Graduate Program in Mathematics
Boğaziçi University
2013

ACKNOWLEDGEMENTS

I would like to thank my advisor Özlem Beyarslan for her patience and help.

I thank the defense jury, Olcay Coşkun for his support and David Pierce for his thorough review of the thesis and his remarks.

I am grateful to everyone at Boğaziçi University as they made me who I am.

Thanks to my family for their encouragement.

Thanks to TÜBİTAK for graduate scholarship.

ABSTRACT

DEFINABLE SETS OVER FINITE FIELDS

Counting the number of points of an algebraic set over a finite field has been studied by Hasse [1], Lang and Weil [2], and is an important theme in algebra. In this thesis, we present the results found by Chatzidakis, van den Dries and Macintyre in the article *Definable Sets over Finite Fields* [3] and their applications. These results give estimates of the number of points of definable sets over finite fields. Main theorem of the thesis says that given a formula with n variables, the number of points of the set defined by this formula in a finite field $\mathbb{F}_q$ with q elements is approximately μq^d . The constants μ and d can take only finitely many values independent of the field $\mathbb{F}_q$ the formula is defined in.

ÖZET

SONLU CİSİMLER ÜZERİNDE TANIMLANABİLİR KÜMELER

Cebirsel kümelerin sonlu cisimlerdeki sıfırlarını saymak Hasse [1], Lang ve Weil [2] tarafından çalışılmıştır ve cebirin önemli bir konusudur. Bu tezde Chatzidakis, van den Dries ve Macintyre'ın *Definable Sets over Finite Fields* [3] makalesinde buldukları sonuçları ve uygulamalarını sunacağız. Bu sonuçlar sonlu cisimler üzerinde bir formül ile tanımlanan kümelerin noktalarının sayısı için yaklaşık değerler verir. Bu tezin ana teoremi, n değişkenli bir formül verildiğinde q elemanlı sonlu $\mathbb{F}_q$ cisminin üzerinde bu formülün tanımladığı kümenin eleman sayısının yaklaşık μq^d olacağını söyler. Bu μ ve q sabitleri, formülün tanımlandığı cisim $\mathbb{F}_q$ 'dan bağımsız olarak yalnızca sonlu sayıda değer alabilir.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	iii
ABSTRACT	iv
ÖZET	v
LIST OF SYMBOLS	1
1. INTRODUCTION	2
2. ALGEBRAIC GEOMETRY	5
2.1. Preliminaries	5
3. MODEL THEORY	14
3.1. Basic Definitions in Model Theory	14
3.2. Some Results in Model Theory	24
3.3. Ultrafilters and Ultraproducts	31
3.4. Model Theory of Finite and Pseudo-Finite Fields	39
3.5. Enriched Fields	43
4. THE DECOMPOSITION-INTERSECTION PROCEDURE	50
4.1. Constructing the F -Absolute Kernel	50
4.2. Decomposing the F -Absolute Kernel	55
5. MAIN THEOREM	60
5.1. Lang-Weil Theorem	60
5.2. Estimates for Arbitrary Varieties	61
5.3. Estimates for Quantifier Free Definable Sets	66
5.4. Estimates for Arbitrary Definable Sets over Finite Fields	68
6. APPLICATIONS	76
6.1. Interpretation of the Numbers d and μ : Dimension and Measure	76
REFERENCES	79

LIST OF SYMBOLS

$\bar{a}$	Denotes a tuple $(a_1, \dots, a_n)$
$\text{card}(S), S $	Cardinality of the set S
$\text{Diag}(\mathcal{M})$	Atomic diagram of $\mathcal{M}$
$\text{Diag}_{\text{el}}(\mathcal{M})$	Elementary diagram of $\mathcal{M}$
$\tilde{F}$	Algebraic closure of the field F
$\mathbb{F}_q$	Finite field of cardinality q
$\mathcal{M} \models \varphi(\bar{a})$	Denotes that $\mathcal{M}$ satisfies $\varphi(\bar{a})$
$x, y, z, \dots$	Elements of a specified set
$X, Y, Z, \dots$	Variable symbols
φ, ψ	Formulas

1. INTRODUCTION

Model theory is a flourishing area of mathematics which establishes connections between algebra and mathematical logic. It has become more popular among mathematicians after model theoretic techniques are used to prove some important results, like Mordell-Lang conjecture and Ax-Kochen theorem. Recently, an unexpected application of model theory to the additive number theory was done by Hrushovski [4], a discussion about these applications can be found in [5].

Here in this thesis, we present one of the cornerstones of model theory history: the article *Definable Sets over Finite Fields* by Zoé Chatzidakis, Lou van den Dries and Angus Macintyre. This article extends the results of Lang and Weil [2] about estimating the number of points of a variety over a finite field to a much more general context. They give estimates for the number of points of definable sets over finite fields.

The estimate which is found in *Definable Sets over Finite Fields* can be explained as follows: Let $\varphi(\bar{X}, \bar{Y})$ be an arbitrary $\mathcal{L}_{\text{ring}}$ -formula where $\bar{X} = (X_1, \dots, X_m)$ and $\bar{Y} = (Y_1, \dots, Y_n)$ are variables. Then there is a number $C > 0$ and a finite set D of pairs (d, μ) with $d \in \{1, \dots, n\}$ and $\mu \in \mathbb{Q}^+$ such that for any finite field F of cardinality q and for any $\bar{x} \in F^m$ there is a pair (d, μ) in the finite set D satisfying the inequality

$$|\text{card}(\varphi(\bar{x}, F^n)) - \mu q^d| \leq C q^{d-1/2} \quad (1.1)$$

That is, there are only finitely many pairs (d, μ) such that no matter which finite field F and which point $\bar{x} \in F^m$ you pick, one of these (d, μ) will satisfy (1.1)

Let us motivate the result by giving an example. Let $\varphi(Y)$ be the $\mathcal{L}_{\text{ring}}$ -formula $\varphi(Y) = \exists X (Y = X^3)$, the formula which defines the cubes in a given structure. Let $\mathbb{F}_q$ be the finite field of cardinality q , and $\mathbb{F}_q^*$ be the multiplicative group of units in $\mathbb{F}_q$. The function $\psi : X \mapsto X^3$ is a group homomorphism on $\mathbb{F}_q^*$. The kernel of ψ has either

one element or three elements. In the former case, every $y \in \mathbb{F}_q$ is in the set defined by ψ , that is every element in $\mathbb{F}_q$ is a cube. In the latter case, $\text{card}(\psi(\mathbb{F}_q)) = \frac{q-1}{3} + 1$. Note that this case can happen only if $3|q-1$. We see that we have two choices for μ in the main theorem, which are 1 and $1/3$ and two possibilities for d which are 0 and 1.

The organization of the thesis is as follows. There are three main parts in this thesis: the part on model theory, the part on algebraic geometry and the part in which the main theorem is proved and the results of the main theorem are explained.

In Chapter 2, we will give basic definitions and results from algebraic geometry. After defining algebraic sets and irreducible algebraic sets, we are going to prove Hilbert's basis theorem (Theorem 2.1), we will show that every algebraic set can be decomposed uniquely to irreducible algebraic sets (Theorem 2.2) and we will define the dimension of an algebraic set (Definition 2.8).

In Chapter 3, we will start by introducing the main definitions in model theory. Some of the important theorems we present in this chapter are Theorem 3.1, compactness theorem (Theorem 3.2), and Łoś's theorem (Theorem 3.4). To prove the compactness theorem we will use Łoś's theorem, which will be proved in Ultraproducts section. Finally, we will investigate finite and pseudo-finite fields in a model theoretic perspective.

Chapter 4 will introduce the decomposition-intersection procedure. This procedure allows us to work on absolutely irreducible algebraic sets instead of arbitrary algebraic sets. The decomposition-intersection procedure will be the main tool to generalize Lang-Weil theorem (Theorem 5.1) to an arbitrary algebraic set instead of an absolutely irreducible algebraic set. The F -absolute kernel of an F -algebraic set V , that is the F -algebraic set V^* the decomposition-intersection procedure gives when applied to V , will be defined and we will show that the polynomials defining V^* can be calculated using the polynomials defining the absolutely irreducible components of V . Second section of this chapter deals with finding a formula that carries infor-

mation about the polynomials defining the V^* , the number of absolutely irreducible components of V^* and their degrees. This formula will allow us to bound the number of absolutely irreducible components of V^* and the number and the degrees of polynomials defining the absolutely irreducible components of V^* .

The main theorem will be proved in Chapter 5. This proof consists of four steps. We will start by stating that by Lang-Weil theorem we can find an estimate for the number of F -points of absolutely irreducible algebraic sets. Then we will generalize this result to any algebraic set with the help of the decomposition-intersection procedure (Proposition 5.1). Next, we will use the fact that any definable set defined by a quantifier-free formula is defined by some polynomial equalities and inequalities (Lemma 5.1). By adding new variables, we can change inequalities to equalities. Moreover, the number of F -points of the algebraic set defined by these polynomials will be the same as the cardinality of the definable set. The previous result allows us to get an estimate for the cardinality of this definable set. Lastly, by using Lemma 3.2 from model theory, we will state that any definable set is a projection of some algebraic set and prove the result for the general case.

2. ALGEBRAIC GEOMETRY

Algebraic geometry studies polynomial equations and their solutions [6, p. v]. Main objects of study in algebraic geometry are the algebraic sets (sets of solutions of some polynomial equations). Investigating algebraic sets and their interactions is an essential theme in algebraic geometry.

The focus of this thesis is to find an estimate for the number of points of definable sets over finite fields. Since the definable sets over finite fields are closely related to projections of algebraic sets, it is vital to study algebraic geometry. This chapter will provide basic definitions and some results that are needed in the sequel.

Throughout this chapter, F will denote a field, $\tilde{F}$ denote its algebraic closure, and $\bar{X}$ will denote the tuple $(X_1, \dots, X_n)$ for some positive integer n .

2.1. Preliminaries

This section will give basic definitions from algebraic geometry. The definitions are from [6–8]. We will mostly use the notation from [8] as long as it agrees with our fixed notation.

Definition 2.1. [8, Definition 0.1] Let A be a nonempty subset of $F[\bar{X}]$ with $\bar{X} = (X_1, \dots, X_n)$. The *zero locus of A* is the set $V(A) := \{\bar{x} \in \tilde{F}^n : f(\bar{x}) = 0, \forall f \in A\}$.

For any field F , the zero locus of the set $\{X\} \subseteq F[X]$ is the set $V(X) = \{0\}$. The zero locus of $\{0\}$ is the whole field $V(\{0\}) = F$, while $V(F[X]) = \emptyset$. The zeroes of the polynomial $X^2 + Y^2 - 1 \in \mathbb{R}[X, Y]$ form the unit circle in the real plane.

Definition 2.2. An *algebraic set* is a set of the form $V(A)$ for some $A \subseteq F[X]$. To emphasize that polynomials in A have coefficients in F , we will call $V(A)$ an *F -algebraic set*.

Consider the polynomial $X^2 - 2 \in \mathbb{Q}$. Then, $V(\{X^2 - 2\}) = \{\sqrt{2}, -\sqrt{2}\}$ is a $\mathbb{Q}$ -algebraic set. However, $\{\sqrt{2}\}$ is not $\mathbb{Q}$ -algebraic as $\sqrt{2}$ cannot be the only zero of a set of polynomials with coefficients in $\mathbb{Q}$. Yet we can consider $\{\sqrt{2}\}$ as a $\mathbb{Q}(\sqrt{2})$ -algebraic set.

Definition 2.3. [7, p. 178] A point $\bar{a} \in \tilde{F}^n$ in an F -algebraic set V is an F -point if all the coordinates of $\bar{a}$ are in F .

In the previous example, $V(\{X^2 - 2\})$ does not have any $\mathbb{Q}$ -points as $\{\sqrt{2}, -\sqrt{2}\}$ does not have any point in $\mathbb{Q}$. We are interested in the set of F -points of F -algebraic varieties because we would like to study the F -points of sets defined over the field F .

The following result states that instead of considering the zero loci of arbitrary subsets A of $F[\bar{X}]$, we may study the zero loci of ideals of $F[\bar{X}]$.

Proposition 2.1. *Let $A \subseteq F[\bar{X}]$ be nonempty and let J be the ideal generated by A . Then $V(A) = V(J)$.*

Proof. For any $\bar{x} \in V(J)$, $f(\bar{x}) = 0$ for every $f \in J$. We know that $A \subseteq J$, therefore $f(\bar{x}) = 0$ for every $f \in A$ and hence $\bar{x} \in V(A)$.

Now let $\bar{x} \in V(A)$. Let $f \in J$ be arbitrary. J is generated by A therefore $f = g_1 h_1 + \dots + g_n h_n$ for some $n \in \mathbb{N}$, $g_1, \dots, g_n \in F[\bar{X}]$, and $h_1, \dots, h_n \in A$. Since $\bar{x} \in V(A)$, we have $h_i(\bar{x}) = 0$ for $1 \leq i \leq n$. Therefore, $f(\bar{x}) = 0$. Thus $\bar{x} \in V(J)$. $\square$

Definition 2.4. For $V \subseteq \tilde{F}^n$, we have $I(V) := \{f \in \tilde{F}[\bar{X}] : f(\bar{x}) = 0 \forall \bar{x} \in V\}$.

Any set of the form $I(V)$ is an ideal in $\tilde{F}[\bar{X}]$. Let $f, g \in I(V)$ and $h \in \tilde{F}[\bar{X}]$. Then we get $(f - g)(\bar{x}) = f(\bar{x}) - g(\bar{x}) = 0$ and $(hf)(\bar{x}) = (fh)(\bar{x}) = f(\bar{x})h(\bar{x}) = 0$ for any $\bar{x} \in V$. Also, note that $I(\tilde{F}^n) = \{0\}$ and $I(V)$ is always nonempty as $0 \in I(V)$.

V and I can be considered as maps between the set of subsets of $\tilde{F}^n$ and the set of ideals of $\tilde{F}[\bar{X}]$.

Proposition 2.2. [8] Let J and J' be subsets of $\tilde{F}[\bar{X}]$, W and W' be subsets of $\tilde{F}^n$. Then the maps I and V satisfy the following:

- (i) $W \subseteq W' \Rightarrow I(W') \subseteq I(W)$
- (ii) $J \subseteq J' \Rightarrow V(J') \subseteq V(J)$
- (iii) $W \subseteq V(I(W))$ (equality holds if and only if W is an algebraic set)
- (iv) $J \subseteq I(V(J))$.

Proof. (i) Let $W \subseteq W'$. Let $f \in I(W')$. Then $f(\bar{x}) = 0$ for any $\bar{x} \in W'$ and hence for any $\bar{x} \in W$, so $f \in I(W)$.

(ii) Let $J \subseteq J'$. If $V(J') = \emptyset$, there is nothing to prove. Let $\bar{x} \in V(J')$ be arbitrary. Then for any $f \in J'$, $f(\bar{x}) = 0$. Hence, for any $f \in J$, we have $f(\bar{x}) = 0$, so $f \in V(J)$.

(iii) If $W = \emptyset$, there is nothing to prove. Let $\bar{x} \in W$. Let $f \in I(W)$ be arbitrary. By definition, $f(\bar{x}) = 0$. Since f was arbitrary, we have for all $f \in I(W)$, $f(\bar{x}) = 0$, that is, $\bar{x} \in V(I(W))$.

If $W = V(I(W))$, then by definition, W is an algebraic set. Now, suppose that W is an algebraic set. Then, there exists a nonempty set $K \subseteq \tilde{F}[\bar{X}]$ such that $W = V(K)$. For any $f \in K$, we have $f(\bar{x}) = 0$ for each $\bar{x} \in W$, so $f \in I(W)$, that is, $K \subseteq I(W)$. By part (ii), $V(I(W)) \subseteq V(K) = W$. Since $W \subseteq V(I(W))$ is already true, we have the equality.

(iv) Let $f \in J$. If $V(J) = \emptyset$, then $I(V(J)) = \tilde{F}[\bar{X}]$. Suppose that $V(J) \neq \emptyset$. Then, $f(\bar{x}) = 0$ for all $\bar{x} \in V(J)$, that is, $f \in I(V(J))$ □

The algebraic subsets of F^n can be seen as the closed sets for a topology, namely the *Zariski Topology*, as the following proposition shows.

Proposition 2.3. [8, Lemma 1.1]

- (i) F and $\emptyset$ are algebraic sets.
- (ii) If V and W are algebraic, then so is $V \cup W$.
- (iii) If V_α , $\alpha \in \Lambda$ is a family of algebraic sets, then $\bigcap_{\alpha \in \Lambda} V_\alpha$ is an algebraic set.

Proof. (i) $V(\{0\}) = F$ and $V(F[X]) = \emptyset$.

(ii) Let $V = V(I)$ and $W = V(J)$. Then, $V(I \cap J) = V \cup W$. Indeed, by Proposition 2.2 part (ii), we see that $V(I) \subseteq V(I \cap J)$ and $V(J) \subseteq V(I \cap J)$, therefore $V \cup W = V(I) \cup V(J) \subseteq V(I \cap J)$. Now let $x \notin V \cup W$. Then $x \notin V$ and $x \notin W$. So, there are $f \in I$ and $g \in J$ with $f(x) \neq 0$ and $g(x) \neq 0$. Therefore, $fg(x) \neq 0$. Since $fg \in I \cap J$, $x \notin V(I \cap J)$.

(iii) Let $V_\alpha = V(I_\alpha)$ for each $\alpha \in \Lambda$. Then, $\bigcap_{\alpha \in \Lambda} V_\alpha = V(\langle \bigcup_{\alpha \in \Lambda} I_\alpha \rangle)$. Again, by Proposition 2.2 part (ii), $I_\alpha \subseteq \bigcup_{\alpha \in \Lambda} I_\alpha$ implies $V(\langle \bigcup_{\alpha \in \Lambda} I_\alpha \rangle) \subseteq V(I_\alpha)$ for each $\alpha \in \Lambda$. Therefore, $V(\langle \bigcup_{\alpha \in \Lambda} I_\alpha \rangle) \subseteq \bigcap_{\alpha \in \Lambda} V_\alpha$. Now, let $x \in \bigcap_{\alpha \in \Lambda} V_\alpha$. Let $f \in \langle \bigcup_{\alpha \in \Lambda} I_\alpha \rangle$. Then $f = g_1 h_1 + \dots + g_n h_n$ for some $n \in \mathbb{N}$, $g_1, \dots, g_n \in F[X]$ and $h_1, \dots, h_n \in \bigcup_{\alpha \in \Lambda} I_\alpha$. So, $f(x) = 0$, and thus, $\bigcap_{\alpha \in \Lambda} V_\alpha \subseteq V(\langle \bigcup_{\alpha \in \Lambda} I_\alpha \rangle)$. $\square$

We will now define an irreducible algebraic set. As we will see later on, each algebraic set can be decomposed uniquely into a finite number of irreducible subsets. Furthermore, irreducible algebraic sets have properties that make them easier to study, which is why they are an important tool in algebraic geometry.

Definition 2.5. [8, Definiton 1.6] An F -algebraic set V is called *reducible* if $V = A \cup B$ for some F -algebraic sets $A, B \subsetneq V$. Otherwise, V is called *irreducible*.

Our previous example $V = V(\{X^2 - 2\})$ is irreducible as a $\mathbb{Q}$ -algebraic set since $\{\sqrt{2}, -\sqrt{2}\}$ cannot be written as a union of two proper $\mathbb{Q}$ -algebraic subsets. On the other hand, the $\mathbb{Q}$ -algebraic set $V(\{X^2 - Y^2\})$ is the union of the $\mathbb{Q}$ -algebraic sets $V(\{X - Y\})$ and $V(\{X + Y\})$.

The next definition explains what an absolutely irreducible algebraic set is. These sets are needed for the proof of the main theorem. Lang-Weil theorem gives an estimate for the number of elements of absolutely irreducible algebraic sets over finite fields. This results will be generalized to arbitrary algebraic sets and then to sets defined over finite fields by quantifier-free formulas. Last step is to prove the main theorem using these results.

Definition 2.6. [7, p.175] An F -algebraic set V is called *absolutely irreducible* if it is irreducible for every extension of F . In other words, if F' is an extension field of F and if A and B are F' -algebraic sets, then V is irreducible if $V = A \cup B$ implies that $A = V$ or $B = V$.

The example above is not absolutely irreducible as $\{\sqrt{2}, -\sqrt{2}\}$ can be written as a union of two proper $\mathbb{Q}(\sqrt{2})$ -algebraic subsets, namely $V(\{X - \sqrt{2}\})$ and $V(\{X + \sqrt{2}\})$.

We will state Hilbert's basis theorem, but before that we will introduce some preliminaries.

Definition 2.7. [9, Definition viii,1.1] A ring R is *Noetherian* if every ascending chain $I_0 \subseteq I_1 \subseteq \dots$ of ideals of R stabilizes, *i.e.* there exists a natural number n such that $i \geq n$ implies that $I_i = I_n$.

The following lemma gives an equivalent definition of Noetherian rings.

Lemma 2.1. [10, x,1] *R is a Noetherian ring if and only if every ideal of R is finitely generated.*

Proof. Let R be a Noetherian ring. We will first show that every nonempty set of ideals of R has a maximal element. Let $\mathcal{S}$ be a nonempty set of ideals of R . Let $I_0 \subseteq I_1 \subseteq \dots$ be an arbitrary chain in $\mathcal{S}$. Since R is Noetherian, this chain has an upper bound, say $I_n \in \mathcal{S}$. Since every chain in $\mathcal{S}$ has an upper bound, $\mathcal{S}$ has a maximal element by Zorn's Lemma.

Now we will show that every ideal of R is finitely generated. Let I be an ideal in R . Let $\mathcal{S}$ be the set of finitely generated ideals of R that are also subsets of I . This set has a maximal element, say J . Since J is finitely generated, there are $r_1, \dots, r_n \in J$ such that $J = \langle r_1, \dots, r_n \rangle$. For each $x \in I$, define $J_x := \langle r_1, \dots, r_n, x \rangle$. J_x is a finitely generated ideal and $J_x \subseteq I$ for all $x \in I$, hence $J_x \in \mathcal{S}$ for all $x \in I$. Moreover, $J \subseteq J_x$.

Since J was maximal, we get $J_x = J$ for each x . Thus, every element of I must be an element of J , *i.e.* $I \subseteq J$. Therefore $I = J$ and I is a finitely generated ideal of R .

For the converse, assume that every ideal of the ring R is finitely generated. Let $I_0 \subseteq I_1 \subseteq \dots$ be a chain of ideals in R . Put $I := \bigcup_{i=0}^{\infty} I_i$. I is an ideal of R and hence finitely generated. Suppose that $I = \langle x_1, \dots, x_n \rangle$. From the way I is defined, we see that each x_j is in some I_i . Thus, there exists some N such that $\{x_1, \dots, x_n\} \subseteq I_N$. Therefore $I \subseteq I_N$ and hence $I = I_N$. The chain $I_0 \subseteq I_1 \subseteq \dots$ terminates at I_N . Since this is true for any chain, R is Noetherian. $\square$

Now we can state and prove Hilbert's basis theorem.

Theorem 2.1. [10, iv, Theorem 4.1] *If R is a commutative Noetherian ring, then $R[\bar{X}]$ is a Noetherian ring.*

Proof. It is sufficient to prove the theorem for $R[X]$ because by induction on the number of components of $\bar{X}$, we can prove the general case. Assume that for any Noetherian ring R , $R[X]$ is Noetherian. Suppose that R is Noetherian and $R[X_1, \dots, X_{n-1}]$ is also Noetherian. Since $R[X_1, \dots, X_n] = R[X_1, \dots, X_{n-1}][X_n]$, and since $R[X_1, \dots, X_n]$ is Noetherian, we see that $R[X_1, \dots, X_n]$ is Noetherian by induction hypothesis.

Let R be a Noetherian ring. We will show that any ideal I of $R[X]$ is finitely generated. Define new sets $I_i := \{a \in R : aX^i + a_{i-1}X^{i-1} + \dots + a_0X^0 \in I\} \cup \{0\}$. These sets are ideals as we will see now.

For a fixed i , let $a, b \in I_i$ and $r \in R$. If both a and b are zero or $ra - b = 0$, then $ra - b \in I_i$ and we are done. Suppose now that $ra - b \neq 0$. If both a and b are nonzero, then there are polynomials $f(X), g(X) \in I$ of degree i with leading coefficients a and b , respectively. Since I is an ideal of $R[X]$, $rf(X) - g(X) \in I$. The leading coefficient of $rf(X) - g(X)$ is $ra - b$ and $\deg(rf(X) - g(X)) = i$, hence $ra - b \in I_i$. If a is zero and b is nonzero, then $-g(X) \in I$, hence $ra - b \in I_i$. If b is zero and a is nonzero, then $rf(X) \in I$, hence $ra - b \in I_i$.

Now, we will show that $I_0 \subseteq I_1 \subseteq \dots$ is a chain of ideals. Fix i and let $a \in I_i$. If $a = 0$, then $a \in I_{i+1}$. Suppose that $a \neq 0$. Then there is a polynomial $f(X) \in I$ of degree i with leading coefficient a . Since I is an ideal, $Xf(X) \in I$. Moreover, $Xf(X)$ is a polynomial of degree $i + 1$ with leading coefficient a . Hence $a \in I_{i+1}$.

Note that $I_0 \subseteq I_1 \subseteq \dots$ is a chain of ideals of the Noetherian ring R , thus there is an N such that $I_i = I_N$ for $i \geq N$. Furthermore, each I_i is finitely generated. Let for each $i \in \{0, 1, \dots, N\}$, $I_i = \langle a_{i,1}, \dots, a_{i,n_i} \rangle$. For every nonzero $a_{i,j}$, fix a polynomial $f_{i,j}(X) \in I$ of degree i with leading coefficient $a_{i,j}$. If there is no nonzero $a_{i,j}$ for some certain i , then put $f_{i,j}(X) = 0$. We claim that these $f_{i,j}(X)$ generates I .

Define $F := \{f_{i,j}(X) : 0 \leq i \leq N, 1 \leq j \leq n_i\}$. This set is a subset of I , so $\langle F \rangle \subseteq I$. We will show that $I \subseteq \langle F \rangle$ by proving that any polynomial in I can be written as a linear combination of polynomials from F . We will use induction on the degrees of polynomials in I . If $f(X)$ is of degree 0, then $f(X) = a_{0,j}$ for some $j \in \{1, \dots, n_0\}$. Suppose that any polynomial of degree $< i$ can be written as a linear combination of polynomials from F . Let $f(X) := c_i X^i + \dots + c_0 X^0$, with $c_i \neq 0$. If $i \geq N$, then $c_i = r_1 a_{N,1} + \dots + r_{n_N} a_{N,n_N}$. Put $g(X) := f(X) - X^{i-N} (r_1 f_{N,1} + \dots + r_{n_N} f_{N,n_N})$. If $i < N$, then $c_i = r_1 a_{i,1} + \dots + r_{n_i} a_{i,n_i}$. Put $g(X) := f(X) - (r_1 f_{i,1} + \dots + r_{n_i} f_{i,n_i})$. In either case, $g(X)$ is an element of I and is of degree $< i$. By induction hypothesis, $g(X)$ is a linear combination of polynomials in F . Thus, $f(X)$ is a linear combination of polynomials in F . Therefore, $I \subseteq \langle F \rangle$. Hence, $I = \langle F \rangle$, i.e. I is finitely generated. $\square$

Hilbert's basis theorem and Definition 2.4 allows us to get the following results:

Corollary 2.1. (i) Every descending chain $V_0 \supseteq V_1 \supseteq \dots$ of $\tilde{F}$ -algebraic sets stabilizes at some V_N .

(ii) Every nonempty set of $\tilde{F}$ -algebraic sets has a minimal element.

Proof. (i) For any descending chain $V_0 \supseteq V_1 \supseteq \dots$ of algebraic sets, we have an ascending chain $I(V_0) \subseteq I(V_1) \subseteq \dots$ of ideals in $\tilde{F}[\bar{X}]$ by Proposition 2.2 part (i). Since $\tilde{F}$ is a field, it is a Noetherian ring. Therefore, by Theorem 2.1 $\tilde{F}[\bar{X}]$

is Noetherian. This means that the chain $I(V_0) \subseteq I(V_1) \subseteq \dots$ will stop at some $I(V_N)$. By Proposition 2.2 part (ii) we get $V(I(V_0)) \supseteq V(I(V_1)) \supseteq \dots$ stops at $V(I(V_N))$, and by Proposition 2.2 part (iii) we get $V_0 \supseteq V_1 \supseteq \dots$ stops at V_N since V_i are algebraic sets.

- (ii) Let Σ be a set of $\tilde{F}$ -algebraic sets. Define $\mathcal{S} := \{I(V) : V \in \Sigma\}$. $\mathcal{S}$ is a set of ideals in $\tilde{F}[\bar{X}]$, and hence has a maximal element as we have seen in the proof of Lemma 2.1. By Proposition 2.2, Σ has a minimal element. $\square$

The following theorem states that every algebraic sets can be decomposed uniquely into a finite number of irreducible algebraic sets as we mentioned earlier. This enables us to deal with more complex algebraic sets.

Theorem 2.2. [8, Proposition 1.9] *Any F -algebraic set V has a decomposition $V = V_1 \cup \dots \cup V_n$ into irreducible F -algebraic subsets such that if $i \neq j$ then $V_i \not\subseteq V_j$. Moreover, this decomposition is unique up to ordering.*

Proof. We will first prove that every F -algebraic set can be written as a finite union of irreducible F -algebraic sets. Suppose that there are F -algebraic sets V that cannot be decomposed into a finite number of irreducible F -algebraic sets. Let Σ be the set of such F -algebraic sets. By Corollary 2.1, this set has a minimal element, say V . V should be reducible because otherwise it could be written as a finite union of irreducible F -algebraic sets. Since V is reducible, there are F -algebraic sets V_1 and V_2 such that $V = V_1 \cup V_2$, where $V_1, V_2 \subsetneq V$. V is minimal in Σ , therefore V_1 and V_2 cannot be in Σ . Hence, they can be decomposed into a finite number of irreducible F -algebraic sets. Thus, V can be written as a union of irreducible components of V_1 and V_2 . Hence, V has a finite decomposition.

For the uniqueness, assume an F -algebraic set V can be decomposed into irreducible F -algebraic sets as $V = V_1 \cup \dots \cup V_n$ and $V = W_1 \cup \dots \cup W_m$, where $V_i \not\subseteq V_j$ and $W_i \not\subseteq W_j$ if $i \neq j$. Fix i , consider the set $V_i = V \cap V_i$. We see that $V_i = \bigcup_{j=1}^m (V_i \cap W_j)$. Since V_i is irreducible, there is j such that $V_i \cap W_j = V_i$ and hence $V_i \subseteq W_j$. Similarly,

$W_j = \bigcup_{k=1}^n (W_j \cap V_k)$ and there is k with $W_j \cap V_k = W_j$ so $W_j \subseteq V_k$. These two subset relations give us $V_i \subseteq V_k$. By the assumption we get $i = k$, thus $V_i = W_j$. Since i was arbitrary, each V_i is equal to some W_j . Applying the same technique for each W_j , we see that there is some V_k with $W_j = V_k$. Therefore, two decompositions are the same up to the ordering of the irreducible algebraic sets. $\square$

We will now define the dimension of an algebraic set. There are other equivalent definitions to the one given below but this definition can be given with the tools in hand.

Definition 2.8. The dimension of a nonempty F -algebraic set $V \subseteq F^n$ is the length of the longest chain $V_0 \subsetneq V_1 \subsetneq \dots \subsetneq V_d \subsetneq V$ of nonempty irreducible F -algebraic sets V_i . In this case the dimension of V is $\dim(V) = d$.

The dimension of the algebraic set $V(\{X, Y, Z\}) \subseteq F[X, Y, Z]$ is 0. $V(\{X, Y, Z\}) = \{(0, 0, 0)\}$ and the only chain of irreducible algebraic sets is of length 0. The dimension of $V(\{Z\}) \subseteq F[X, Y, Z]$ is 2. The elements of $V(\{Z\})$ forms the xy plane in F^3 . One example of a chain of length 2 is $V(\{X, Y, Z\}) \subseteq V(\{Y, Z\}) \subseteq V(\{Z\})$.

Remark 2.1. Here are a few remarks on the properties of the dimension.

- (i) $\dim(F^n) = n$ for any field F .
- (ii) For any two F -algebraic sets V_1 and V_2 , if $V_1 \subseteq V_2$, then $\dim(V_1) \leq \dim(V_2)$.
- (iii) If V_1 and V_2 are absolutely irreducible F -algebraic sets, then $\dim(V_1 \cap V_2) < \dim(V_1)$ and $\dim(V_1 \cap V_2) < \dim(V_2)$ [7].

3. MODEL THEORY

Model theory studies the relations between mathematical structures and the set of sentences that are true for these structures (these sets are called theories). The relations work for both ways. We can investigate properties of theories in general and then apply the result to specific structures, or we can use certain properties of the structures to gather information about the forms of sentences over these structures.

This chapter will give some basic definitions and theorems on model theory. Moreover, we will prove the model theoretic facts we are going to use in the following chapters.

3.1. Basic Definitions in Model Theory

This section introduces basic definitions and concepts in model theory. Main tools in model theory are formulas and sentences. These terms are defined inductively, before defining them we need to define their components. We will start by defining the set of symbols that are needed to form formulas.

Definition 3.1. [11, Definition 1.1.1] A *language* $\mathcal{L}$ is a triple $(\mathcal{C}, \mathcal{F}, \mathcal{R})$ where $\mathcal{C}$, $\mathcal{F}$, and $\mathcal{R}$ are sets. $\mathcal{C}$ is called the set of constant symbols, $\mathcal{F}$ is called the set of function symbols, and $\mathcal{R}$ is called the set of relation symbols. Each element c of $\mathcal{C}$ is a constant symbol, each element (f, n_f) of $\mathcal{F}$ is an n_f -ary function symbol, and each element (R, n_R) of $\mathcal{R}$ is an n_R -ary relation symbol.

Whenever we can, we will write a language $\mathcal{L}$ as $\{c_1, c_2, \dots, f_1, f_2, \dots, R_1, R_2, \dots\}$, where $c_i \in \mathcal{C}$, $f_i \in \mathcal{F}$, and $R_i \in \mathcal{R}$ for each i . When we say $\alpha \in \mathcal{L}$, we mean that α is in one of the sets $\mathcal{C}$, $\mathcal{F}$, or $\mathcal{R}$. Using this notation, we will now give some examples of languages.

If we want to talk about graphs, we must use vertices and edges. An edge is a

binary relation over vertices, so we need a relation symbol R in our formulas. The language of graphs is the set $\mathcal{L}_{\text{graph}} = \{R\}$ where R is a relation symbol. This language does not have any constant or function symbols. Note that the same language can also be used to speak about ordered sets.

The language of additive groups is the set $\mathcal{L}_{\text{group}} = \{0, +, -\}$ where 0 is a constant symbol, $+$ is a binary function symbol and $-$ is a unary function symbol. To make statements about rings, we extend this set to $\mathcal{L}_{\text{ring}} = \{0, 1, +, -, \cdot\}$, where 1 is a constant symbol and $\cdot$ is a relation symbol.

With a language in hand, we can form formulas over some sets. These sets will be called $\mathcal{L}$ -structures.

Definition 3.2. [12, Definition 1.1.2] Let $\mathcal{L} = (\mathcal{C}, \mathcal{F}, \mathcal{R})$ be a language. An $\mathcal{L}$ -structure $\mathcal{M}$ is a pair $(M, (\cdot^{\mathcal{M}}))$, where M is a nonempty set called the *underlying set*, and $\cdot^{\mathcal{M}}$ is a function over $\mathcal{L}$ having the following properties:

- (i) If $c \in \mathcal{C}$, then $c^{\mathcal{M}}$ is an element in M
- (ii) If $(f, n_f) \in \mathcal{F}$, then $f^{\mathcal{M}}$ is a function from M^{n_f} to M
- (iii) If $(R, n_R) \in \mathcal{R}$, then $R^{\mathcal{M}}$ is a relation on M^{n_R}

For any $\alpha \in \mathcal{L}$, $\alpha^{\mathcal{M}}$ is called the $\mathcal{M}$ -interpretation of α .

As we discussed before, for any graph G , the element R of the language $\mathcal{L}_{\text{graph}}$ is interpreted over G as the edges between the vertices. For example, we can form the sentence $\forall X \forall Y (X R^{\mathcal{G}} Y \leftrightarrow Y R^{\mathcal{G}} X)$, stating that having an edge is a symmetric relation. The pair $\mathcal{G} = (G, \cdot^{\mathcal{G}})$ is an $\mathcal{L}_{\text{group}}$ -structure. We will sometimes say that the set M is an $\mathcal{L}$ -structure, whenever there is no confusion about the interpretations.

The set of integers $\mathbb{Z}$ is both an $\mathcal{L}_{\text{group}}$ and an $\mathcal{L}_{\text{ring}}$ -structure. The interpretations of elements in these languages are in their usual meanings.

Now we will define the maps between $\mathcal{L}$ -structures that preserves the interpretations.

Definition 3.3. [12, Definition 1.1.3] Let $\mathcal{L} = (\mathcal{C}, \mathcal{F}, \mathcal{R})$ be a language, $\mathcal{M}$ and $\mathcal{N}$ be two $\mathcal{L}$ -structures. A function $\eta : M \rightarrow N$ is called an $\mathcal{L}$ -homomorphism if the following conditions hold:

- (i) For every $c \in \mathcal{C}$, $\eta(c^{\mathcal{M}}) = c^{\mathcal{N}}$.
- (ii) For every $(f, n_f) \in \mathcal{F}$, $\eta(f^{\mathcal{M}}(a_1, \dots, a_{n_f})) = f^{\mathcal{N}}(\eta(a_1), \dots, \eta(a_{n_f}))$.
- (iii) For every $(R, n_R) \in \mathcal{R}$, $(a_1, \dots, a_{n_R}) \in R^{\mathcal{M}}$ implies $(\eta(a_1), \dots, \eta(a_{n_R})) \in R^{\mathcal{N}}$.

If η is injective and satisfy the condition

- (i) For every $(R, n_R) \in \mathcal{R}$, $(a_1, \dots, a_{n_R}) \in R^{\mathcal{M}}$ if and only if $(\eta(a_1), \dots, \eta(a_{n_R})) \in R^{\mathcal{N}}$,

then η is called an $\mathcal{L}$ -embedding. If an $\mathcal{L}$ -embedding η is surjective, then it is called an $\mathcal{L}$ -isomorphism.

A substructure can be defined using an $\mathcal{L}$ -embedding.

Definition 3.4. [12, Definition 1.1.5] Let $\mathcal{L}$ be a language, $\mathcal{M}$ and $\mathcal{N}$ be $\mathcal{L}$ -structures. Suppose that $M \subseteq N$. If the inclusion map $\eta : M \rightarrow N$, defined as $\eta(m) = m$ for all $m \in M$, is an $\mathcal{L}$ -embedding, then we say that $\mathcal{M}$ is a substructure of $\mathcal{N}$.

We have stated earlier that a formula will be defined inductively. Most basic elements of formulas are symbols coming from a fixed language. The next step is to use some of these symbols and extra variable symbols to form terms.

Definition 3.5. Let $\mathcal{L} = (\mathcal{C}, \mathcal{F}, \mathcal{R})$ be a language. We call t an $\mathcal{L}$ -term if

- (i) $t = c$ for some $c \in \mathcal{C}$
- (ii) $t = X$ for some variable symbol X
- (iii) $t = f(t_1, \dots, t_{n_f})$ for some terms $t_1, \dots, t_{n_f}$, and for some $(f, n_f) \in \mathcal{F}$.

Let X and Y be variable symbols. Then $0, X, X - (X + Y)$ are all $\mathcal{L}_{\text{group}}$ -terms. For the language $\mathcal{L}_{\text{ring}}$, each $\mathcal{L}_{\text{ring}}$ -term is a polynomial over $\mathbb{Z}$, *e.g.* for X and Y variable symbols, $X^2 - 3XY + Y + 2$ is an $\mathcal{L}_{\text{ring}}$ -term, which can be written also as $(X \cdot X) + (-((1 + 1 + 1) \cdot X \cdot Y)) + Y + (1 + 1)$. For brevity, we will usually use the name “variable” instead of variable symbol.

Variable symbols are a new addition to our set of symbols and we have not yet defined their interpretation over a structure. They will be considered as a varying element from the structure. The next definition indicates how to interpret terms, and hence variables, over a structure.

Definition 3.6. Let $\mathcal{L}$ be a language, $\mathcal{M}$ be an $\mathcal{L}$ -structure. Let t be a term over $\mathcal{L}$ and $\bar{a} = (a_1, \dots, a_n) \in M^n$ be arbitrary. Then $t^{\mathcal{M}}(\bar{a})$ is defined as:

- (i) $c^{\mathcal{M}}$ if $t = c$ for some constant symbol c
- (ii) a_i if $t = X_i$ for some variable symbol X_i
- (iii) $f^{\mathcal{M}}(t_1^{\mathcal{M}}(\bar{a}), \dots, t_m^{\mathcal{M}}(\bar{a}))$ if $t = f(t_1, \dots, t_m)$, f is an m -ary function symbol and $t_1, \dots, t_m$ are $\mathcal{L}$ -terms.

The next item uses equality signs and relation symbols, in addition to the symbols from terms.

Definition 3.7. Let $\mathcal{L} = (\mathcal{C}, \mathcal{F}, \mathcal{R})$ be a language. φ is called an *atomic $\mathcal{L}$ -formula* if

- (i) $\varphi = (t_1 = t_2)$ for some terms t_1 and t_2
- (ii) $\varphi = R(t_1, \dots, t_{n_R})$ for a relation symbol $(R, n_R) \in \mathcal{R}$ and terms $t_1, \dots, t_{n_R}$.

Each atomic formula in $\mathcal{L}_{\text{ring}}$ is of the form $f(X_1, \dots, X_n) = g(X_1, \dots, X_n)$ for some polynomials $f, g \in \mathbb{Z}[X_1, \dots, X_n]$ because there are no relation symbols in $\mathcal{L}_{\text{ring}}$. We will see later that in the theory of rings, this atomic formula can also be expressed by the equivalent atomic formula $h(X_1, \dots, X_n) = 0$ where $h = f - g$. Therefore, any atomic formula states that a polynomial is equal to zero in the theory of rings.

We are now able to define an $\mathcal{L}$ -formula.

Definition 3.8. Let $\mathcal{L}$ be a language. An $\mathcal{L}$ -formula is one of the following

- (i) an atomic formula φ
- (ii) $\neg\varphi$ if φ is a formula or an atomic formula
- (iii) $(\varphi \wedge \psi)$ if φ and ψ are formulas or atomic formulas
- (iv) $\exists X \varphi$ if φ is a formula or an atomic formula.

We will use the following familiar abbreviations for the given formulas:

- $\varphi \vee \psi$ denotes $\neg(\neg\varphi \wedge \neg\psi)$
- $\varphi \rightarrow \psi$ denotes $\neg\varphi \vee \psi$
- $\varphi \leftrightarrow \psi$ denotes $(\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi)$
- $\forall X \varphi$ denotes $\neg(\exists X \neg\varphi)$.

The inductive definition of formulas allows us to make inductive proofs. We will start by proving the desired statement for terms. Then assuming the statement is true for all terms, we will prove the statement for atomic formulas and lastly for formulas.

The formulas that do not involve quantifier symbols are of particular interest, so they are given a special name.

Definition 3.9. A formula that does not have any quantifiers is called *quantifier-free*.

The following definitions are needed to define a sentence.

Definition 3.10. [12, Definition 1.2.10] A variable X is called *free* in the formula

- (i) φ if φ is an atomic formula and X appears in φ
- (ii) $\neg\varphi$ if X is free in φ
- (iii) $\varphi \wedge \psi$ if X is free in φ or ψ
- (iv) $\exists Y \varphi$ if $X \neq Y$ and X is free in φ .

Definition 3.11. A variable X is *in the scope of a quantifier* in the formula φ if X appears in ψ and $\varphi = \exists X \psi$. A variable in a formula is called *bound* if it is in the scope of a quantifier.

These definitions allow a variable to be both free and bound in a formula, *e.g.* X is both free and bound in $(\exists X (X \cdot Y = 1)) \wedge (X \cdot Y = 0)$. This can be prevented by replacing free occurrences of the variable by a new variable. The formula in the previous example becomes $(\exists X (X \cdot Y = 1)) \wedge (Z \cdot Y = 0)$, we replaced the last X with Z . We will denote a formula φ as $\varphi(X_1, \dots, X_n)$ where $\{X_1, \dots, X_n\}$ is the set of all free variables occurring in φ .

We have got the adequate information to define a sentence.

Definition 3.12. A formula is called a *sentence* if it does not have any free variables.

In the language of graphs, $\forall X \forall Y (X R Y)$ is a sentence, stating that “there is an edge between any two vertices”. The sentence $\forall X \exists Y (X \cdot Y = 1)$ is in the language of rings, and it says that “every element has a multiplicative inverse”. Notice that there are graphs that do not satisfy the first statement and there are rings that do not satisfy the second statement. We will now define what does it mean for a structure to satisfy a formula, and hence a sentence.

Definition 3.13. Let $\mathcal{L}$ be a language and $\mathcal{M}$ be an $\mathcal{L}$ -structure. Let $\varphi(X_1, \dots, X_n)$ be an $\mathcal{L}$ -formula, $\bar{a} = (a_1, \dots, a_n) \in M^n$ be arbitrary. We say that $\mathcal{M}$ satisfies $\varphi(\bar{a})$, and write $\mathcal{M} \models \varphi(\bar{a})$, if

- (i) $\varphi = (t_1 = t_2)$ and $t_1^{\mathcal{M}}(\bar{a}) = t_2^{\mathcal{M}}(\bar{a})$ where t_1 and t_2 are terms,
- (ii) $\varphi = R(t_1, \dots, t_m)$ and $(t_1^{\mathcal{M}}(\bar{a}), \dots, t_m^{\mathcal{M}}(\bar{a})) \in R^{\mathcal{M}}$ where R is an m -ary relation symbol and $t_1, \dots, t_m$ are terms
- (iii) $\varphi = \neg\psi$ and $\mathcal{M} \not\models \psi(\bar{a})$, where ψ is a formula or an atomic formula,
- (iv) $\varphi = (\psi_1 \wedge \psi_2)$ and both $\mathcal{M} \models \psi_1(\bar{a})$ and $\mathcal{M} \models \psi_2(\bar{a})$, where ψ_1 and ψ_2 are formulas or atomic formulas.

(v) $\varphi = (\exists Y \psi)$ and there is $a' \in M$ with $\mathcal{M} \models \psi(a_1, \dots, a_m, a')$, where ψ is a formula or an atomic formula.

Note that if φ is a sentence, then $\mathcal{M} \models \varphi(\bar{a})$ is either true for every $\bar{a} \in M$ or false for every $\bar{a} \in M$ because there are no free variables in φ . Therefore we either write $\mathcal{M} \models \varphi$ or $\mathcal{M} \not\models \varphi$, omitting $\bar{a}$.

Abbreviations $\vee, \rightarrow, \leftrightarrow$ and $\forall$ agree with their usual meanings. We will only show that $\mathcal{M} \models (\forall X \varphi)(\bar{a})$ means that $\mathcal{M}$ satisfies $\varphi(\bar{a}, X)$ for all values of X . $\forall X \varphi$ abbreviates $\neg(\exists X \neg\varphi)$. Then $\mathcal{M} \not\models (\exists X \neg\varphi)(\bar{a})$, that is, it is not the case that there is $a' \in M$ with $\mathcal{M} \models \neg\varphi(\bar{a}, a')$. Hence, for all $a' \in M$, $\mathcal{M} \models \neg\varphi(\bar{a}, a')$ is false. However $\mathcal{M} \models \neg\varphi(\bar{a}, a')$ means that $\mathcal{M} \not\models \varphi(\bar{a}, a')$. Thus we get for all $a' \in M$, $\mathcal{M} \models \varphi(\bar{a}, a')$ as we wanted.

Sentences have a truth value over a structure, regardless of the specific elements of the structure. So if we fix a sentence, or a set of sentences, we can talk about the structures that satisfy the sentence or sentences. This subject deserves more attention, so we need new definitions.

Definition 3.14. Let $\mathcal{L}$ be a language. An $\mathcal{L}$ -theory is a set T of $\mathcal{L}$ -sentences.

Theories are usually named by the structures they define (We will explain this in the following definition). For example, the theory of graphs is the set T_{graph} consisting of the sentence $\forall X \forall Y (X R^g Y \leftrightarrow Y R^g X)$, written using the language $\mathcal{L}_{\text{graph}}$. A theory with more elements is the theory of rings, written T_{ring} . The sentences in T_{ring} are

- $\forall X \forall Y \forall Z ((X + Y) + Z = X + (Y + Z))$
- $\forall X ((X + 0 = X) \wedge (0 + X = X))$
- $\forall X (X + (-X) = 0)$
- $\forall X \forall Y (X + Y = Y + X)$
- $\forall X \forall Y \forall Z ((X \cdot Y) \cdot Z = X \cdot (Y \cdot Z))$

- $\forall X ((1 \cdot X = X) \wedge (X \cdot 1 = X))$
- $\forall X \forall Y \forall Z (X \cdot (Y + Z) = (X \cdot Y) + (X \cdot Z))$
- $\forall X \forall Y \forall Z ((Y + Z) \cdot X = (Y \cdot X) + (Z \cdot X)).$

The first three sentences in the above list form the theory of groups. If we add the fourth sentence to this theory, we get the theory of abelian groups.

The next definition explains how a theory defines a structure.

Definition 3.15. Let $\mathcal{L}$ be a language and T be an $\mathcal{L}$ -theory. An $\mathcal{L}$ -structure $\mathcal{M}$ is called a *model* of T if $\mathcal{M} \models \varphi$ for all $\varphi \in T$.

We will write $T \models \varphi$ whenever $\mathcal{M} \models \varphi$ for any model $\mathcal{M}$ of T .

We can now justify that the formulas $\varphi(\bar{X}) = (f(X_1, \dots, X_n) = g(X_1, \dots, X_n))$ and $\psi(\bar{X}) = ((f - g)(X_1, \dots, X_n) = 0)$ are equivalent in the theory of rings. We say that two formulas $\varphi(\bar{X})$ and $\psi(\bar{X})$ are *equivalent in the theory T* if $T \models \forall \bar{X} (\varphi(\bar{X}) \leftrightarrow \psi(\bar{X}))$. Now, fix a model $\mathcal{M}$ for the theory of rings. We must show that for every $\bar{a} \in M^n$, $\mathcal{M} \models (\varphi(\bar{a}) \leftrightarrow \psi(\bar{a}))$. Let $\bar{a} \in M^n$ be arbitrary. $f(\bar{a})$ and $g(\bar{a})$ are two elements of the set M , where the functions $+$, $-$ and the element 0 satisfy the sentences in the theory of rings. Therefore, it is easy to see that $f(\bar{a}) = g(\bar{a})$ if and only if $(f - g)(\bar{a}) = 0$.

The following definition introduces the concept of definable sets, which appears on the title of this thesis. Definable sets will be used frequently in the proof of the main theorem.

Definition 3.16. [11, Definition 1.3.1] Let $\mathcal{L}$ be a language, $\mathcal{M}$ be an $\mathcal{L}$ -structure. A subset $S \subseteq M^n$ is called a *definable set* if there exists a point $(b_1, \dots, b_m) \in M^m$ and an $\mathcal{L}$ -formula $\varphi(X_1, \dots, X_n, Y_1, \dots, Y_m)$ such that

$$S = \{(a_1, \dots, a_n) : \mathcal{M} \models \varphi(a_1, \dots, a_n, b_1, \dots, b_m)\}$$

In this case, we say that $\varphi(\bar{X}, \bar{b})$ defines the set S . We write $\varphi(M^n, \bar{b}) = S$.

Two $\mathcal{L}$ -formulas $\varphi(\bar{X})$ and $\psi(\bar{X})$ are said to be *equivalent* if for any $\mathcal{L}$ -structure $\mathcal{M}$, we have $\varphi(M^n) = \psi(M^n)$. In other words, φ and ψ are equivalent if the sets they define are equal for every structure [12, p.10].

Some forms of formulas are easier to work with than the others. The next definition introduces one such form.

Definition 3.17. [12, p.11] A formula is said to be in *negation normal form* if it is written using atomic formulas and the negations of atomic formulas together with the symbols $\exists$, $\forall$, $\wedge$ and $\vee$.

For example, $\exists X (XY = 1 \wedge X + Y = 0)$ is in negation normal form, while $\neg(\forall X XY \neq 1 \vee X + Y \neq 0)$ is not. However, these two formulas are equivalent. The next lemma shows that what we see here is in fact more general.

Lemma 3.1. [12, Lemma 1.2.14] *For every formula, there is an equivalent formula in negation normal form.*

Proof. We will proceed by induction on the total number of symbols $\wedge$, $\vee$, $\exists$ and $\forall$ used in a formula. As the basis step, we will show that any formula including only one of these symbols has an equivalent formula in negation normal form. If it is already in negation normal form, then there is nothing to prove. We only need to prove the result for formulas that are not in negation normal form.

- (i) If ψ_1 and ψ_2 are atomic formulas or negations of atomic formulas and if φ is $\neg(\psi_1 \wedge \psi_2)$, then φ is equivalent to $\neg\psi_1 \vee \neg\psi_2$, which is in negation normal form. To see the equivalence, note that if $\bar{a} \in \varphi(M^n)$ then $\mathcal{M} \not\models (\psi_1(\bar{a}) \wedge \psi_2(\bar{a}))$. This is equivalent to say that $\mathcal{M} \models \psi_1(\bar{a})$ and $\mathcal{M} \models \psi_2(\bar{a})$ cannot be both true. Again, this is equivalent to $\mathcal{M} \not\models \psi_1(\bar{a})$ or $\mathcal{M} \not\models \psi_2(\bar{a})$. The previous statement holds if and only if $\bar{a} \in (\neg\psi_1 \vee \neg\psi_2)(M^n)$.

- (ii) If ψ_1 and ψ_2 are atomic formulas or negations of atomic formulas and if φ is $\neg(\psi_1 \vee \psi_2)$, then φ is equivalent to $\neg\psi_1 \wedge \neg\psi_2$, which is in negation normal form.
- (iii) If ψ is an atomic formula or negation of an atomic formula, and if φ is $\neg(\exists X \psi)$ then φ is equivalent to $\forall X \neg\psi$, which is in negation normal form.
- (iv) If ψ is an atomic formula or negation of an atomic formula, and if φ is $\neg(\forall X \psi)$ then φ is equivalent to $\exists X \neg\psi$, which is in negation normal form.

Assume that any formula that includes m of the symbols $\wedge$, $\vee$, $\exists$ and $\forall$ has an equivalent formula in negation normal form. Suppose that φ is a formula written using $m + 1$ of these symbols.

- (i) If φ is $\neg(\varphi_1 \wedge \varphi_2)$, then it is equivalent to $\neg\varphi_1 \vee \neg\varphi_2$. Since the number of symbols in $\neg\varphi_1$ and $\neg\varphi_2$ are less than $m + 1$, they are equivalent to the formulas ψ_1 and ψ_2 that are in negation normal form, respectively. Hence, φ is equivalent to the formula $\psi_1 \vee \psi_2$ that is in negation normal form.
- (ii) If φ is $\neg(\varphi_1 \vee \varphi_2)$, then it is equivalent to $\neg\varphi_1 \wedge \neg\varphi_2$. Since the number of symbols in $\neg\varphi_1$ and $\neg\varphi_2$ are less than $m + 1$, they are equivalent to the formulas ψ_1 and ψ_2 that are in negation normal form, respectively. Hence, φ is equivalent to the formula $\psi_1 \wedge \psi_2$ that is in negation normal form.
- (iii) If φ is $\neg(\exists X \varphi_1)$, then it is equivalent to $\forall X \neg\varphi_1$. Since the number of symbols in $\neg\varphi_1$ is less than $m + 1$, it is equivalent to the formula ψ_1 that is in negation normal form. Hence, φ is equivalent to the formula $\forall X \psi_1$ that is in negation normal form.
- (iv) If φ is $\neg(\forall X \varphi_1)$, then it is equivalent to $\exists X \neg\varphi_1$. Since the number of symbols in $\neg\varphi_1$ is less than $m + 1$, it is equivalent to the formula ψ_1 that is in negation normal form. Hence, φ is equivalent to the formula $\exists X \psi_1$ that is in negation normal form.

By induction on the number of symbols, the result follows. □

The next definition names two types of formulas in negation normal form. These

formulas will be needed later on, we will see these in Theorem 3.6.

Definition 3.18. [12, Definition 1.2.15] A formula in negation normal form is called *existential* if it includes the symbol $\exists$ and does not include the symbol $\forall$. A formula in negation normal form is called *universal* if it includes the symbol $\forall$ and does not include the symbol $\exists$.

3.2. Some Results in Model Theory

In this section we will justify some important model theoretic facts, and we will prove some results that will be used in the sequel.

We will start by defining the maps between structures that preserve the interpretations.

Definition 3.19. [12, Definition 1.1.3] Let $\mathcal{L} = (\mathcal{C}, \mathcal{F}, \mathcal{R})$ be a language, $\mathcal{M}$ and $\mathcal{N}$ be two $\mathcal{L}$ -structures. A function $\eta : M \rightarrow N$ is called an $\mathcal{L}$ -homomorphism if the following conditions hold:

- (i) For every $c \in \mathcal{C}$, $\eta(c^{\mathcal{M}}) = c^{\mathcal{N}}$.
- (ii) For every $(f, n_f) \in \mathcal{F}$, $\eta(f^{\mathcal{M}}(a_1, \dots, a_{n_f})) = f^{\mathcal{N}}(\eta(a_1), \dots, \eta(a_{n_f}))$.
- (iii) For every $(R, n_R) \in \mathcal{R}$, $(a_1, \dots, a_{n_R}) \in R^{\mathcal{M}}$ implies $(\eta(a_1), \dots, \eta(a_{n_R})) \in R^{\mathcal{N}}$.

If η is injective and satisfy the condition

- (i) For every $(R, n_R) \in \mathcal{R}$, $(a_1, \dots, a_{n_R}) \in R^{\mathcal{M}}$ if and only if $(\eta(a_1), \dots, \eta(a_{n_R})) \in R^{\mathcal{N}}$,

then η is called an $\mathcal{L}$ -embedding. If an $\mathcal{L}$ -embedding η is surjective, then it is called an $\mathcal{L}$ -isomorphism.

A substructure can be defined using an $\mathcal{L}$ -embedding.

Definition 3.20. [12, Definition 1.1.5] Let $\mathcal{L}$ be a language, $\mathcal{M}$ and $\mathcal{N}$ be $\mathcal{L}$ -structures.

Suppose that $M \subseteq N$. If the inclusion map $\eta : M \rightarrow N$, defined as $\eta(m) = m$ for all $m \in M$, is an $\mathcal{L}$ -embedding, then we say that $\mathcal{M}$ is a substructure of $\mathcal{N}$.

Here is a result that works on quantifier-free formulas. We will see that quantifier-free formulas have some nice properties.

Proposition 3.1. *[11, Proposition 1.1.8] Let $\mathcal{L}$ be a language. Let $\mathcal{N}$ be an $\mathcal{L}$ -structure and $\mathcal{M}$ be a substructure of $\mathcal{N}$. Suppose that $\varphi(\bar{X})$ is a quantifier-free formula. Then for any $\bar{a} \in M^n$, $\mathcal{M} \models \varphi(\bar{a})$ if and only if $\mathcal{N} \models \varphi(\bar{a})$*

Proof. We will prove this fact by induction on formulas. This means that we will prove the statement for each possible form of formulas, from the simplest to the more complex ones.

Let $\bar{a} \in M^n$ be arbitrary. Since $\mathcal{M}$ is a substructure of $\mathcal{N}$, the inclusion map $\eta : M \rightarrow N$ is an $\mathcal{L}$ -embedding. Before starting induction on formulas, we must first show that the interpretations of terms are the same in both structures.

Claim 3.1. *For any $\mathcal{L}$ -term t and for any $\bar{a} \in M^n$, we have $t^{\mathcal{M}}(\bar{a}) = t^{\mathcal{N}}(\bar{a})$*

This will be proved by induction on terms.

(i) If $t(\bar{X})$ is a constant symbol c , then

$$t^{\mathcal{M}}(\bar{a}) = c^{\mathcal{M}} = \eta(c^{\mathcal{M}}) = c^{\mathcal{N}} = t^{\mathcal{N}}(\bar{a})$$

(ii) If $t(\bar{X})$ is a variable symbol X_i , then

$$t^{\mathcal{M}}(\bar{a}) = a_i = t^{\mathcal{N}}(\bar{a})$$

- (iii) Suppose that statement holds for terms $t_1, \dots, t_m$. If $t(\bar{X})$ is $f(t_1, \dots, t_m)$ for a function symbol f , then

$$\begin{aligned}
 t^{\mathcal{M}}(\bar{a}) &= f^{\mathcal{M}}(t_1^{\mathcal{M}}(\bar{a}), \dots, t_m^{\mathcal{M}}(\bar{a})) \\
 &= \eta(f^{\mathcal{M}}(t_1^{\mathcal{M}}(\bar{a}), \dots, t_m^{\mathcal{M}}(\bar{a}))) \\
 &= f^{\mathcal{N}}(t_1^{\mathcal{N}}(\bar{a}), \dots, t_m^{\mathcal{N}}(\bar{a})) \\
 &= t^{\mathcal{N}}(\bar{a})
 \end{aligned}$$

We showed that $\mathcal{L}$ -terms have the same interpretations over $\mathcal{M}$ and $\mathcal{N}$.

Now we can start the induction on formulas. Let $\varphi(\bar{X})$ be a quantifier-free formula. Assume that $\mathcal{M} \models \varphi(\bar{a})$.

- (i) If φ is $t_1 = t_2$, then

$$\mathcal{M} \models \varphi(\bar{a}) \Leftrightarrow t_1^{\mathcal{M}}(\bar{a}) = t_2^{\mathcal{M}}(\bar{a}) \Leftrightarrow t_1^{\mathcal{N}}(\bar{a}) = t_2^{\mathcal{N}}(\bar{a}) \Leftrightarrow \mathcal{N} \models \varphi(\bar{a})$$

- (ii) If φ is $R(t_1, \dots, t_m)$, then

$$\begin{aligned}
 \mathcal{M} \models \varphi(\bar{a}) &\Leftrightarrow (t_1^{\mathcal{M}}(\bar{a}), \dots, t_m^{\mathcal{M}}(\bar{a})) \in R^{\mathcal{M}} \\
 &\Leftrightarrow \eta(t_1^{\mathcal{M}}(\bar{a}), \dots, t_m^{\mathcal{M}}(\bar{a})) \in R^{\mathcal{N}} \\
 &\Leftrightarrow (t_1^{\mathcal{N}}(\bar{a}), \dots, t_m^{\mathcal{N}}(\bar{a})) \in R^{\mathcal{N}} \\
 &\Leftrightarrow \mathcal{N} \models \varphi(\bar{a})
 \end{aligned}$$

- (iii) Suppose that the statement holds for ψ and suppose that φ is $\neg\psi$. Then

$$\mathcal{M} \models \varphi(\bar{a}) \Leftrightarrow \mathcal{M} \not\models \psi(\bar{a}) \Leftrightarrow \mathcal{N} \not\models \psi(\bar{a}) \Leftrightarrow \mathcal{N} \models \varphi(\bar{a})$$

(iv) Suppose that the statement holds for ψ_1 and ψ_2 , and suppose that φ is $\psi_1 \wedge \psi_2$.

Then

$$\begin{aligned}
 \mathcal{M} \models \varphi(\bar{a}) &\Leftrightarrow \mathcal{M} \models (\psi_1(\bar{a}) \wedge \psi_2(\bar{a})) \\
 &\Leftrightarrow \mathcal{M} \models \psi_1(\bar{a}) \text{ and } \mathcal{M} \models \psi_2(\bar{a}) \\
 &\Leftrightarrow \mathcal{N} \models \psi_1(\bar{a}) \text{ and } \mathcal{N} \models \psi_2(\bar{a}) \\
 &\Leftrightarrow \mathcal{N} \models (\psi_1(\bar{a}) \wedge \psi_2(\bar{a})) \\
 &\Leftrightarrow \mathcal{N} \models \varphi(\bar{a})
 \end{aligned}$$

These cover all the possible cases for a quantifier-free formula. Hence for any quantifier-free formula $\varphi(\bar{X})$ and for any $\bar{a} \in M^n$, $\mathcal{M} \models \varphi(\bar{a})$ if and only if $\mathcal{N} \models \varphi(\bar{a})$. $\square$

There are theories in which every formula is equivalent to a quantifier-free formula. This property is given the following name.

Definition 3.21. [11, Definition 3.1.1] Let T be a theory. If for every formula $\varphi(\bar{X})$ there exists a quantifier-free formula $\psi(\bar{X})$ with $T \models \forall \bar{X} (\varphi(\bar{X}) \leftrightarrow \psi(\bar{X}))$, then we say that T has *quantifier elimination*.

Definition 3.22. [12, p.11] Let $\mathcal{L} = (\mathcal{C}, \mathcal{F}, \mathcal{R})$ be a language, $\mathcal{M}$ be an $\mathcal{L}$ -structure. Let $\mathcal{L}(M)$ be the language $(\mathcal{C} \cup M, \mathcal{F}, \mathcal{R})$, that is the language obtained from $\mathcal{L}$ by adding each $m \in M$ as a constant symbol. The *atomic diagram* of $\mathcal{M}$ is the set $\text{Diag}(\mathcal{M}) = \{\phi(m_1, \dots, m_n) : \phi \text{ is either an atomic } \mathcal{L}_M\text{-formula without free variables or the negation of an atomic } \mathcal{L}_M\text{-formula without free variables and } \mathcal{M} \models \phi(m_1, \dots, m_n)\}$.

Lemma 3.2. *Let $\mathcal{M}$ be an $\mathcal{L}$ -structure and $\mathcal{N}$ be an $\mathcal{L}$ -structure with $\mathcal{N} \models \text{Diag}(\mathcal{M})$. Then there is an $\mathcal{L}$ -embedding from $\mathcal{M}$ into $\mathcal{N}$.*

Proof. Define $j : M \rightarrow N$ to be the function $j(m) = m^{\mathcal{N}}$ for each $m \in M$.

To show that j is one-to-one, let $m_1, m_2 \in M$ be distinct elements. Therefore

we have $M \models m_1 \neq m_2$ and hence $m_1 \neq m_2 \in \text{Diag}(\mathcal{M})$. Since $\mathcal{N} \models \text{Diag}(\mathcal{M})$ by assumption, $m_1^{\mathcal{N}} \neq m_2^{\mathcal{N}}$. Therefore $j(m_1) \neq j(m_2)$.

Now, we will prove that j is an $\mathcal{L}$ -embedding.

Let f be a function symbol in $\mathcal{L}$. Then $f^{\mathcal{M}}$ is a function from M^n to M . Let $m_1, \dots, m_n \in M$ be arbitrary. Then there exists $m_{n+1} \in M$ with $f^{\mathcal{M}}(m_1, \dots, m_n) = m_{n+1}$. Thus, $f(m_1, \dots, m_n) = m_{n+1} \in \text{Diag}(\mathcal{M})$. Since $\mathcal{N} \models \text{Diag}(\mathcal{M})$ by assumption, we have $f^{\mathcal{N}}(m_1^{\mathcal{N}}, \dots, m_n^{\mathcal{N}}) = m_{n+1}^{\mathcal{N}}$, i.e. $j(f^{\mathcal{M}}(m_1, \dots, m_n)) = f^{\mathcal{N}}(j(m_1), \dots, j(m_n))$.

Let R be a relation symbol in $\mathcal{L}$. Suppose that $\bar{m} \in R^{\mathcal{M}}$. Then, $R(\bar{m}) \in \text{Diag}(\mathcal{M})$ and by assumption $\mathcal{N} \models R^{\mathcal{N}}(\bar{m}^{\mathcal{N}})$. Therefore, $j(\bar{m}) \in R^{\mathcal{N}}$.

Let c be a constant symbol in $\mathcal{L}$. Then $c^{\mathcal{M}}$ is an element $m \in M$, which is a constant symbol in $\mathcal{L}$. Therefore $c = m \in \text{Diag}(\mathcal{M})$. Hence $\mathcal{N} \models c^{\mathcal{N}} = m^{\mathcal{N}}$. Therefore $j(c^{\mathcal{M}}) = c^{\mathcal{N}}$.

The function j defined above is an $\mathcal{L}$ -embedding. □

Theorem 3.1. [11, Theorem 3.1.4] *Let $\mathcal{L}$ be a language containing a constant symbol c . Let T be an $\mathcal{L}$ -theory and $\phi(\bar{v})$ be an $\mathcal{L}$ -formula. The following are equivalent:*

- (i) $\phi(\bar{v})$ is equivalent to a quantifier-free $\mathcal{L}$ -formula $\psi(\bar{v})$
- (ii) For any two models $\mathcal{M}$ and $\mathcal{N}$ of T having a common substructure $\mathcal{A}$, we have $\mathcal{M} \models \phi(\bar{a})$ if and only if $\mathcal{N} \models \phi(\bar{a})$ for all $\bar{a} \in \mathcal{A}$.

Proof. We will first show that (i) $\Rightarrow$ (ii). Assume that $\psi(\bar{v})$ is a quantifier-free formula such that $T \models \forall \bar{v} (\phi(\bar{v}) \leftrightarrow \psi(\bar{v}))$. Let $\mathcal{M}$ and $\mathcal{N}$ be models of T , $\mathcal{A}$ be an $\mathcal{L}$ -structure, $\mathcal{A} \subseteq \mathcal{M}$, and $\mathcal{A} \subseteq \mathcal{N}$. Let $\bar{a} \in \mathcal{A}$ be arbitrary.

Suppose that $\mathcal{M} \models \phi(\bar{a})$. By assumption, this is equivalent to $\mathcal{M} \models \psi(\bar{a})$. By Proposition 3.1, this is true if and only if $\mathcal{A} \models \psi(\bar{a})$. Again, by Proposition 3.1, we

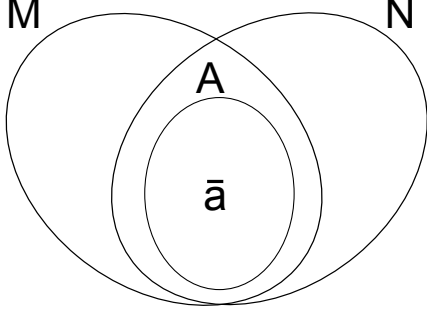


Figure 3.1. $\mathcal{A}$ is a substructure of both $\mathcal{M}$ and $\mathcal{N}$.

have $\mathcal{A} \models \psi(\bar{a})$ if and only if $\mathcal{N} \models \psi(\bar{a})$. Finally, by assumption, this is equivalent to say that $\mathcal{N} \models \phi(\bar{a})$.

Now, we will show that $(ii) \Rightarrow (i)$.

If $T \models \forall \bar{v} \phi(\bar{v})$, then $T \models \forall \bar{v} (\phi(\bar{v}) \leftrightarrow c = c)$. Since $c = c$ is a quantifier-free formula, we are done.

If $T \models \forall \bar{v} \neg \phi(\bar{v})$, then $T \models \forall \bar{v} (\phi(\bar{v}) \leftrightarrow c \neq c)$. Again, since $c \neq c$ is a quantifier-free formula, we are done.

Now, suppose that neither is the case. Then, T satisfies both $\exists \bar{v} \phi(\bar{v})$ and $\exists \bar{v} \neg \phi(\bar{v})$. In other words, both $T \cup \{\exists \bar{v} \phi(\bar{v})\}$ and $T \cup \{\exists \bar{v} \neg \phi(\bar{v})\}$ have a model.

Define the set $\Gamma(\bar{v}) := \{\psi(\bar{v}) : \psi \text{ is quantifier-free and } T \models \forall \bar{v} (\phi(\bar{v}) \rightarrow \psi(\bar{v}))\}$. Let $d_1, \dots, d_m$ be new constant symbols, $\bar{d} = (d_1, \dots, d_m)$.

Claim 3.2. $T \cup \Gamma(\bar{d}) \models \phi(\bar{d})$

Suppose on the contrary that $\mathcal{M} \models T \cup \Gamma(\bar{d}) \cup \{\neg \phi(\bar{d})\}$. Let $\mathcal{A}$ be the substructure of $\mathcal{M}$ generated by $\bar{d}$.

Let $\Sigma = T \cup \text{Diag}(\mathcal{A}) \cup \phi(\bar{d})$. We assumed that $T \cup \{\neg\phi(\bar{d})\}$ is satisfiable. Also, $\mathcal{M} \models T \cup \text{Diag}(\mathcal{A})$. Therefore, if Σ is unsatisfiable, then there are quantifier-free formulas $\psi_1(\bar{d}), \dots, \psi_n(\bar{d}) \in \text{Diag}(\mathcal{A})$ such that

$$T \models \forall \bar{v} \left(\bigwedge_{i=1}^n \psi_i(\bar{v}) \rightarrow \neg\phi(\bar{v}) \right)$$

or equivalently,

$$T \models \forall \bar{v} \left(\phi(\bar{v}) \rightarrow \bigvee_{i=1}^n \neg\psi_i(\bar{v}) \right)$$

Since $\bigvee_{i=1}^n \neg\psi_i(\bar{v})$ is quantifier-free, it is in $\Gamma(\bar{v})$. By Proposition 3.1, $\mathcal{A} \models \bigvee_{i=1}^n \neg\psi_i(\bar{v})$, contradicting $\psi_1(\bar{d}), \dots, \psi_n(\bar{d}) \in \text{Diag}(\mathcal{A})$. Hence, Σ must be satisfiable.

Suppose that $\mathcal{N} \models \Sigma$. Then $\mathcal{N} \models \phi(\bar{d})$. By Lemma 3.2, since $\mathcal{N} \models \text{Diag}(\mathcal{A})$, we have $\mathcal{A} \subseteq \mathcal{N}$. By our assumption we have $\mathcal{M} \models \neg\phi(\bar{d})$ and by (ii) we get $\mathcal{N} \models \neg\phi(\bar{d})$, contradicting $\mathcal{N} \models \phi(\bar{d})$. $\square$

We now extend our discussion at the end of the previous section. Each quantifier-free formula is equivalent in the theory of rings to a conjunction of the following form:

$$\left(\bigwedge_{i=1}^n f_i(\bar{X}) = 0 \right) \wedge \left(\bigwedge_{i=1}^m g_i(\bar{X}) \neq 0 \right)$$

where $f_i(\bar{X}), g_i(\bar{X}) \in \mathbb{Z}[\bar{X}]$. This is easy to see, as any quantifier-free formula is obtained from atomic formulas by negation and conjunction operations.

The following theorem is an important result in model theory and is needed in

the proof of the main theorem.

Theorem 3.2 (Compactness Theorem). *[11, Theorem 2.1.4] A theory T has a model if and only if every finite subset of T has a model.*

This theorem will be proved in Section 3.3 using ultraproducts.

3.3. Ultrafilters and Ultraproducts

This section introduces ultrafilters and ultraproducts. At the end of the section we will prove Łoś's theorem and the compactness theorem.

We start by defining filters.

Definition 3.23. [7, 7.5] A nonempty collection $\mathcal{F}$ of subsets of a set I is called a filter on I if the following conditions hold:

- (i) $\emptyset \notin \mathcal{F}$
- (ii) if $A, B \in \mathcal{F}$, then $A \cap B \in \mathcal{F}$
- (iii) if $A \in \mathcal{F}$, $B \subseteq I$, and $A \subseteq B$, then $B \in \mathcal{F}$.

I is called an *index set*.

Let P be a nonempty subset of $\mathbb{N}$. The set $\mathcal{F} := \{S \subseteq \mathbb{N} : P \subseteq S\}$ is a filter. First condition holds as $P \not\subseteq \emptyset$. For the second condition, suppose that $A, B \in \mathcal{F}$. Then $P \subseteq A$ and $P \subseteq B$. Therefore $P \subseteq (A \cap B)$ and hence $(A \cap B) \in \mathcal{F}$. Finally, for any $A \in \mathcal{F}$ and for any $B \subseteq \mathbb{N}$ with $A \subseteq B$, we have $P \subseteq B$, so $B \in \mathcal{F}$.

The set $\mathcal{G} := \{S \subseteq \mathbb{N} : \text{card}(\mathbb{N} \setminus S) < \infty\}$ is also a filter on $\mathbb{N}$. $\text{card}(\mathbb{N} \setminus \emptyset)$ is not finite, so $\emptyset \notin \mathcal{G}$. Let $A, B \in \mathcal{G}$. Then $\mathbb{N} \setminus A$ and $\mathbb{N} \setminus B$ are finite sets. With this fact and the fact that $\mathbb{N} \setminus (A \cap B) = (\mathbb{N} \setminus A) \cup (\mathbb{N} \setminus B)$, we get $\text{card}(\mathbb{N} \setminus (A \cap B)) < \infty$ and thus $(A \cap B) \in \mathcal{G}$. Now fix $A \in \mathcal{G}$ and $B \subseteq \mathbb{N}$ with $A \subseteq B$. Then $(\mathbb{N} \setminus B) \subseteq (\mathbb{N} \setminus A)$. Since $A \in \mathcal{G}$, $\mathbb{N} \setminus A$ is finite. Therefore $\mathbb{N} \setminus B$ is finite, too. Thus $B \in \mathcal{G}$.

The following lemma gives a way to construct a larger filter from a given filter. This will be needed in Lemma 3.4 to prove a maximality result.

Lemma 3.3. *Let $\mathcal{F}$ be a filter on I . Let $X \subseteq I$ with $X \notin \mathcal{F}$. Define $\mathcal{G} := \{S \subseteq I : (Z \setminus X) \subseteq S \text{ for some } Z \in \mathcal{F}\}$. Then $\mathcal{G}$ is a filter, $\mathcal{F} \subseteq \mathcal{G}$ and $I \setminus X \in \mathcal{G}$.*

Proof. Fix a filter $\mathcal{F}$ and a set $X \subseteq I$ with $X \notin \mathcal{F}$. Since $\mathcal{F} \neq \emptyset$, there is $A \in \mathcal{F}$. $(A \setminus X) \subseteq A$, hence $A \in \mathcal{G}$. This proves that $\mathcal{F} \subseteq \mathcal{G}$ and $\mathcal{G} \neq \emptyset$.

Now, suppose that $\emptyset \in \mathcal{G}$. Then, there is $Z \in \mathcal{F}$ such that $Z \setminus X \subseteq \emptyset$, which implies $Z \subseteq X$. Since $\mathcal{F}$ is a filter and $Z \in \mathcal{F}$, we have $X \in \mathcal{F}$, contradicting our assumption. Therefore, $\emptyset \notin \mathcal{G}$.

Let $A, B \in \mathcal{G}$. Then there are $Z, W \in \mathcal{F}$ such that $Z \setminus X \subseteq A$ and $W \setminus X \subseteq B$. We have $A \cap B \supseteq (Z \setminus X) \cap (W \setminus X) = (Z \cap W) \setminus X$. Since $Z, W \in \mathcal{F}$ and since $\mathcal{F}$ is a filter we have $Z \cap W \in \mathcal{F}$. Hence, $A \cap B \in \mathcal{G}$.

Finally, suppose that $A \in \mathcal{G}$ and $A \subseteq B$. There exists a $Z \in \mathcal{F}$ such that $Z \setminus X \subseteq A \subseteq B$. Thus, $B \in \mathcal{G}$. This proves that $\mathcal{G}$ is a filter.

Let $A \in \mathcal{F}$. Then $A \subseteq I$ and by definition $I \in \mathcal{F}$. Since $I \in \mathcal{F}$ and $I \setminus X \subseteq I \setminus X$, we have $I \setminus X \in \mathcal{G}$. □

We now define ultrafilters, which allow us to define ultraproducts.

Definition 3.24. A filter $\mathcal{U}$ on I is called an ultrafilter if for every $S \subseteq I$ either $S \in \mathcal{U}$ or $I \setminus S \in \mathcal{U}$.

Let for any subset S of the index set I , S^C denote the set $I \setminus S$. Our previous example $\mathcal{F} := \{S \subseteq \mathbb{N} : P \subseteq S\}$ becomes an ultrafilter on $\mathbb{N}$ if we choose P to be a singleton $\{a\}$ for some $a \in \mathbb{N}$. Let $S \subseteq \mathbb{N}$ be arbitrary. If $a \in S$, then $S \in \mathcal{F}$. If $a \notin S$, then $a \in S^C$ and $S^C \in \mathcal{F}$. These filters are called *principal ultrafilters*.

The following theorem gives an equivalent definition for ultrafilters.

Theorem 3.3. *$\mathcal{U}$ is an ultrafilter on I if and only if for any $A, B \subseteq I$, if $A \cup B \in \mathcal{U}$ then $A \in \mathcal{U}$ or $B \in \mathcal{U}$.*

Proof. Let $\mathcal{U}$ be an ultrafilter on I . Let $A, B \subseteq I$ be arbitrary, such that $A \cup B \in \mathcal{U}$.

If $A \in \mathcal{U}$, we are done. Suppose now that $A \notin \mathcal{U}$. Then $A^C \in \mathcal{U}$ since $\mathcal{U}$ is an ultrafilter. So, $(A \cup B) \cap A^C = B \cap A^C \in \mathcal{U}$. Since $B \cap A^C \subseteq B$, we have $B \in \mathcal{U}$.

Now, assume that $A \cup B \in \mathcal{U} \Rightarrow A \in \mathcal{U}$ or $B \in \mathcal{U}$. Let $S \subseteq I$ be arbitrary. $I = S \cup S^C$ and $I \in \mathcal{U}$ as we have seen in the previous proof. Then by assumption either S or S^C is in $\mathcal{U}$. $\square$

For a fixed set I , the set of filters on I is a partially ordered set with the subset relation. The next lemma shows that ultraproducts are the maximal elements of this poset.

Lemma 3.4. *[7, Lemma 7.5.2] A filter $\mathcal{U}$ on I is an ultrafilter if and only if it is maximal with respect to $\subseteq$.*

Proof. Let $\mathcal{U}$ be an ultrafilter on I . Assume that $\mathcal{F}$ is a filter on I with $\mathcal{U} \subseteq \mathcal{F}$. Let $A \in \mathcal{F}$. Since $\mathcal{U}$ is an ultrafilter, either A or A^C is in $\mathcal{U}$. Suppose that $A^C \in \mathcal{U}$. Then $A^C \in \mathcal{F}$. However, $\mathcal{F}$ is a filter and hence $A \cap A^C$ should be in $\mathcal{F}$. This is impossible as $\emptyset \notin \mathcal{F}$. Hence $A \in \mathcal{U}$. This shows that $\mathcal{U} = \mathcal{F}$.

Assume for the converse that $\mathcal{U}$ is maximal. Suppose that $A \notin \mathcal{U}$. By Lemma 3.3, the filter $\mathcal{F}$ defined as the set $\{S \subseteq I : (Z \setminus A) \subseteq S \text{ for some } Z \in \mathcal{U}\}$ contains $\mathcal{U}$. Since $\mathcal{U}$ is maximal, $\mathcal{U} = \mathcal{F}$. Moreover, $A^C \in \mathcal{U}$ and hence $\mathcal{U}$ is an ultrafilter. $\square$

We now move on to construct ultraproducts. Let I be an infinite set, and let A_i be a set for each $i \in I$. Suppose that $\mathcal{U}$ is an ultrafilter on I . Consider the set

$\prod_{i \in I} A_i := \{(a_i)_{i \in I} : a_i \in A_i\}$. Define a relation on $\prod_{i \in I} A_i$ as follows

$$(a_i) \sim (b_i) \Leftrightarrow \{i \in I : a_i = b_i\} \in \mathcal{U}$$

We now show that this is an equivalence relation on $\prod_{i \in I} A_i$. Let $(a_i)_{i \in I}$, $(b_i)_{i \in I}$, and $(c_i)_{i \in I}$ be elements of $\prod_{i \in I} A_i$.

- (i) $\{i \in I : a_i = a_i\} = I \in \mathcal{U}$, so $(a_i)_{i \in I} \sim (a_i)_{i \in I}$.
- (ii) $(a_i)_{i \in I} \sim (b_i)_{i \in I}$ implies that $\{i \in I : a_i = b_i\} \in \mathcal{U}$. Hence $(b_i)_{i \in I} \sim (a_i)_{i \in I}$.
- (iii) $(a_i)_{i \in I} \sim (b_i)_{i \in I}$ and $(b_i)_{i \in I} \sim (c_i)_{i \in I}$ implies that $\{i \in I : a_i = b_i\} \in \mathcal{U}$ and $\{i \in I : b_i = c_i\} \in \mathcal{U}$. Then, the set $\{i \in I : a_i = b_i\} \cap \{i \in I : b_i = c_i\} \in \mathcal{U}$. Moreover, $\{i \in I : a_i = b_i\} \cap \{i \in I : b_i = c_i\} \subseteq \{i \in I : a_i = c_i\}$. Therefore $\{i \in I : a_i = c_i\} \in \mathcal{U}$ and hence $(a_i)_{i \in I} \sim (c_i)_{i \in I}$.

The ultraproduct of A_i can now be defined.

Definition 3.25. The *ultraproduct* of the collection $\{A_i : i \in I\}$ with respect to $\mathcal{U}$ is the set of equivalence classes of $\prod_{i \in I} A_i$ with respect to the equivalence relation $\sim$ defined as

$$(a_i) \sim (b_i) \Leftrightarrow \{i \in I : a_i = b_i\} \in \mathcal{U}$$

This ultraproduct will be denoted by $\prod_{i \in I} A_i / \mathcal{U}$.

Fix a language $\mathcal{L}$. Let $\mathcal{M}_i$ be an $\mathcal{L}$ -structure for all i in the infinite set I . Suppose that $\mathcal{U}$ is an ultrafilter on I . We can define the ultraproduct of the collection $\{\mathcal{M}_i\}$ with respect to $\mathcal{U}$ in the following manner to make it into an $\mathcal{L}$ -structure $\mathcal{M}$ [7, p.142]. Let $M := \prod_{i \in I} M_i / \mathcal{U}$, this will be the underlying set for our structure. We now define the interpretations of the elements in the language.

- (i) If c is a constant symbol in $\mathcal{L}$, then $c^{\mathcal{M}}$ is the equivalence class of the element $(c^{\mathcal{M}_i})_{i \in I}$.
- (ii) If f is an n -ary function symbol in $\mathcal{L}$, then $f^{\mathcal{M}}$ is a function on M such that

$$f^{\mathcal{M}}((a_{1i})_{i \in I} / \sim, \dots, (a_{ni})_{i \in I} / \sim) = (f^{\mathcal{M}_i}(a_{1i}, \dots, a_{ni}))_{i \in I} / \sim$$

where $a_{ji} \in M_i$ for all $i \in I$, $(a_{ji})_{i \in I} / \sim$ denotes the equivalence class of the element $(a_{ji})_{i \in I}$ and $(f^{\mathcal{M}_i}(a_{1i}, \dots, a_{ni}))_{i \in I} / \sim$ denotes the equivalence class of the element $(f^{\mathcal{M}_i}(a_{1i}, \dots, a_{ni}))_{i \in I}$ for $1 \leq j \leq n$.

- (iii) If R is an n -ary relation symbol in $\mathcal{L}$, then $R^{\mathcal{M}}$ is the relation defined on M as

$$((a_{1i})_{i \in I} / \sim, \dots, (a_{ni})_{i \in I} / \sim) \in R^{\mathcal{M}} \Leftrightarrow \{i \in I : (a_{1i}, \dots, a_{ni}) \in R^{\mathcal{M}_i}\} \in \mathcal{U}$$

where $a_{ji} \in M_i$ for all $i \in I$ and $(a_{ji})_{i \in I} / \sim$ denotes the equivalence class of the element $(a_{ji})_{i \in I}$ where $1 \leq j \leq n$.

In this case, $\mathcal{M}$ is an $\mathcal{L}$ -structure. $\mathcal{M}$ will be denoted by $\prod_{i \in I} \mathcal{M}_i / \mathcal{U}$.

Here is the statement and proof of Łoś's theorem.

Theorem 3.4 (Łoś's Theorem). *[11, Exercise 2.5.19] Let $\mathcal{L}$ be a language, $\mathcal{M}_i$ be an $\mathcal{L}$ -structure for each i in the infinite set I . Suppose that $\mathcal{U}$ is an ultrafilter on I and $\mathcal{M}$ is the ultraproduct $\prod_{i \in I} \mathcal{M}_i / \mathcal{U}$. Let $\varphi(\bar{X})$ be an $\mathcal{L}$ -formula with $\bar{X} = (X_1, \dots, X_n)$. Then, for all $i \in I$ and for all $a_{1i}, \dots, a_{ni} \in M_i$, $\mathcal{M} \models \varphi((a_{1i})_{i \in I} / \sim, \dots, (a_{ni})_{i \in I} / \sim)$ if and only if $\{i \in I : \mathcal{M}_i \models \varphi(a_{1i}, \dots, a_{ni})\} \in \mathcal{U}$.*

Proof. Let $\bar{a}$ denote the tuple $((a_{1i})_{i \in I} / \sim, \dots, (a_{ni})_{i \in I} / \sim) \in M^n$, and $\bar{a}_i$ denote the tuple $(a_{1i}, \dots, a_{ni}) \in M_i^n$.

(i) Suppose that φ is $t_1(\bar{X}) = t_2(\bar{X})$ for some terms t_1 and t_2 .

$$\begin{aligned}\mathcal{M} \models \varphi(\bar{a}) &\Leftrightarrow t_1^{\mathcal{M}}(\bar{a}) = t_2^{\mathcal{M}}(\bar{a}) \\ &\Leftrightarrow \{i \in I : t_1^{\mathcal{M}_i}(\bar{a}_i) = t_2^{\mathcal{M}_i}(\bar{a}_i)\} \in \mathcal{U} \\ &\Leftrightarrow \{i \in I : \mathcal{M}_i \models \varphi(\bar{a}_i)\} \in \mathcal{U}\end{aligned}$$

(ii) Suppose that φ is $R(t_1(\bar{X}), \dots, t_m(\bar{X}))$ for some terms $t_1, \dots, t_m$ and a relation symbol R .

$$\begin{aligned}\mathcal{M} \models \varphi(\bar{a}) &\Leftrightarrow (t_1^{\mathcal{M}}(\bar{a}), \dots, t_m^{\mathcal{M}}(\bar{a})) \in R^{\mathcal{M}} \\ &\Leftrightarrow \{i \in I : (t_1^{\mathcal{M}_i}(\bar{a}_i), \dots, t_m^{\mathcal{M}_i}(\bar{a}_i)) \in R^{\mathcal{M}_i}\} \in \mathcal{U} \\ &\Leftrightarrow \{i \in I : \mathcal{M}_i \models \varphi(\bar{a}_i)\} \in \mathcal{U}\end{aligned}$$

(iii) Suppose that φ is $\neg\psi$ for some ψ satisfying the theorem. Then

$$\begin{aligned}\mathcal{M} \models \varphi(\bar{a}) &\Leftrightarrow \mathcal{M} \not\models \psi(\bar{a}) \\ &\Leftrightarrow \{i \in I : \mathcal{M}_i \models \psi(\bar{a}_i)\} \notin \mathcal{U} \\ &\Leftrightarrow \{i \in I : \mathcal{M}_i \not\models \psi(\bar{a}_i)\} \in \mathcal{U} \\ &\Leftrightarrow \{i \in I : \mathcal{M}_i \models \varphi(\bar{a}_i)\} \in \mathcal{U}\end{aligned}$$

(iv) Suppose that φ is $\psi_1 \wedge \psi_2$ for some ψ_1 and ψ_2 satisfying the theorem. Then

$$\begin{aligned}
\mathcal{M} \models \varphi(\bar{a}) &\Leftrightarrow \mathcal{M} \models \psi_1(\bar{a}) \text{ and } \mathcal{M} \models \psi_2(\bar{a}) \\
&\Leftrightarrow \{i \in I : \mathcal{M}_i \models \psi_1(\bar{a}_i)\} \in \mathcal{U} \text{ and } \{i \in I : \mathcal{M}_i \models \psi_2(\bar{a}_i)\} \in \mathcal{U} \\
&\Leftrightarrow \{i \in I : \mathcal{M}_i \models \psi_1(\bar{a}_i)\} \cap \{i \in I : \mathcal{M}_i \models \psi_2(\bar{a}_i)\} \in \mathcal{U} \\
&\Leftrightarrow \{i \in I : \mathcal{M}_i \models \psi_1(\bar{a}_i) \text{ and } \mathcal{M}_i \models \psi_2(\bar{a}_i)\} \in \mathcal{U} \\
&\Leftrightarrow \{i \in I : \mathcal{M}_i \models (\psi_1(\bar{a}_i) \wedge \psi_2(\bar{a}_i))\} \in \mathcal{U} \\
&\Leftrightarrow \{i \in I : \mathcal{M}_i \models \varphi(\bar{a}_i)\} \in \mathcal{U}
\end{aligned}$$

(v) Suppose that $\varphi(\bar{X})$ is $\exists Y \psi(\bar{X}, Y)$ for some ψ satisfying the formula. Then

$$\begin{aligned}
\mathcal{M} \models \varphi(\bar{a}) &\Leftrightarrow \text{There exists } (b_i)_{i \in I} / \sim \in M \text{ such that } \mathcal{M} \models \psi(\bar{a}, (b_i)_{i \in I} / \sim) \\
&\Leftrightarrow \text{There exists } (b_i)_{i \in I} / \sim \in M \text{ such that } \{i \in I : \mathcal{M}_i \models \psi(\bar{a}_i, b_i)\} \in \mathcal{U}
\end{aligned}$$

First, let $(b_i^0)_{i \in I} / \sim \in M$ be such that $\{i \in I : \mathcal{M}_i \models \psi(\bar{a}_i, b_i^0)\} \in \mathcal{U}$. Let $j \in \{i \in I : \mathcal{M}_i \models \psi(\bar{a}_i, b_i^0)\}$ be arbitrary. Then $\mathcal{M}_j \models \psi(\bar{a}_j, b_j^0)$. Therefore, $j \in \{i \in I : \text{There exists } b_i \in M_i \text{ such that } \mathcal{M}_i \models \psi(\bar{a}_i, b_i)\}$. This shows that $\{i \in I : \mathcal{M}_i \models \psi(\bar{a}_i, b_i^0)\} \subseteq \{i \in I : \mathcal{M}_i \models \varphi(\bar{a}_i)\}$. Since $\mathcal{U}$ is an ultrafilter, $\{i \in I : \mathcal{M}_i \models \varphi(\bar{a}_i)\} \in \mathcal{U}$.

Now put $A := \{i \in I : \mathcal{M}_i \models \varphi(\bar{a}_i)\}$, assume that $A \in \mathcal{U}$. Note that

$$A = \{i \in I : \text{There exists } b_i \in M_i \text{ such that } \mathcal{M}_i \models \psi(\bar{a}_i, b_i)\}$$

Now construct an element $(b_i^0)_{i \in I}$ in M such that if $j \in A$, then choose b_j^0 to be an element satisfying $\mathcal{M}_j \models \psi(\bar{a}_j, b_j^0)$. If $j \notin A$, then choose $b_j^0 \in M_j$ arbitrarily. Then, the set $\{i \in I : \mathcal{M}_i \models \psi(\bar{a}_i, b_i^0)\}$ includes all the elements of A . Therefore, $\{i \in I : \mathcal{M}_i \models \psi(\bar{a}_i, b_i^0)\} \in \mathcal{U}$. By the above arguments, $\mathcal{M} \models \varphi(\bar{a})$ holds.

By induction on formulas, we see that $\mathcal{M} \models \varphi((a_{1i})_{i \in I} / \sim, \dots, (a_{ni})_{i \in I} / \sim)$ if and only if $\{i \in I : \mathcal{M}_i \models \varphi(a_{1i}, \dots, a_{ni})\} \in \mathcal{U}$. $\square$

We need the following lemma before proving the compactness theorem.

Lemma 3.5. *[11, Exercise 2.5.20.(a)] Let T be an infinite theory such that each finite subset of T has a model. Let I be the set of finite subsets of T . Put $\tilde{T} := \{\psi_1 \wedge \dots \wedge \psi_n : n \in \mathbb{N} \text{ and } \psi_1, \dots, \psi_n \in T\}$. For each $\varphi \in \tilde{T}$, define S_φ to be the set of finite subsets of T that satisfy φ . Then the set $\mathcal{F} := \{Y \subseteq I : S_\varphi \subseteq Y \text{ for some } \varphi \in \tilde{T}\}$ is a filter on I .*

Proof. Since T is infinite, the set I is also infinite. We will start the proof by showing that $\mathcal{F}$ is not empty. Let $\varphi \in \tilde{T}$. Then $S_\varphi \in \mathcal{F}$.

Now, let A and B be two sets in $\mathcal{F}$. Then there are φ and $\psi \in \tilde{T}$ such that $S_\varphi \subseteq A$ and $S_\psi \subseteq B$. Then $S_\varphi \cap S_\psi \subseteq A \cap B$. The set $S_\varphi \cap S_\psi$ is equal to the set $S_{\varphi \wedge \psi}$. Since $\varphi \wedge \psi \in \tilde{T}$, we have $A \cap B \in \mathcal{F}$.

Finally, let $A \in \mathcal{F}$ and let $B \subseteq I$ be such that $A \subseteq B$. Then there is $\varphi \in \tilde{T}$ with $S_\varphi \subseteq A$. Hence $S_\varphi \subseteq B$ and so $B \in \mathcal{F}$. Therefore, $\mathcal{F}$ is a filter on I . $\square$

We now prove the compactness theorem with the help of above results.

Theorem 3.5 (Compactness Theorem). *[11, Theorem 2.1.4] A theory T has a model if and only if every finite subset of T has a model.*

Proof. If $\mathcal{M} \models T$, then every subset $S \subseteq T$ has a model, hence the finite subsets of T have a model.

Assume that every finite subset of T has a model. If T is finite, then it is a finite subset of itself and therefore, has a model. Suppose now that T is infinite. By Lemma 3.5, $\mathcal{F} := \{Y \subseteq I : S_\varphi \subseteq Y \text{ for some } \varphi \in \tilde{T}\}$ is a filter on $I := \{S : S \subseteq T, S \text{ is finite}\}$,

where $S_\varphi := \{S \in I : S \models \varphi\}$. By Zorn's Lemma, we see that there is an ultrafilter $\mathcal{U}$ that contains $\mathcal{F}$. To see this, let P be the set of filters on I containing $\mathcal{F}$. Let $\mathcal{F} \subseteq \mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \dots$ be a chain in P . We must show that this chain has an upper bound. Consider the set $\mathcal{G} := \bigcup_i \mathcal{F}_i$. This set is a filter. As for each i we have $\emptyset \notin \mathcal{F}_i$, we get $\emptyset \notin \mathcal{G}$. If A and B are in $\mathcal{G}$, then there is a j such that $A, B \in \mathcal{F}_j$ so $A \cap B \in \mathcal{F}_j$. Therefore $A \cap B \in \mathcal{G}$. Lastly, let $A \in \mathcal{G}$ and $A \subseteq B \subseteq I$. Then there is j with $A \in \mathcal{F}_j$. Since $\mathcal{F}_j$ is a filter, we have $B \in \mathcal{F}_j$ and thus $B \in \mathcal{G}$. By Zorn's Lemma, P has a maximal element $\mathcal{U}$. By Lemma 3.4, $\mathcal{U}$ is an ultrafilter. Since $\mathcal{U} \in P$, $\mathcal{U}$ contains $\mathcal{F}$.

For each $S \in I$, let $\mathcal{M}_S$ be a structure such that $\mathcal{M}_S \models S$. As in [11, Exercise 2.5.20(b)], we claim that the ultraproduct $\prod_{S \in I} \mathcal{M}_S / \mathcal{U}$ satisfies T . Fix a sentence $\varphi \in T$. Put $A := \{S \in I : \mathcal{M}_S \models \varphi\}$. For any $S \in S_\varphi$, we have $S \models \varphi$. This means that any model of S satisfies φ , in particular $\mathcal{M}_S \models \varphi$. Hence $S \in A$. By the definition of the filter $\mathcal{F}$, we see that $A \in \mathcal{F}$. Therefore, $A \in \mathcal{U}$. By Łoś's theorem 3.4, we get $\prod_{S \in I} \mathcal{M}_S / \mathcal{U} \models \varphi$. This holds for any $\varphi \in T$, hence $\prod_{S \in I} \mathcal{M}_S / \mathcal{U} \models T$. $\square$

In this proof, we showed that every filter is contained in an ultrafilter. Consider the filter $\mathcal{G} := \{S \subseteq \mathbb{N} : \text{card}(\mathbb{N} \setminus S) < \infty\}$ that we introduced in the beginning of this section. This filter must be contained in an ultrafilter $\mathcal{U}$, too. Moreover, this ultrafilter cannot be a principal ultrafilter. Suppose on the contrary that it is the principal ultrafilter $\mathcal{F}_a = \{S \subseteq \mathbb{N} : a \in S\}$. The set $\mathbb{N} \setminus \{a\}$ is in $\mathcal{U}$ but not in $\mathcal{F}_a$. We call this filter a *nonprincipal* ultrafilter.

3.4. Model Theory of Finite and Pseudo-Finite Fields

In this section we will provide some results about finite and pseudo-finite fields using methods of model theory. These structures are studied by many mathematicians including Hrushovski [13], a collection of results on these structures can be found in [14]. One result of particular interest in this section will be Lemma 3.2, which states that any formula in the enriched language of rings is equivalent to a conjunction of formulas that include exactly one existential quantifier and no universal quantifiers. We will

start by defining pseudo-algebraically closed fields and pseudo-finite fields. Then, we will construct an enriched language to turn $\forall$ symbols into $\exists$ symbols in formulas. Finally, we will use the result from [15] and generalize them to prove Lemma 3.2.

Consider the theory consisting of the sentences that are satisfied by all finite fields. This theory will be called the *theory of finite fields*, denoted T_f [14, p.96]. The pseudo-finite fields are defined as follows:

Definition 3.26. Infinite models of the theory of finite fields are called *pseudo-finite fields*.

We will give the axioms for pseudo-finite fields after we give the necessary definitions.

Definition 3.27. [7, p.192] A field F is called a *pseudo algebraically closed* (abbreviated as *PAC*) if each nonempty absolutely irreducible F -algebraic set has an F -point.

If F is an algebraically closed field, then every nonempty absolutely irreducible F -algebraic set has an F -point.

Recall the following definition from the field theory:

Definition 3.28. A field F is called *perfect* if F is of characteristic 0 or F is of characteristic p and $F^p = F$.

We now state an equivalent definition of pseudo-finite fields given by Ax:

Theorem 3.6. [16] *A pseudo-finite field is perfect, PAC, and has a unique field extension of degree n for each $n > 1$.*

An example of pseudo-finite fields is a nonprincipal ultraproduct of finite fields. The first fact we are going to use to prove the previous statement is that being pseudo-finite can be defined using $\mathcal{L}_{\text{ring}}$ -sentences [7, Proposition 20.4.4]. Being perfect can be

defined by the disjunction of the sentences $p \cdot 1 = 0 \wedge \forall X \exists Y X^p = Y$ where p runs over all prime numbers. Having a unique extension for each degree can also be expressed by an $\mathcal{L}_{\text{ring}}$ -sentence. First, note that for a polynomial f of the form $a_n X^n + \dots + a_1 X + a_0$, being reducible can be expressed in the following way:

$$\bigvee_i \exists b_0, \dots, b_i, c_0, \dots, c_{n-i} ((a_0 = b_0 c_0) \wedge (a_1 = a_0 b_1 + a_1 b_0) \wedge \dots \wedge (a_n = b_i c_{n-i}))$$

Therefore, there is a formula $\text{Irr}(X_0, \dots, X_n)$ that defines the coefficients of irreducible polynomials over a field. If a field has a unique extension of each degree n , then any irreducible polynomial of degree n has a root in this unique extension.

The $\mathcal{L}_{\text{ring}}$ -sentences for being PAC needs more work. These sentences use the fact that absolutely irreducible F -algebraic sets correspond to the prime ideals of $\tilde{F}$. Each element g of degree less than or equal to e in an ideal generated by the polynomials $f_1, \dots, f_r$ of degrees less than or equal to e can be defined by an $\mathcal{L}_{\text{ring}}$ -formula $g = h_1 f_1 + \dots + h_r f_r$ due to the bounds on the degrees of h_i 's given by Seidenberg [17], Schmidt and van den Dries [18]. Being a prime ideal can then be defined, we just form the sentence $f \cdot g \in I \rightarrow f \in I \vee g \in I$. Lastly, we state that given a number e , if the ideal generated by any set of polynomials $f_1, \dots, f_r$ of degrees less than or equal to e is a prime ideal, then there is an $\bar{X}$ such that $f_1(\bar{X}) = 0 \wedge \dots \wedge f_r(\bar{X}) = 0$. The conjunction of these sentences over the natural numbers e is the desired sentence, the sentence that defines PAC fields.

All finite fields are perfect and have a unique extension of each degree. However, they are not PAC. Still, each sentence defining PAC fields is true for all finite fields of cardinality greater than some number determined by the sentence. This follows from Lang-Weil theorem:

Theorem 3.7 (Lang-Weil Theorem). *[2, Theorem 1] Let e, n and r be natural numbers. Then there exists a $C > 0$ that depends on e, n and r such that for any finite field k of cardinality q and for any polynomials $f_1, \dots, f_r \in k[X_1, \dots, X_n]$ of degree less than*

or equal to e , if the algebraic set $V := V(f_1, \dots, f_r)$ is absolutely irreducible and if $\dim(V) = d$, then

$$|\text{card}(V(k)) - q^d| \leq Cq^{(d-1/2)}$$

Given any set of polynomials $f_1, \dots, f_r \in k[X_1, \dots, X_n]$ of degree less than or equal to e , if the ideal generated by these polynomials is prime in $\tilde{k}$, then $V = V(f_1, \dots, f_r)$ is absolutely irreducible over k . The set of $\bar{x}$'s in k^n with $f_1(\bar{x}) = \dots = f_r(\bar{x}) = 0$ is exactly the set $V(k)$. By Lang-Weil theorem,

$$-Cq^{(d-1/2)} + q^d \leq \text{card}(V(k)) \leq Cq^{(d-1/2)} + q^d$$

If the number $-Cq^{(d-1/2)} + q^d$ is greater than zero, then we can say that $V(k)$ is nonempty. This holds when $\sqrt{q} > C$. Hence, any single sentence of being PAC determined by e holds for sufficiently large finite fields.

We can finally construct pseudo-finite fields. Let $\mathcal{U}$ be a nonprincipal ultrafilter over the set I of nonnegative prime powers. Consider the ultraproduct $k := \prod_{q \in I} \mathbb{F}_q / \mathcal{U}$. We have just seen that being perfect, PAC, and having a unique extension of each degree can be stated using $\mathcal{L}_{\text{ring}}$ -sentences. These sentences hold for all but finitely many of the $\mathbb{F}_q$, hence the set of q 's where F_q satisfies one of the sentences is in the filter $\mathcal{G}$, hence in the ultrafilter $\mathcal{U}$. By Łoś's theorem, the ultraproduct k is also a perfect PAC field having only one extension of each degree, *i.e.* a pseudo-finite field. Note that a pseudo finite field can be of any positive characteristic p as well as of characteristic 0.

3.5. Enriched Fields

In this section we will show that any formula in the language of rings is equivalent to a much simpler formula, namely a conjunction of the formulas of the form $\exists T (g(X, T) = 0)$ modulo the theory of pseudo finite fields. Since pseudofinite fields are infinite models of finite fields, this result will still hold modulo large enough finite fields. That is; in a very large finite field, every formula is a conjunction of the formulas of the form $\exists T (g(X, T) = 0)$.

Let F be a perfect field that has a unique extension of each degree. The obvious way to state that a polynomial $g(X) \in F[X]$ does not have a zero in F is to use the statement $F \models \forall X \neg(g(X) = 0)$. Our purpose in the following paragraphs is to convert the universal formula $\forall X \neg(g(X) = 0)$ to an existential formula, so that by using the following lemma of van den Dries, it will be equivalent to a conjunction of formulas of the form $\exists T (g(X, T) = 0)$.

Lemma 3.6. *[15] Let $\phi(X)$ be an existential formula in the language of rings. Then, $\phi(X)$ is equivalent, uniformly for all PAC-fields, to a conjunction of the formulas $\exists T (g(X, T) = 0)$, with $g(X, T) \in \mathbb{Z}[X, T]$, T a single variable.*

Converting $\forall X \neg(g(X) = 0)$ to an existential formula will be done as follows; instead of saying that the roots of the polynomial $g(X)$ are not in the field F , we will say that the roots of $g(X)$ are in $F_n \setminus F$, where F_n is a field extension containing the splitting field of $g(X)$. To express this we should be able to talk about the elements of F_n . This will be possible by adding certain constant symbols to the language $\mathcal{L}$ of rings, to get the enriched language $\mathcal{L}(\bar{c})$.

Here is how the new constant symbols are added: remember that F is a perfect field having a unique extension of degree n for each $n \in \mathbb{N}$. Denote this unique extension of F by F_n . Since F is perfect $F_n = F(\alpha)$ for some element $\alpha \in F_n$. The set $\{1, \alpha, \alpha^2, \dots, \alpha^{n-1}\}$ forms a basis for F_n over F . Let $f(X) = T^n + c_{n1}T^{n-1} + \dots + c_{nn}$ be the minimal polynomial of α over F , the constants c_{nj} lets us define F_n as a vector

space over F .

Let $g(T) = a_d T^d + a_{d-1} T^{d-1} + \dots + a_0$ be a polynomial in $F[T]$. If this polynomial does not have a root in F , then it must have a root in some extension of F . The polynomial g is of degree d , so each root of g must lie in the extension of degree $d!$. Let $n = d!$ and F_n be the unique extension of F of degree n .

Suppose that $y_1, \dots, y_d$ are the roots of g . These y_i are in F_n . Since $\{1, \alpha, \dots, \alpha^{n-1}\}$ is a basis for F_n over F , for each y_i there exists $z_{i0}, z_{i1}, \dots, z_{i,n-1} \in F$ such that $y_i = z_{i0} + \alpha z_{i1} + \dots + \alpha^{n-1} z_{i,n-1}$. If g does not have any root in F , then some z_{ij} must be nonzero for $j \neq 0$. Now we can write the formula that states that g has no zero in F .

Let $y_i \notin F$ be the abbreviation for the formula $\exists z_{i1}, \dots, z_{i,n-1} (\bigwedge_{1 \leq j \leq n-1} z_{ij} \neq 0)$. Let $\psi_d(X_0, \dots, X_d, Y_1, \dots, Y_n, Z_{10}, Z_{11}, \dots, Z_{d,n-1})$ be the conjunction of formulas that states that the polynomials $X_d T^d + X_{d-1} T^{d-1} + \dots + X_0$ and $X_d \prod (T - y_i)$ are equal, y_i is the abbreviation for the term $Z_{i0} + \alpha Z_{i1} + \dots + \alpha^{n-1} Z_{i,n-1}$, α_n is a variable symbol that satisfies $\alpha^n + Y_1 \alpha^{n-1} + \dots + Y_n = 0$.

Then the statement that g does not have a zero in F is given by

$$F \models \exists \bar{z} \left(\bigwedge_{1 \leq i \leq d} y_i \notin F \right) \wedge \psi_d(\bar{a}, \bar{c}_n, \bar{z})$$

where $\bar{a} = (a_0, a_1, \dots, a_d)$ is the tuple formed by the coefficients of g , $\bar{c}_n = (c_{n1}, \dots, c_{nn})$ is the tuple formed by the coefficients of f , $\bar{z} = (z_{ij})_{1 \leq i \leq d, 0 \leq j \leq n-1}$. This formula is an existential formula as desired.

The above arguments together with Lemma 3.6 allow us to prove that every formula in the enriched language $\mathcal{L}(\bar{c})$ of rings is equivalent over a perfect PAC-field having a unique extension of each degree, in other words a pseudo finite field, to a

conjunction of formulas of the form $\exists T g(\bar{c}, \bar{X}, T) = 0$ where $g(\bar{C}, \bar{X}, T) \in \mathbb{Z}[\bar{C}, \bar{X}, T]$.

Proposition 3.2. [3, Proposition 2.7] Let $\mathcal{L}(\bar{c})$ be the enriched language of rings, $\varphi(\bar{X})$ be an $\mathcal{L}(\bar{c})$ -formula. Then $\varphi(\bar{X})$ is equivalent to a formula of the form

$$\bigwedge (\exists T (g(\bar{c}, \bar{X}, T) = 0))$$

in all enriched pseudo-finite fields and sufficiently large finite fields, where $\bar{C} = (C_{ni})_{1 \leq i \leq n}$ for some n and $g(\bar{C}, \bar{X}, T) \in \mathbb{Z}[\bar{C}, \bar{X}, T]$.

Proof. We start by investigating different kinds of formulas. The result for existential formulas can be deduced from Lemma 3.6. To apply it on the enriched language $\mathcal{L}(\bar{c})$, note that every constant symbol c_{ni} can be replaced by a variable symbol C_{ni} and $\varphi(\bar{X})$ becomes a formula $\varphi(\bar{C}, \bar{X})$ in the language $\mathcal{L}$. Now let $\varphi(\bar{X})$ be a quantifier-free formula. Then, it is a conjunction of polynomial equalities and inequalities over integers augmented with constant symbols c_{ni} . If there are no inequalities, *i.e.* if $\varphi(\bar{X})$ is equivalent to $\bigwedge_i f_i(\bar{c}, \bar{X}) = 0$, then we get the result automatically by putting $g_i(\bar{c}, \bar{X}, T) := f_i(\bar{c}, \bar{X})$. If there are some polynomial inequalities $f(\bar{X}) \neq 0$, then replace these with the formulas $\exists Y (f(\bar{X})Y - 1 = 0)$. This makes $\varphi(\bar{X})$ an existential formula and we can apply Lemma 3.6.

Finally, we prove the result for formulas that are not existential and not quantifier-free. We will proceed by induction on the number of quantifiers. Let $\varphi(\bar{X})$ be a formula of the form $Q_1 Y_1 \dots Q_n Y_n \psi(\bar{X}, \bar{Y})$, where $Q_i \in \{\forall, \exists\}$ and $\psi(\bar{X}, \bar{Y})$ is a quantifier-free formula. Let j be the greatest positive integer $\leq n$ such that Q_j is the symbol $\forall$ and Q_i is the symbol $\exists$ for $i > j$. That is, the formula $\varphi(\bar{X})$ is the formula

$$Q_1 Y_1 \dots Q_{j-1} Y_{j-1} \forall Y_j \exists Y_{j+1} \dots \exists Y_n \psi(\bar{X}, \bar{Y})$$

Consider the formula $\forall Y_j \exists Y_{j+1} \dots \exists Y_n \psi(\bar{X}, \bar{Y})$. Using the meaning of $\forall$, we see that

this formula is the formula $\neg(\exists Y_j \neg(\exists Y_{j+1} \dots \exists Y_n \psi(\bar{X}, \bar{Y})))$. Moreover, by Lemma 3.6, this is equivalent to $\neg(\exists Y_j \neg(\bigwedge \exists T g(\bar{c}, \bar{X}, Y_1, \dots, Y_j, T) = 0))$.

Previous work showed that a formula of the form $\forall T g(T) \neq 0$ is equivalent to a formula of the form $\exists \bar{Z} (\bigwedge (\bigvee Z_{ij} \neq 0)) \wedge \psi_d(\bar{a}, \bar{c}_n, \bar{Z})$. This formula is an existential formula, so we get an equivalent formula of the form $\bigwedge \exists T g(\bar{c}, \bar{X}, T)$. We now go back to the discussion in the previous paragraph. The formula

$$\neg(\exists Y_j \neg(\bigwedge \exists T g(\bar{c}, \bar{X}, Y_1, \dots, Y_j, T) = 0))$$

is equivalent to the formula

$$\neg(\exists Y_j (\bigvee \neg(\exists T g(\bar{c}, \bar{X}, Y_1, \dots, Y_j, T) = 0)))$$

We have already showed that the formula $\neg \exists T g(T) = 0$ is equivalent to an existential formula. We need to generalize this result to our polynomial $g(\bar{c}, \bar{X}, Y_1, \dots, Y_j, T)$. We can see this polynomial as a polynomial on the single variable T by rewriting it as $\sum h(\bar{c}, \bar{X}, Y_1, \dots, Y_j) T^\nu$, where $h_\nu(\bar{C}, \bar{X}, Y_1, \dots, Y_j) \in \mathbb{Z}[\bar{C}, \bar{X}, Y_1, \dots, Y_j]$. For certain elements $\bar{c}, (y_1, \dots, y_j)$ and $\bar{x}$, $h_\nu(\bar{c}, \bar{x}, y_1, \dots, y_j)$ might be equal to zero for some ν . This may change the degree of g but there are finite number of possibilities for the degree. Therefore, the formula $\neg(\exists T g(\bar{c}, \bar{X}, Y_1, \dots, Y_j, T) = 0)$ is equivalent to a disjunction

$$\bigvee_d \exists \bar{Z} (\bigwedge (\bigvee Z_{ij} \neq 0)) \wedge \psi_d(\bar{X}, Y_1, \dots, Y_j, \bar{c}_n, \bar{Z})$$

where d runs over all possible degrees of g . This is equivalent to an existential formula, too. Therefore $\varphi(\bar{X})$ becomes

$$Q_1 Y_1 \dots Q_{j-1} Y_{j-1}, \exists T_1, \dots, \exists T_l \theta(\bar{X}, Y_1, \dots, Y_{j-1}, T_1, \dots, T_l)$$

where θ is a quantifier-free formula. Inductively, we can show that the whole formula φ is equivalent to a formula in the desired form.

The statement of the proposition also holds for sufficiently large finite fields. Suppose on the contrary that there are infinitely many finite fields in which the statement does not hold. Then the statement does not hold for the ultraproduct of these fields, which is a pseudo finite field. Hence, there can only be finitely many finite fields that do not satisfy the statement of the proposition. $\square$

We need the following lemma before proving the next proposition.

Lemma 3.7. *Let $A_{1,i}, A_{2,i}$ be formulas for $i = 1, \dots, k$. Then,*

$$\bigwedge_{i=1}^k (A_{1,i} \vee A_{2,i}) = \bigvee_{S \subseteq K} \left(\left(\bigwedge_{j \in S} A_{1,j} \right) \wedge \left(\bigwedge_{j \notin S} A_{2,j} \right) \right)$$

where $K = \{1, \dots, k\}$.

Proof. By induction on k .

For $k = 2$, we have

$$\begin{aligned} (A_{1,1} \vee A_{2,1}) \wedge (A_{1,2} \vee A_{2,2}) &= ((A_{1,1} \vee A_{2,1}) \wedge A_{1,2}) \vee ((A_{1,1} \vee A_{2,1}) \wedge A_{2,2}) \\ &= (A_{1,1} \wedge A_{1,2}) \vee (A_{2,1} \wedge A_{1,2}) \vee (A_{1,1} \vee A_{2,2}) \wedge (A_{2,1} \vee A_{2,2}) \end{aligned}$$

Suppose that the claim holds for k . Let $K' = \{1, \dots, k, k+1\}$. Then for $k+1$, we have

$$\begin{aligned}
\bigwedge_{i=1}^{k+1} (A_{1,i} \vee A_{2,i}) &= \left(\bigwedge_{i=1}^k (A_{1,i} \vee A_{2,i}) \right) \wedge (A_{1,k+1} \vee A_{2,k+1}) \\
&= \left(\bigvee_{S \subseteq K} \left(\left(\bigwedge_{j \in S} A_{1,j} \right) \wedge \left(\bigwedge_{j \notin S} A_{2,j} \right) \right) \right) \wedge (A_{1,k+1} \vee A_{2,k+1}) \\
&= \left(\bigvee_{S \subseteq K} \left(\left(\bigwedge_{j \in S} A_{1,j} \right) \wedge \left(\bigwedge_{j \notin S} A_{2,j} \right) \wedge A_{1,k+1} \right) \right) \\
&\quad \vee \left(\bigvee_{S \subseteq K} \left(\left(\bigwedge_{j \in S} A_{1,j} \right) \wedge \left(\bigwedge_{j \notin S} A_{2,j} \right) \wedge A_{2,k+1} \right) \right) \\
&= \left(\bigvee_{T \subseteq K', k+1 \in T} \left(\left(\bigwedge_{j \in T} A_{1,j} \right) \wedge \left(\bigwedge_{j \notin T} A_{2,j} \right) \right) \right) \\
&\quad \vee \left(\bigvee_{T \subseteq K', k+1 \notin T} \left(\left(\bigwedge_{j \in T} A_{1,j} \right) \wedge \left(\bigwedge_{j \notin T} A_{2,j} \right) \right) \right) \\
&= \bigvee_{T \subseteq K'} \left(\left(\bigwedge_{j \in T} A_{1,j} \right) \wedge \left(\bigwedge_{j \notin T} A_{2,j} \right) \right) \quad \square
\end{aligned}$$

The following proposition will be used in an application of the main theorem of this thesis.

Proposition 3.3. *[3, Lemma 2.9] For any field F , let $\mathcal{L}(F)$ be the language of rings together with constant symbols for each element in F .*

For any pseudo-finite field F and for any $S \subseteq F^m$ defined by an $\mathcal{L}(F)$ formula there are natural numbers e and n such that S is a finite union of sets $\pi(V(F))$. Here, $V \subseteq \tilde{F}^{m+n}$ is an F -algebraic set, $V(F)$ is the set of F -points of V , $\pi: \tilde{F}^{m+n} \rightarrow \tilde{F}^m$ is the projection $\pi: (x_1, \dots, x_m, y_1, \dots, y_n) \mapsto (x_1, \dots, x_m)$, and the set $V_{\bar{x}} := \{y \in \tilde{F}^n : (\bar{x}, y) \in V\}$ has at most e elements for any $\bar{x} \in \tilde{F}^m$.

Proof. The language $\mathcal{L}(F)$ has the same properties with the enriched language $\mathcal{L}(\bar{c})$, the proof of Lemma 3.2 holds over $\mathcal{L}(F)$. Therefore, by Lemma 3.2 the formula defining S

is a conjunction of the formulas of the form $\exists T f(\bar{X}, T) = 0$, where $f(\bar{X}, T) \in F[\bar{X}, T]$ and $\bar{X} = (X_1, \dots, X_m)$.

In order to bound the number of elements of the sets $V_{\bar{x}}$, the cases in which $f(\bar{x}, T) = 0$ for all T must be handled. Consider $f(\bar{X}, T)$ as a polynomial in T with coefficients $f_i(\bar{X})$. Now, the formula $\exists T f(\bar{X}, T) = 0$ is equivalent to

$$\bigwedge_i f_i(\bar{x}) = 0 \vee \exists T \exists U (f(\bar{X}, T) = 0 \wedge \prod_i (f_i(\bar{X}) \cdot U - 1) = 0)$$

In Lemma 3.7, if we choose $A_{1,i}$ to be the formula $\bigwedge_i f_i(\bar{x}) = 0$ and $A_{2,i}$ to be the formula $\exists T \exists U (f(\bar{X}, T) = 0 \wedge \prod_i (f_i(\bar{X}) \cdot U - 1) = 0)$, we see that their conjunction is equivalent to a disjunction of the formulas

$$\left(\bigwedge_{j=1}^l g_j(\bar{X}) = 0 \right) \wedge \exists \bar{Y} \left(\bigwedge_{k=1}^n h_k(\bar{X}, Y_k) \right)$$

where $g_j(\bar{X}) \in F[\bar{X}]$ and $h_k(\bar{X}, Y_k) \in F[\bar{X}, Y_k]$ for $1 \leq k \leq n$. Moreover, the set $\{\bar{y} \in \tilde{F}^n : h_k(\bar{x}, \bar{y}) = 0 \text{ for all } k\}$ has at most e elements for any $\bar{x} \in F^m$. Put V to be the F -algebraic set $V(g_1, \dots, g_l, h_1, \dots, h_n)$. By the previous remark, we see that $\text{card}(V_x) \leq e$. Furthermore, $\pi(V) = S$. $\square$

4. THE DECOMPOSITION-INTERSECTION PROCEDURE

We want to find estimates on the number of elements of definable sets over finite fields. We have already proved in Proposition 3.2, any definable set over a sufficiently large finite field F is the set of F -points of some projection of an F -algebraic set. Let V be an F -algebraic set. The absolutely irreducible components of V are not necessarily F -algebraic. Still, in this chapter we will investigate a method that uniformly generates an F -algebraic set V^* such that $V^*(F) = V(F)$ and absolutely irreducible components are F -algebraic sets. This method is called the decomposition-intersection procedure. The results from Chapter 2 will be used frequently.

4.1. Constructing the F -Absolute Kernel

Let F be a field and $\tilde{F}$ be its algebraic closure. Suppose that $f_1, \dots, f_r \in F[\bar{X}]$ where $\bar{X} = (X_1, \dots, X_n)$. Let V be the F -algebraic set $V(f_1, \dots, f_r)$.

By Theorem 2.2, V can be decomposed uniquely into its absolutely irreducible components. Let $V_1, \dots, V_k$ be the absolutely irreducible components of V . Then $V = V_1 \cup \dots \cup V_k$. Some of these V_i 's may not be F -algebraic. However, there is a finite extension F' such that all V_i are F' -algebraic.

Let $G(F) = \text{Aut}(\tilde{F}|F)$. If $\sigma \in G$, then for any $i = 1, \dots, k$ there is a $j = 1, \dots, k$ such that $\sigma(V_i) = V_j$. To see this, note that $\sigma(V_i)$ is a K -algebraic set for some finite degree extension K of F , and V_j can be written as $\bigcup_i \sigma(V_i) \cap V_j$. Since V_j is absolutely irreducible, one of these $\sigma(V_i) \cap V_j$ should be equal to V_j and the rest should be empty.

Define the sets $W_i := \bigcap_{\sigma \in G} \sigma(V_i)$ for $i = 1, \dots, k$. With this definition, we get $V_i \cap F^n = W_i \cap F^n$. To see this, let $x \in V_i \cap F^n$ be arbitrary, then $\sigma(x) = x$ for any $\sigma \in G$. Therefore, $x \in \bigcap_{\sigma \in G} \sigma(V_i) = W_i$. Also, $W_i \subseteq V_i$, so $W_i \cap F^n \subseteq V_i \cap F^n$.

Let $\sigma \in G$ be arbitrary. $\sigma(W_i) = \sigma(\bigcap_{\rho \in G} \rho(V_i)) = \bigcap_{\rho \in G} \sigma\rho(V_i) = W_i$, therefore W_i is F -algebraic.

Define $V' := W_1 \cup \dots \cup W_k$. V' is a F -algebraic as it is a union of F -algebraic sets. V' is a subset of V , since $W_i \subseteq V_i$ for each i . Moreover,

$$\begin{aligned} V \cap F^n &= (V_1 \cup \dots \cup V_k) \cap F^n = (V_1 \cap F^n) \cup \dots \cup (V_k \cap F^n) \\ &= (W_1 \cap F^n) \cup \dots \cup (W_k \cap F^n) \\ &= (W_1 \cup \dots \cup W_k) \cap F^n \\ &= V' \cap F^n \end{aligned}$$

that is, V and V' have the same F -points.

V' is an F -algebraic set, so we can construct the F -algebraic set V'' in the same way. These sets form a chain $V \supseteq V' \supseteq V'' \supseteq \dots$. By Corollary 2.1, this chain terminates at some $V^{(r)}$. This set will be called the *F -absolute kernel of V* and denoted V^* .

Proposition 4.1. [3, p.111] *The absolutely irreducible components of V^* are F -algebraic sets.*

Proof. Let $V_1, \dots, V_k$ be the absolutely irreducible components of V^* , let $W_i = \bigcap_{\sigma \in G} \sigma(V_i)$. Since $(V^*)' = V^*$, we have $W_1 \cup \dots \cup W_k = V_1 \cup \dots \cup V_k$. We also know that $W_i \subseteq V_i$ for any $i = 1, \dots, k$. Now we will prove that $W_i = V_i$. Suppose on the contrary that

$V_i \setminus W_i \neq \emptyset$. We have

$$\begin{aligned}
 V_i &= V_i \cap (W_1 \cup \dots \cup W_i \dots \cup W_k) \\
 &= (V_i \cap W_1) \cup \dots \cup (V_i \cap W_i) \dots \cup (V_i \cap W_k) \\
 &= (V_i \cap W_i) \cup \left(\bigcup_{j \neq i} V_i \cap W_j \right) \\
 &= W_i \cup \left(\bigcup_{j \neq i} V_i \cap W_j \right)
 \end{aligned}$$

Each $V_i \cap W_j$ is an algebraic set, and $V_i \setminus W_i \neq \emptyset$, therefore $\bigcup_{j \neq i} V_i \cap W_j$ is a nonempty algebraic set. This contradicts V_i being an absolutely irreducible algebraic set because it is a union of two algebraic sets. Hence, $W_i = V_i$ for all i . Since W_i 's are F -algebraic, V_i 's are F -algebraic, too. $\square$

Proposition 4.2. [3, p.111] *Every F -algebraic absolutely irreducible subset of V is contained in V^* .*

Proof. Let V_1 be an F -algebraic absolutely irreducible component of V . This means that $V_1 = V(\{f_1(Y), \dots, f_k(Y) : f_i(Y) \in F[Y]\})$. If $\alpha \in V_1$, then for any $i \in \{1, \dots, k\}$ and for any $\sigma \in G$, $f_i(\sigma(\alpha)) = \sigma(f_i(\alpha)) = 0$. Therefore, $\cap_{\sigma} \sigma(V_1) = V_1$, which in turn implies $V_1 = W_1$. So, V_1 is an absolutely irreducible subset of V' . Hence, V_1 is an absolutely irreducible subset of V^* . $\square$

We have stated earlier that there is a finite degree extension F' of F such that all the absolutely irreducible components V_i of V are algebraic in F' . Since $[F' : F]$ is finite, and F' is a vector space over F , we can find a basis $b_1, \dots, b_N$ that generates F' over F .

Proposition 4.3. [3, p.112] *Let F' be an extension of F generated by the basis $\{b_1, \dots, b_N\}$. Let $V_i := V(g_{i1}, \dots, g_{ir(i)})$, where $g_{ij} \in F'[\bar{X}]$. Then there exists $g_{ijv} \in F[\bar{X}]$ such that $g_{ij} = \sum_v b_v g_{ijv}$.*

Proof. Let $g_{ij}(X_1, \dots, X_n) = \sum_l a_l X_1^{\alpha_{l1}} \dots X_n^{\alpha_{ln}}$, where $a_l \in F'$ and $\alpha_{li} \in \{0, 1, 2, \dots\}$. Suppose that $a_l = \sum_v b_v a_{lv}$, where $a_{lv} \in F$.

$$\begin{aligned}
 g_{ij}(X_1, \dots, X_n) &= \sum_l \left(\sum_v b_v a_{lv} \right) X_1^{\alpha_{l1}} \dots X_n^{\alpha_{ln}} \\
 &= \sum_l \left(\sum_v b_v a_{lv} X_1^{\alpha_{l1}} \dots X_n^{\alpha_{ln}} \right) \\
 &= \sum_v \left(\sum_l b_v a_{lv} X_1^{\alpha_{l1}} \dots X_n^{\alpha_{ln}} \right) \\
 &= \sum_v b_v \left(\sum_l a_{lv} X_1^{\alpha_{l1}} \dots X_n^{\alpha_{ln}} \right)
 \end{aligned}$$

Put $g_{ijv}(X_1, \dots, X_n) = \sum_l a_{lv} X_1^{\alpha_{l1}} \dots X_n^{\alpha_{ln}}$. □

Recall that the set W_i is defined as $W_i = \cap_{\sigma} \sigma(V_i)$ and the set V' is defined as $V' = W_1 \cup \dots \cup W_k$.

Proposition 4.4. [3, p.113] Suppose that the polynomials $g_{ijv} \in F[\bar{X}]$ are as in Proposition 4.3. Then $W_i = V(\{g_{ijv} : 1 \leq j \leq r(i), 1 \leq v \leq n\})$.

Proof. Fix $i \in \{1, \dots, k\}$. Let $(x_1, \dots, x_n) \in V(\{g_{ijv} : 1 \leq j \leq r(i), 1 \leq v \leq n\})$, where g_{ijv} is defined as in Proposition 4.3. Let $\sigma \in G$ be arbitrary.

$$\begin{aligned}
 g_{ij}(\sigma^{-1}(x_1), \dots, \sigma^{-1}(x_n)) &= \sum_v b_v g_{ijv}(\sigma^{-1}(x_1), \dots, \sigma^{-1}(x_n)) \\
 &= \sum_v b_v \sigma^{-1}(g_{ijv}(x_1, \dots, x_n)) \\
 &= \sigma^{-1} \left(\sum_v \sigma(b_v) g_{ijv}(x_1, \dots, x_n) \right) \\
 &= 0
 \end{aligned}$$

So, $\sigma^{-1}((x_1, \dots, x_n)) \in V_i$, in other words, $(x_1, \dots, x_n) \in \sigma(V_i)$. Since σ was arbitrary, $(x_1, \dots, x_n) \in W_i$.

Now, let $(x_1, \dots, x_n) \in W_i$. Since $W_i \subseteq V_i$, we have $(x_1, \dots, x_n) \in V_i$, and therefore $g_{ij}(x_1, \dots, x_n) = 0$ for any $j \in \{1, \dots, r(i)\}$. By Proposition 4.3, we get

$$\begin{aligned} g_{ij}(x_1, \dots, x_n) &= \sum_v b_v g_{ijv}(x_1, \dots, x_n) \\ &= 0 \end{aligned}$$

Since $\{b_1, \dots, b_N\}$ is a basis, $g_{ijv} = 0$ for each j and v . □

Proposition 4.5. [3, p.113] Suppose that $g_{i,j,v}$ are defined as in Proposition 4.4. Then

$$V' = V \left(\left\{ \prod_{1 \leq i \leq k} g_{ij(i)v(i)} : 1 \leq j(i) \leq r(i), 1 \leq v(i) \leq N \text{ for each } i \in \{1, \dots, k\} \right\} \right)$$

Proof. Let $(x_1, \dots, x_n)$ be an element of the set in the right hand side of the above equality. Suppose also that $(x_1, \dots, x_n) \notin W_i$ for any $i \in \{1, \dots, k\}$. So, for each i there are $j_i^0 \in \{1, \dots, r(i)\}$ and $v_i^0 \in \{1, \dots, N\}$ such that $g_{ij_i^0 v_i^0}(x_1, \dots, x_n) \neq 0$. Define the functions j^0 and v^0 on $\{1, \dots, k\}$ such that $j^0(i) = j_i^0$ and $v^0(i) = v_i^0$ for all i . Now, we have $\prod_{1 \leq i \leq k} g_{ij^0(i)v^0(i)}(x_1, \dots, x_n) \neq 0$, contradicting our assumption.

Let $(x_1, \dots, x_n) \in V'$. There exists $i^0 \in \{1, \dots, k\}$ such that $g_{i^0 j v}(x_1, \dots, x_n) = 0$ for each $j \in \{1, \dots, r(i^0)\}$ and for each $v \in \{1, \dots, N\}$. Therefore,

$$\begin{aligned} \prod_{1 \leq i \leq k} g_{ij(i)v(i)}(x_1, \dots, x_k) &= g_{i^0 j(i^0)v(i^0)}(x_1, \dots, x_n) \prod_{i \neq i^0} g_{ij(i)v(i)}(x_1, \dots, x_k) \\ &= 0 \end{aligned}$$

which means that $(x_1, \dots, x_n) \in V \left(\left\{ \prod_{1 \leq i \leq k} g_{ij(i)v(i)} \right\} \right)$. □

4.2. Decomposing the F -Absolute Kernel

In the previous section, we defined, for a given variety V the F -absolute kernel V^* of V so that V^* has the same F -points as V and absolutely irreducible components of V^* are F -algebraic sets.

We need to find a bound for the number of absolutely irreducible components of V^* in order to utilize the methods used in the main theorem. In this section, we include root functions to the language of rings. These functions take the coefficients of a polynomial as an input and give a root of the polynomial as an output. Then we will show that using this language we get quantifier elimination. It becomes possible to prove that there is an upper bound on the number of absolutely irreducible components of V^* independent of the field V is defined and the parameters we picked to define V .

We begin by defining root functions.

Definition 4.1. [3, p.113] Let $d \in \{2, 3, 4, \dots\}$, $\varrho_d : F^d \rightarrow F$ be a function satisfying

$$X^d + a_1 X^{d-1} + \dots + a_{d-1} X + a_d \rightarrow \varrho_d(\bar{a})^d + a_1 \varrho_d(\bar{a})^{d-1} + \dots + a_{d-1} \varrho_d(\bar{a}) + a_d = 0$$

where $\bar{a} = (a_1, \dots, a_d)$. We call ϱ_d a *root function*.

Let $\bar{X} = (X_1, \dots, X_m)$, $\bar{Y} = (Y_1, \dots, Y_n)$, and $f_1(\bar{X}, \bar{Y}), \dots, f_r(\bar{X}, \bar{Y}) \in \mathbb{Z}[\bar{X}, \bar{Y}]$. For a field F and a point $\bar{x} = (x_1, \dots, x_m) \in F^m$, define

$$V_{F, \bar{x}} := \{\bar{y} \in \tilde{F}^n : f_1(\bar{x}, \bar{y}) = \dots = f_r(\bar{x}, \bar{y}) = 0\}$$

This set is an F -algebraic set as $f_i(\bar{x}, \bar{Y})$ is a polynomial for $1 \leq i \leq r$. Now, consider the field generated by $x_1, \dots, x_m$. Let E be the relative algebraic closure in F of this field. Then $\tilde{E} \subseteq \tilde{F}$. Let V be the algebraic set $V_{E, \bar{x}}$ for the moment. The methods in

the previous section can be applied to V to get the polynomials g_{ijv} as before. We can summarize the properties of these polynomials as follows:

- (i) We have a polynomial $q(T) \in E[T]$ of degree N that is monic, irreducible and separable.
- (ii) For a root α of q , we can consider the set $\{1, \alpha, \dots, \alpha^{N-1}\}$ as a basis of $E(\alpha)$ over E . If $g_{ij} = \sum g_{ijv} \alpha^v$, then the irreducible components of $V_{E, \bar{x}}$ are the sets $V_i := V(g_{i1}, \dots, g_{ir(i)})$.

Consider the ring R generated by $x_1, \dots, x_m$. Let $\bar{R}$ be the integral closure of R in E . Then E is the field of fractions of $\bar{R}$. To see this, note that if $\varepsilon \in F$ is algebraic over the field generated by $x_1, \dots, x_m$, then there is a polynomial $p(T) \in \mathbb{Q}(x_1, \dots, x_m)[T]$ such that $p(\varepsilon) = 0$. There exists some $r \in R$ such that $rp(T) \in R[T]$. Let $rp(T) = a_n T^n + a_{n-1} T^{n-1} + \dots + a_0$. Then $a_n \varepsilon$ is a root of the monic polynomial $T^n + a_{n-1} T^{n-1} + a_n a_{n-2} T^{n-2} + \dots + a_n^{n-1} a_0$. Hence, $a_n \varepsilon$ is integral over R and hence an element of $\bar{R}$. Therefore, any $\varepsilon \in E$ is a ratio of two elements from $\bar{R}$.

Now, we add root functions to the language of rings, call this language $\mathcal{L}(\varrho)$. Elements of R are of the form $\tau(\bar{x})$, where $\tau(\bar{X})$ is an $\mathcal{L}_{\text{ring}}$ -term. Let $r \in \bar{R}$, then there is a monic polynomial p over R such that $p(r) = 0$. We can find a root function ϱ_d in $\mathcal{L}(\varrho)$ such that ϱ_d maps the tuple of the coefficients of p to r . Hence, each element of $\bar{R}$ is $\tau(\bar{x})$ where $\tau(\bar{X})$ is an $\mathcal{L}(\varrho)$ -term.

As we stated in the beginning of this chapter, we want a formula that allows us to talk about absolutely irreducible components of $V_{F, \bar{x}}^*$. The following lemma states that there are formulas such that given a field F and a point $\bar{x} \in F^m$, these formulas hold for $\bar{x}$ if and only if the set $V_{F, \bar{x}}$ is absolutely irreducible and the dimension of $V_{F, \bar{x}}$ is of dimension d . This result is related to some of the works of Kronecker.

Lemma 4.1. *[3, Lemma 1.12] There are quantifier-free formulas $\text{AI}(\bar{X})$ and $\text{Dim}_d(\bar{X})$ for each*

$d \in \{-\infty, 0, \dots, n\}$, in the language of rings such that for each field F and $\bar{x} \in F^m$ we

have:

- (i) $V_{F,\bar{x}} \subseteq \tilde{F}^n$ is absolutely irreducible $\Leftrightarrow F \models \text{AI}(\bar{x})$,
- (ii) $\dim(V_{F,\bar{x}}) = d \Leftrightarrow F \models \text{Dim}_d(\bar{x})$

We have another quantifier-free formula which is in the language $\mathcal{L}(\varrho)$ that determines whether a polynomial is irreducible or not.

Lemma 4.2. [3, Lemma 1.10] *Let $\mathcal{L}(\varrho)$ be the language of rings with root functions. For every $d \in \mathbb{N}$ there is a quantifier-free $\mathcal{L}(\varrho)$ -formula $\text{Irr}_d(\bar{x})$ such that for each field F with root functions and each $\bar{a} \in F^d$, we have*

$$F \models \text{Irr}_d(\bar{a}) \Leftrightarrow \text{the polynomial } T^d + a_1 T^{d-1} + \dots + a_d \in F[T] \text{ is irreducible.}$$

The next definitions give us a tool to partition F^m into subsets

Definition 4.2. [3, p.114] Suppose that $\mathcal{D}$ is the tuple $(q(\bar{X}, T), g_1(\bar{X}, \bar{Y}), \dots, g_k(\bar{X}, \bar{Y}))$, where

- (i) $q(\bar{X}, T)$ is a monic polynomial in T of degree N with $\mathcal{L}(\varrho)$ -term coefficients and
- (ii) $g_i(\bar{X}, \bar{Y})$ is the tuple $(g_{i,j,v}(\bar{X}, \bar{Y}))_{1 \leq j \leq r(i), 1 \leq v \leq N}$ with each $g_{i,j,v}(\bar{X}, \bar{Y})$ is a polynomial in $\bar{Y}$ with $\mathcal{L}(\varrho)$ -term coefficients.

Then $\mathcal{D}$ is called a *potential decomposition* for f over F .

Definition 4.3. [3, p.115] Let $\mathcal{D}$ be a potential decomposition for f over F . Fix $\bar{x} \in F^m$. Suppose that

- (i) $q(\bar{X}, T)$ is a separable and irreducible polynomial in $F[T]$ and for an $\alpha \in \tilde{F}$ with $q(\bar{x}, \alpha) = 0$, and
- (ii) for $g_{ij} := \sum g_{ijv} \alpha^{v-1}$, the sets defined as $V_i := V(g_{i,1}, \dots, g_{i,r(i)})$ are the absolutely irreducible components of $V_{F,\bar{x}}$.

Then $\mathcal{D}$ is called a *decomposition* for f over F at $\bar{x}$.

Let us now fix a field F with root functions. Firstly, we will show that a decomposition for f over F at $\bar{x}$ exists. We have already seen in the discussion on $V_{E,\bar{x}}$ that we can find a monic, separable and irreducible polynomial $q(T)$ of degree N , and polynomials $g_{ijv}(\bar{Y}) \in E[\bar{Y}]$ such that $V_1, \dots, V_k$ are the irreducible components of $V_{E,\bar{x}}$ where $V_i = V(g_{i1}, \dots, g_{ir(i)})$ and $g_{ij} = \sum g_{ijv} \alpha^{v-1}$. Moreover, since E is the fraction field of $\bar{R}$, we can take $q(T) \in \bar{R}[T]$ and $g_{ijv}(\bar{Y}) \in \bar{R}[\bar{Y}]$. Finally, since $\bar{R}$ is the set of $\tau(\bar{x})$'s for each $\mathcal{L}(\varrho)$ -term τ , we notice that q and g_{ijv} 's are polynomials with $\mathcal{L}(\varrho)$ -term coefficients. Last things to show here are that $q(T)$ is irreducible over F and that absolutely irreducible components of $V_{E,\bar{x}}$ and $V_{F,\bar{x}}$ are the same. Irreducibility of $q(T)$ stems from the fact that E is relatively algebraically closed in F . Suppose that $q(T) = q_1(T)q_2(T)$ for some $q_1(T), q_2(T) \in F[T]$. Then the coefficients of q_1 and q_2 must be algebraic over E . Hence they belong to E . Since q is not irreducible over E , one of q_1 or q_2 should be a unit. Lastly by Lemma 4.1 we see that being absolutely irreducible is given by a quantifier-free $\mathcal{L}_{\text{ring}}$ -formula $\text{AI}(\bar{X})$. Moreover, E is a substructure of F as an $\mathcal{L}_{\text{ring}}$ -structure. By Proposition 3.1, for $\bar{x} \in E^m$ we have $E \models \text{AI}(\bar{x}) \Leftrightarrow F \models \text{AI}(\bar{x})$.

Secondly, we will state that there is a quantifier-free $\mathcal{L}(\varrho)$ -formula that determines if a given potential decomposition for f over F is a decomposition for f over F at $\bar{x}$. Let $\mathcal{D} = (q, g_1, \dots, g_k)$ be a decomposition. Irreducibility of q is given by the quantifier-free $\mathcal{L}(\varrho)$ -formula Irr by Lemma 4.2. Absolute irreducibility of the algebraic sets $V_i := V(g_{i,1}, \dots, g_{i,r(i)})$ are given by the quantifier-free $\mathcal{L}_{\text{ring}}$ -formulas AI and Dim by Lemma 4.1. Separability of q can also be expressed in a quantifier-free way in $\mathcal{L}(\varrho)$. Let for any polynomial f and any potential decomposition $\mathcal{D}$, $\text{Dec}_{\mathcal{D},f}(\bar{X})$ be the quantifier-free formula satisfying

$$F \models \text{Dec}_{\mathcal{D},f}(\bar{x}) \Leftrightarrow \mathcal{D} \text{ is a decomposition for } f \text{ over } F \text{ at } x$$

By compactness, we see that there are only finitely many potential decompositions,

say $\mathcal{D}_1, \dots, \mathcal{D}_J$, for a fixed f such that for any field F and any point $\bar{x} \in F^m$, there is some $1 \leq i \leq J$ such that $F \models \text{Dec}_{\mathcal{D},f}(\bar{x})$

Theorem 4.1. [3, Theorem 1.15] *There are quantifier-free $\mathcal{L}(\varrho)$ -formulas $c_i(\bar{X})$ for each $i \in \{1, \dots, I\}$, finitely many polynomials $h_{i,j}(\bar{X}, \bar{Y})$ in variable $\bar{Y}$ with $\mathcal{L}(\varrho)$ -term coefficients for each $i \in \{1, \dots, I\}$, and a function $M : 1, \dots, I \rightarrow \{0, 1, \dots\}$ such that for any field F , $F^m \subseteq \bigcup c_i(F^m)$ and for any $\bar{x} \in c_i(F^m)$, the absolutely irreducible components of $V_{F,\bar{x}}^*$ are the sets $V_i := V(h_{i1}, \dots, h_{iM(i)})$.*

Theorem 4.2. [3, Proposition 1.7] *Let $\bar{Y} = (Y_1, \dots, Y_n)$. Let $V = V(f_1, \dots, f_r)$, where $f_1, \dots, f_r \in F[\bar{Y}]$ are all of degree $\leq d$. Then there are natural numbers D, M and R depending on (d, n, r) and independent of the field F and the specific polynomials f_j , such that*

- (i) V^* has at most M absolutely irreducible components,
- (ii) each absolutely irreducible component of V^* is of the form $V(h_1, \dots, h_R)$ for h 's in $F[\bar{Y}]$ of degree $\leq D$.

Theorem 4.2 is a result of Theorem 4.1. For polynomials $f_1, \dots, f_r$, the coefficients of these lie in some $c_i(F^m)$ and hence the number of irreducible components of V^* is bounded by the maximum of the set $\{M(1), \dots, M(I)\}$. Moreover, there are finitely many polynomials h_{ij} defining the absolutely irreducible components of V^* and their degrees are bounded by a number determined by the numbers D, M and R .

The following corollary of Theorem 4.1 will be used for applications of the main theorem.

Corollary 4.1. [3, p.116] *Let $f \in \mathbb{Z}[\bar{X}, \bar{Y}]^r$, $0 \leq d \leq n$, and $1 \leq \mu \leq \max\{M(i) : i \in \{1, \dots, I\}\}$, where $M(i)$'s are as in Theorem 4.1. Then there is an $\mathcal{L}_{\text{ring}}$ -formula $S_{f,d,\mu}$ such that for any field F and for any $\bar{x} \in F^m$, $F \models S_{f,d,\mu}(\bar{x}) \Leftrightarrow \dim(V_{F,\bar{x}}^*) = d$ and $V_{F,\bar{x}}^*$ has exactly μ absolutely irreducible components of dimension d*

5. MAIN THEOREM

In this chapter, we will prove that given an arbitrary $\mathcal{L}$ -formula $\varphi(\bar{X}, \bar{Y})$, there is a positive constant C and a finite set $D \subseteq \mathbb{N} \times \mathbb{Q}^+$ of pairs (d, μ) such that if k is a finite field of cardinality q and $\bar{x} \in k^m$ with $\varphi(\bar{x}, k^n) \neq \emptyset$, then

$$|\text{card}(\varphi(\bar{x}, k^n)) - \mu q^d| \leq Cq^{d-1/2}$$

for some $(d, \mu) \in D$. Here are the steps we will take for the proof of the main theorem: *Step 1:* If $\varphi(\bar{X}, \bar{Y})$ gives an absolutely irreducible algebraic set, we get the result by Lang-Weil estimates (in Theorem 5.1), where $\mu = 1$ and d is the algebraic dimension of the algebraic set. *Step 2:* If $\varphi(\bar{X}, \bar{Y})$ defines an algebraic set V , then with the help of decomposition-intersection procedure we decompose k -rational points of V as the union of k -rational points of absolutely irreducible varieties defined over k and hence we get the above estimates where $\mu \in \mathbb{N}$ and d the algebraic dimension of V (in Proposition 5.1). *Step 3:* Next, we express quantifier-free definable sets as zeros of polynomials over k so that we get the estimates for quantifier-free definable sets (in Lemma 5.1). *Step 4:* Finally, by quantifier elimination of Ax's result we write arbitrary definable sets as projections of absolutely irreducible sets, where each fiber is finite and bounded. Then we use intricate counting arguments to count the points of the projection of an algebraic set hence get estimates for the number of rational points over k (in Theorem 5.2).

In this chapter, unless otherwise stated, k will denote a finite field of cardinality q . We now state the Lang-Weil theorem:

5.1. Lang-Weil Theorem

Lang-Weil theorem was used before in Section 3.4. This time we are using the theorem for finding an estimate for the number of k -points of an absolutely irreducible

k -algebraic set.

Theorem 5.1 (Lang-Weil Theorem). *[2, Theorem 1] Let e, n and r be natural numbers. Then there exists a $C > 0$ that depends on e, n and r such that for any finite field k of cardinality q and for any polynomials $f_1, \dots, f_r \in k[X_1, \dots, X_n]$ of degree less than or equal to e , if the algebraic set $V := V(f_1, \dots, f_r)$ is absolutely irreducible and if $\dim(V) = d$, then*

$$|\text{card}(V(k)) - q^d| \leq Cq^{(d-1/2)}$$

5.2. Estimates for Arbitrary Varieties

The previous result will be generalized now. We will now show that an estimate can be found for the number of k -points of a k -algebraic set V even though V is not absolutely irreducible.

Proposition 5.1. *[3, Proposition 3.3] Let e, n and r be natural numbers. Then there exists a $C > 0$ and a natural number M that depend on e, n and r such that for any finite field k of cardinality q and for any polynomials $f_1, \dots, f_r \in k[X_1, \dots, X_n]$ of degree less than or equal to e , if $V(k) \neq \emptyset$ then we have*

$$|\text{card}(V(k)) - \mu q^d| \leq Cq^{d-(1/2)}$$

for some $d \in \{0, \dots, \dim(V)\}$ and $\mu \in \{1, \dots, M\}$.

Proof. We will use induction on the dimension of V .

If $d = \dim(V) = 0$, then V is a finite set of points in $\tilde{k}^n$. Applying decomposition-intersection procedure to V , we get V^* . Let $V^* = V_1 \cup \dots \cup V_s$, where $V_1, \dots, V_s$ are the absolutely irreducible components of V^* . By Theorem 4.2, s is bounded by a

natural number M depending only on (e, n, r) . Furthermore, $\dim(V_i) \leq 0$ for each i , so $\text{card}(V_i) \leq 1$. $V(k) = V^*(k)$, so $V(k) = V_1(k) \cup \dots \cup V_s(k)$. Each $V_i(k)$ has at most one element because $V_i(k) \subseteq V_i$. So, $\text{card}(V(k)) \leq s$. Rearrange V_i 's so that $V_1, \dots, V_\mu$ are of dimension 0. This implies that V_i has a single element if $1 \leq i \leq \mu$ and is empty if $\mu < i \leq s$. Since V_i are k -algebraic, $V_i = V_i(k)$. Therefore $\text{card}(V_i(k)) = 1$ for $1 \leq i \leq \mu$ and $\text{card}(V_i(k)) = 0$ for $\mu < i \leq s$. Then, $\text{card}(V(k)) = \mu$ and hence

$$|\text{card}(V(k)) - \mu q^d| = |\mu - \mu q^0| = 0$$

Any positive constant C will satisfy the desired inequality.

Suppose that the proposition holds for varieties of dimension up to d .

Now, let $\dim(V) = d$ and $V(k) \neq \emptyset$. Apply intersection-decomposition procedure to V over k to get V^* . Let $V_1, \dots, V_s$ be the distinct absolutely irreducible components of V^* . So, they are k -algebraic. Now, we have

$$V(k) = V \cap k^n = V^* \cap k^n = (V_1 \cup \dots \cup V_s) \cap k^n = V_1(k) \cup \dots \cup V_s(k)$$

By Theorem 4.2, s is bounded by a natural number M depending only on (e, n, r) .

Assume that $V_1, \dots, V_\mu$ are of maximal dimension $d = \dim(V^*) \leq \dim(V)$ and $\dim(V_j) < d$ for $\mu < j \leq s$.

Claim 5.1. *For these values of d and μ , we have:*

$$|\text{card}(V(k)) - \mu q^d| \leq C q^{d-(1/2)}$$

for a positive constant C depending only on (e, n, r) .

By Theorem 4.2, each V_j is of the form $V(h_1, \dots, h_R)$ for polynomials $h_1, \dots, h_R \in k[\bar{Y}]$ of degree $\leq E$, where R and E only depends on (e, n, r) .

- (i) $|\text{card}(V_j(k)) - q^d| \leq C_1 q^{d-(1/2)}$ for $j \in \{1, \dots, \mu\}$ and a positive constant C_1 that depends only on (E, n, R) , since h_l is of degree $\leq E$ for $l \in \{1, \dots, R\}$ by Lang-Weil theorem.
- (ii) $|\text{card}(V_j(k))| \leq C_2 q^{d-1}$ for $\mu < j \leq s$ and a positive constant C_2 that depends only on (E, n, R) , since V_j is absolutely irreducible and $\dim(V_j) := \delta \leq d-1$. By Lang-Weil theorem:

$$\begin{aligned} |\text{card}(V_j(k)) - q^\delta| &\leq C q^{\delta-(1/2)} \\ \text{card}(V_j(k)) &\leq q^\delta + C q^{\delta-(1/2)} \\ \text{card}(V_j(k)) &\leq q^{d-1} + C q^{d-1} \end{aligned}$$

Taking $C_2 := 1 + C$, we get the result.

- (iii) $|\text{card}(V_i(k) \cap V_j(k))| \leq C_3 q^{d-1}$ for a positive constant C_3 that depends only on (E, n, R) , where $1 \leq i < j \leq s$, since $V_i(k) \cap V_j(k)$ is a k -algebraic set of dimension $\delta < d$. Therefore, by inductive hypothesis:

$$\begin{aligned} |\text{card}(V_i(k) \cap V_j(k)) - \mu q^\delta| &\leq C q^{\delta-(1/2)} \\ \text{card}(V_i(k) \cap V_j(k)) &\leq \mu q^\delta + C q^{\delta-(1/2)} \\ \text{card}(V_i(k) \cap V_j(k)) &\leq \mu q^{d-1} + C q^{d-1} \end{aligned}$$

Taking $C_3 := 1 + C$, we get the result. This result can be applied to any intersection of the form $V_{j(1)}(k) \cap \dots \cap V_{j(t)}(k)$ for $1 \leq j(1) < \dots < j(t) \leq s$ because the intersection is a k -algebraic set of dimension $< d$.

Then,

$$\begin{aligned}
& |\text{card}(V(k)) - \mu q^d| = |\text{card}(V_1(k) \cup \dots \cup V_s(k)) - \mu q^d| \\
& = \left| \sum_i \text{card}(V_i(k)) - \sum_{i,j} \text{card}(V_i(k) \cap V_j(k)) + \dots + (-1)^{s+1} \text{card}(V_1 \cap \dots \cap V_s) - \mu q^d \right| \\
& \leq \left| \sum_{1 \leq i \leq \mu} (\text{card}(V_i(k)) - q^d) \right| + \left| \sum_{\mu < i \leq s} \text{card}(V_i(k)) \right| + \left| \sum_{i,j} \text{card}(V_i(k) \cap V_j(k)) \right| + \\
& \dots + |\text{card}(V_1 \cap \dots \cap V_s)| \\
& \leq \mu C_1 q^{d-(1/2)} + (s - \mu) C_2 q^{d-1} + \binom{s}{2} C_3 q^{d-1} + \dots + C_{s+1} q^{d-1} \\
& \leq C q^{d-(1/2)}
\end{aligned}$$

Where we can take $C := \mu C_1 + (s - \mu) C_2 + \binom{s}{2} C_3 + \dots + C_{s+1} q^{d-1}$. This C only depends on s, E, n and R , and s is bounded by M , which depends on (E, n, R) . $\square$

Let $f \in \mathbb{Z}[\bar{X}, \bar{Y}]^r$, $\bar{X} = (X_1, \dots, X_m)$, $\bar{Y} = (Y_1, \dots, Y_n)$ and f_i are of degree less than or equal to e . Then, by Proposition 5.1 there is a positive constant C and a natural number M , both depending only on (e, n, r) , such that if $V_x(k) \neq \emptyset$ then

$$|\text{card}(V_x(k)) - \mu q^d| \leq C q^{d-(1/2)}$$

for some $d \in \{0, \dots, \dim(V_x)\}$ and $\mu \in \{1, \dots, M\}$. To see this, note that $f(\bar{x}, \bar{Y}) \in k[\bar{Y}]$ and $V_{\bar{x}}$ can be considered a k -algebraic set.

Proposition 5.2. *[3, p.122] With the above notation, for a fixed $d \in \{0, \dots, n\}$ and a fixed $\mu \in \{1, \dots, M\}$, there is an $\mathcal{L}$ -formula $S_{f,d,\mu}(\bar{X})$ such that for any field k ,*

$$S_{f,d,\mu}(k^m) = \{\bar{x} \in k^m : |\text{card}(V_{\bar{x}}(k)) - \mu q^d| \leq C q^{d-(1/2)}\}$$

Proof. The formula $S_{f,d,\mu}$ defined in Corollary 4.1 has the desired property. $\square$

Claim 5.2. *If q is sufficiently larger than C , then for each $\bar{x} \in k^m$ with $V_{\bar{x}}(k) \neq \emptyset$, there is a unique pair (d, μ) such that $k \models S_{f,d,\mu}(\bar{x})$.*

Proof. To see this, suppose that there is an $\bar{x} \in k^m$ and $d_1, d_2 \in \{0, \dots, n\}$, $\mu_1, \mu_2 \in \{1, \dots, M\}$ such that $k \models S_{f,d_1,\mu_1}(x)$ and $k \models S_{f,d_2,\mu_2}(\bar{x})$. So

$$|\text{card}(V_{\bar{x}}(k)) - \mu_1 q^{d_1}| \leq C q^{d_1 - (1/2)}$$

and

$$|\text{card}(V_{\bar{x}}(k)) - \mu_2 q^{d_2}| \leq C q^{d_2 - (1/2)}$$

which gives us

$$q^{d_1}(\mu_1 - \frac{C}{\sqrt{q}}) \leq \text{card}(V_{\bar{x}}(k)) \leq q^{d_1}(\mu_1 + \frac{C}{\sqrt{q}})$$

and

$$q^{d_2}(\mu_2 - \frac{C}{\sqrt{q}}) \leq \text{card}(V_{\bar{x}}(k)) \leq q^{d_2}(\mu_2 + \frac{C}{\sqrt{q}})$$

Since M is independent of q , we may assume that q is much larger than μ_1 and μ_2 . If we assume $d_1 < d_2$, then we get

$$q^{d_2}(\mu_2 - \frac{C}{\sqrt{q}}) \leq \text{card}(V_{\bar{x}}(k)) \leq q^{d_1}(\mu_1 + \frac{C}{\sqrt{q}})$$

$$q(\mu_2 - \frac{C}{\sqrt{q}}) \leq q^{d_2-d_1}(\mu_2 - \frac{C}{\sqrt{q}}) \leq \mu_1 + \frac{C}{\sqrt{q}}$$

$$q \leq \frac{\mu_1 + \frac{C}{\sqrt{q}}}{\mu_2 - \frac{C}{\sqrt{q}}} \leq \frac{M+1}{1-(1/2)}$$

which contradicts our assumption. □

If $V_{\bar{x}}(k) = \emptyset$, then

$$\begin{aligned} |\text{card}(V_{\bar{x}}(k)) - \mu q^d| &= \mu q^d \\ &> q^d \\ &> Cq^{d-(1/2)} \end{aligned}$$

So, there is no (d, μ) pair.

5.3. Estimates for Quantifier Free Definable Sets

The next step towards the main theorem is to get an estimate for the number of points of sets defined by quantifier-free formulas. We will convert the quantifier free formulas to a conjunction of polynomial inequalities and create a correspondence between definable sets and algebraic sets. Then, using the results from Proposition 5.1, we will find the estimate.

Lemma 5.1. *[3, Lemma 3.5] Let $\varphi(\bar{X}, \bar{Y})$ be a quantifier-free $\mathcal{L}_{ring}$ -formula, where $\bar{X} = (X_1, \dots, X_m)$ and $\bar{Y} = (Y_1, \dots, Y_n)$. Then there exist a $C > 0$ and a natural number M that depend on φ such that for any finite field k and for any $\bar{x} \in k^m$, if $\varphi(\bar{x}, k^n) \neq \emptyset$ then there exist $d \in \{0, \dots, n\}$ and $\mu \in \{1, \dots, M\}$ with $|\text{card}(\varphi(\bar{x}, k^n)) - \mu q^d| \leq Cq^{d-(1/2)}$.*

Proof. Since φ is quantifier-free, it can be written as $\bigvee_{\nu=1}^N \bigwedge_{j=1}^M \psi_{\nu j}$, where $\psi_{\nu j}$ is an atomic formula or the negation of an atomic formula for each ν and j . In the language of rings, each $\psi_{\nu j}$ is either $f(\bar{X}, \bar{Y}) = 0$ or $f(\bar{X}, \bar{Y}) \neq 0$ where $f(\bar{X}, \bar{Y}) \in \mathbb{Z}[\bar{X}, \bar{Y}]$. Therefore, $\varphi(\bar{X}, \bar{Y})$ is a formula of the form $\bigvee_{1 \leq \nu \leq N} (f_\nu(\bar{X}, \bar{Y}) = 0 \wedge g_\nu(\bar{X}, \bar{Y}) \neq 0)$, where f_ν and g_ν are polynomials in $\mathbb{Z}[\bar{X}, \bar{Y}]^{r(\nu)}$. In fact, $g_\nu(\bar{X}, \bar{Y})$ can be chosen to be in $\mathbb{Z}[\bar{X}, \bar{Y}]$ because instead of $g_1(\bar{X}, \bar{Y}) \neq 0 \wedge \dots \wedge g_r(\bar{X}, \bar{Y}) \neq 0$, then we can use the equivalent formula $\left(\prod_{i=1}^r g_i(\bar{X}, \bar{Y}) \right) \neq 0$.

Working with a positive formula will make the connection between definable sets and algebraic sets, since quantifier-free positive formulas in the language of rings are formed by conjunctions of polynomial equalities. To eliminate polynomial inequalities, new variables $\bar{Y}' := (Y_{n+1}, \dots, Y_{n+N})$ are introduced. Consider the formula $\varphi'(\bar{X}, \bar{Y}, \bar{Y}')$ which is defined as

$$\bigvee_{1 \leq \nu \leq N} (f_\nu(\bar{X}, \bar{Y}) = 0 \wedge g_\nu(\bar{X}, \bar{Y}) \cdot Y_{n+\nu} = 1 \wedge \bigwedge_{\lambda \neq \nu} Y_{n+\lambda} = 0)$$

The formula φ' is a positive quantifier-free formula as desired. Therefore $\varphi'(\bar{x}, \bar{Y}, \bar{Y}')$ defines an F -algebraic set over any field F and for any $\bar{x} \in F^m$.

Proposition 5.1 can be applied to the set $\varphi'(x, k^{n+N})$. Indeed, the numbers e , n and r needed in the proposition are all given in φ' . Recall that e is the bound for the degrees of polynomials defining the algebraic set, n is the number of variables (this is $n + N$ in our case), and r is the number of polynomials defining the algebraic set. Hence, there are $C > 0$ and $M \in \mathbb{Q}^+$ that depend on φ' such that for any finite field k of cardinality q and for any $\bar{x} \in k^m$ if $\varphi'(\bar{x}, k^{n+N})$ is nonempty, then there are $d \in \{0, 1, \dots, \dim(V)\}$ and $\mu \in \{1, \dots, M\}$ satisfying the inequality

$$|\text{card}(\varphi'(\bar{x}, k^{n+N})) - \mu q^d| \leq C q^{d-(1/2)}$$

Note that the variables $\bar{Y}'$ depend on the values of the variables $\bar{Y}$, so $\dim(V) \leq n$. Therefore, d is in the set $\{0, \dots, n\}$.

Although φ' is a different formula than φ , the sets they define have the same number of elements over any finite field. Consider the function

$$\begin{aligned} \eta : \phi'(\bar{x}, k^{n+N}) &\rightarrow \phi(\bar{x}, k^n) \\ (\bar{y}, \bar{y}') &\mapsto \bar{y} \end{aligned}$$

This function is a bijection for every $\bar{x} \in k^m$. We will first prove that this function maps $\varphi'(\bar{x}, k^{n+N})$ to $\varphi(\bar{x}, k^n)$. Let $(\bar{y}, \bar{y}')$ be in $\varphi'(\bar{x}, k^{n+N})$. Let $\bar{y}' = (y'_1, \dots, y'_N)$. Then

$$f_\nu(\bar{x}, \bar{y}) = 0 \wedge g_\nu(\bar{x}, \bar{y}) \cdot y'_\nu = 1 \wedge \bigwedge_{\lambda \neq \nu} y'_\lambda = 0$$

holds for some $\nu \in \{1, \dots, N\}$. Hence, $f_\nu(\bar{x}, \bar{y}) = 0 \wedge g_\nu(\bar{x}, \bar{y}) \neq 0$ holds. Therefore $(\bar{x}, \bar{y}) \in \varphi(\bar{x}, k^n)$. Now we will prove that this map is an injection. Let $(\bar{y}, \bar{\xi})$ and $(\bar{y}, \bar{\zeta})$ be in $\varphi'(\bar{x}, k^{n+N})$. Since the disjuncts defining φ are disjoint, there is a unique $\nu \in \{1, \dots, N\}$ such that $f_\nu(\bar{x}, \bar{y}) = 0 \wedge g_\nu(\bar{x}, \bar{y}) \neq 0$ holds. Therefore $\xi_\nu = \zeta_\nu = 1/g_\nu(\bar{x}, \bar{y})$ and $\xi_i = \zeta_i = 0$ for $i \neq \nu$. Lastly, we will show that η is surjective. Let $\bar{y}$ be in $\varphi(\bar{x}, k^n)$. Again, there is a unique $\nu \in \{1, \dots, N\}$ such that $f_\nu(\bar{x}, \bar{y}) = 0 \wedge g_\nu(\bar{x}, \bar{y}) \neq 0$ holds. Choose $\bar{y}'$ such that $y_{n+\nu} = 1/g_\nu(\bar{x}, \bar{y})$ and $y_{n+i} = 0$ for $i \neq \nu$. Then $\eta(\bar{y}, \bar{y}') = \bar{y}$.

We now see that $\text{card}(\bar{x}, k^n) = \text{card}(\bar{x}, k^{n+N})$. Also notice that φ' is given by the formula φ and hence the constants C and M found above depend on φ . Therefore, we get the desired result. $\square$

5.4. Estimates for Arbitrary Definable Sets over Finite Fields

We have gathered enough information to prove the main theorem:

Theorem 5.2 (Main Theorem). [3, Theorem 3.7] Let $\varphi(X, Y)$ be an arbitrary $\mathcal{L}_{\text{ring}}$ -formula. Then there exists a $C > 0$ and a finite set D of pairs (d, μ) with $d \in \{0, \dots, n\}$ and $\mu > 0$ a rational number, both depending on φ , such that for any $\bar{x} \in k^m$ with $\varphi(\bar{x}, k^n) \neq \emptyset$ there exists some $(d, \mu) \in D$ satisfying $|\text{card}(\phi(x, k^n)) - \mu q^d| \leq C q^{d-(1/2)}$

Proof. If the language $\mathcal{L}_{\text{ring}}$ is enriched as it was done in Section 3.5, then we can use the Proposition 3.2 to conclude that φ is equivalent to a conjunction of formulas of the form $\exists T (g(\bar{c}, \bar{X}, \bar{Y}, T) = 0)$, where $g(\bar{C}, \bar{X}, \bar{Y}, T) \in \mathbb{Z}[\bar{C}, \bar{X}, \bar{Y}, T]$. To work on the language $\mathcal{L}_{\text{ring}}$, we need to remove the constant symbols. This is done by replacing constant symbols $\bar{c} = (c_1, \dots, c_M)$ with variable symbols $\bar{X}' = (X'_1, \dots, X'_M)$. Now we have a formula $G(\bar{X}, \bar{X}', \bar{Y}, T) \in \mathbb{Z}(\bar{X}, \bar{X}', \bar{Y}, T)$, which is equivalent to the formula $g(\bar{c}, \bar{X}, \bar{Y}, T)$. Instead of $(X_1, \dots, X_m, X'_1, \dots, X'_M)$, we will only write $\bar{X} = (X_1, \dots, X_m)$, and for $G(\bar{X}, \bar{Y}, T)$, we will use $g(\bar{X}, \bar{Y}, T)$.

Note that for some $(\bar{x}, \bar{y}) \in F^{m+n}$, we might have $g(\bar{x}, \bar{y}, T) = 0$ for any value of T . This can be prevented if we use the equivalent formula

$$\left(\bigwedge_i g_i(\bar{X}, \bar{Y}) = 0 \right) \vee \left(\exists T, g(\bar{X}, \bar{Y}, T) = 0 \wedge \exists U, \prod_i (g_i(X, Y) \cdot U - 1) = 0 \right)$$

where $g(\bar{X}, \bar{Y}, T) = \sum_i g_i(\bar{X}, \bar{Y}) T^i$. To see the equivalence, note that if there is a T with $g(\bar{X}, \bar{Y}, T) = 0$, then either the polynomial $\sum_i g_i(\bar{X}, \bar{Y}) T^i$ has all its coefficients $g_i(\bar{X}, \bar{Y}) = 0$, or there are some coefficients that are nonzero and $g(\bar{X}, \bar{Y}, \bar{T}) = 0$. In the latter case, the nonzero coefficient has a multiplicative inverse, so there is a U such that $g_i(X, Y) \cdot U - 1 = 0$. Moreover, the two disjuncts in this formula are disjoint. Suppose on the contrary that $(\bar{x}, \bar{y})$ satisfy both disjuncts. Second part of the second disjunct implies that there exists i with $(g_i(\bar{x}, \bar{y}) \cdot U - 1) = 0$ for some U , so $g_i(X, Y) \neq 0$, contradicting $(\bigwedge_i g_i(X, Y) = 0)$.

By applying Lemma 3.7 on the conjunction of the formulas

$$(\bigwedge_i g_i(\bar{X}, \bar{Y}) = 0) \vee \left(\exists T, g(\bar{X}, \bar{Y}, T) = 0 \wedge \exists U \prod_i (g_i(X, Y) \cdot U - 1) = 0 \right)$$

we get a disjunction of formulas of the form

$$f(\bar{X}, \bar{Y}) = 0 \wedge \exists \bar{Z} (\bigwedge_j h_j(\bar{X}, \bar{Y}, Z_j) = 0)$$

where $f \in \mathbb{Z}[\bar{X}, \bar{Y}]^r$, $h_j \in \mathbb{Z}[\bar{X}, \bar{Y}, Z_j]$ for $1 \leq j \leq k$, and $\bar{Z} = (Z_1, \dots, Z_k)$. Each disjunct here is also disjoint. This fact can be observed easily if we use the notation in Lemma 3.7. Recall that we showed

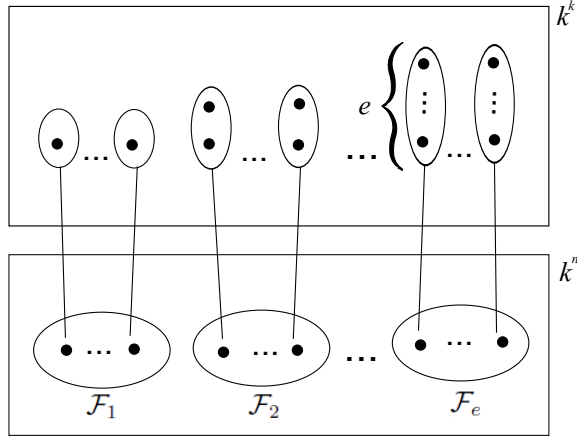
$$\bigwedge_{i=1}^k (A_{1,i} \vee A_{2,i}) = \bigvee_{S \subseteq K} ((\bigwedge_{j \in S} A_{1,j}) \wedge (\bigwedge_{j \notin S} A_{2,j}))$$

In our case, we set $A_{1,j}$ to be the formula $\bigwedge_i g_i(\bar{X}, \bar{Y}) = 0$ and $A_{2,j}$ to be the formula $\exists T, g(\bar{X}, \bar{Y}, T) = 0 \wedge \exists U \prod_i (g_i(X, Y) \cdot U - 1) = 0$. For each j , $A_{1,j}$ and $A_{2,j}$ are disjoint. Each disjunct in $\bigvee_{S \subseteq K} ((\bigwedge_{j \in S} A_{1,j}) \wedge (\bigwedge_{j \notin S} A_{2,j}))$ is given by a different subset of $K = \{1, \dots, k\}$. Let S and S' be two distinct subsets of K . Without loss of generality, assume that there is $j \in S \setminus S'$. So, the disjunct given by S include $A_{1,j}$ while the disjunct given by S' include $A_{2,j}$. Therefore, these disjuncts are disjoint.

Define the formula

$$\Psi(\bar{X}, \bar{Y}, \bar{Z}) := f(\bar{X}, \bar{Y}) = 0 \wedge (\bigwedge_j h_j(\bar{X}, \bar{Y}, Z_j) = 0)$$

We will now deal with the number of $\bar{Z}$'s. Consider the polynomial $h_j \in \mathbb{Z}[\bar{X}, \bar{Y}, Z_j]$.

Figure 5.1. The sets $\mathcal{F}_j$.

For any $(\bar{x}, \bar{y})$ in any field, $h_j(\bar{x}, \bar{y}, Z_j)$ has a bound on the solutions in Z_j , namely the degree of h_j as a polynomial in Z_j over $\mathbb{Z}[\bar{X}, \bar{Y}]$. Since there are finitely many h_j 's and finitely many disjuncts, a number e can be found such that for any field F and any $(\bar{x}, \bar{y}) \in F^{m+n}$, the number of distinct $\bar{z}$'s satisfying $F \models \Psi(\bar{x}, \bar{y}, \bar{z})$ is less than or equal to e .

Fix a field k and an $\bar{x} \in k^m$ with $\Psi(\bar{x}, k^{n+k}) \neq \emptyset$. Fix some $d \in \{0, \dots, n+k\}$ and $\mu \in \{1, \dots, N\}$ satisfying the last statement. Define the sets:

$$\mathcal{F} := \varphi(\bar{x}, k^n)$$

$$\mathcal{F}_j := \{\bar{y} \in k^n : \text{card}(\Psi(\bar{x}, \bar{y}, k^k)) = j\} \text{ for } j=1, \dots, e$$

$$\mathcal{G} := \Psi(\bar{x}, k^{n+k})$$

Since $\mathcal{F} = \bigcup_{j \in \{1, \dots, e\}} \mathcal{F}_j$ and $\mathcal{F}_i \cap \mathcal{F}_j = \emptyset$ for $i \neq j$ we have

$$\text{card}(\mathcal{F}) = \text{card}(\mathcal{F}_1) + \text{card}(\mathcal{F}_2) + \dots + \text{card}(\mathcal{F}_e) \quad (5.1)$$

Now, let $j \in \{1, \dots, e\}$ be given and let $\bar{y} \in \mathcal{F}_j$ be arbitrary. Then the set $\Psi(\bar{x}, \bar{y}, k^k)$ has j elements. Let $\Psi(\bar{x}, \bar{y}, k^k) = \{\bar{z}_1, \dots, \bar{z}_j\}$. This means that $(\bar{y}, \bar{z}_1), \dots, (\bar{y}, \bar{z}_j) \in \mathcal{G}$ and hence

$$\text{card}(\mathcal{G}) = \text{card}(\mathcal{F}_1) + 2 \cdot \text{card}(\mathcal{F}_2) + \dots + e \cdot \text{card}(\mathcal{F}_e). \quad (5.2)$$

Combining Equations 5.1 and 5.2, we get $\text{card}(\mathcal{F}) \geq \text{card}(\mathcal{G})/e$. Using this on the inequality $|\text{card}(\Psi(\bar{x}, k^{n+k})) - \mu q^d| \leq Aq^{d-(1/2)}$, we have

$$|\text{card}(\mathcal{G}) - \mu q^d| \leq Aq^{d-(1/2)} \quad (5.3)$$

$$\mu q^d - Aq^{d-(1/2)} \leq \text{card}(\mathcal{G}) \quad (5.4)$$

$$\frac{\mu}{e} q^d - \frac{A}{e} q^{d-(1/2)} \leq \text{card}(\mathcal{F}) \quad (5.5)$$

For sufficiently large q , we must have $d \leq n$ as $\mathcal{F} \subseteq k^n$.

All the work up to this point gave us a way to calculate the cardinality of $\mathcal{F}$, *i.e.* the set $\varphi(\bar{x}, k^n)$, in terms of the sets $\mathcal{F}_1, \dots, \mathcal{F}_e$. In the rest of the proof we will find estimates for the number of elements of $\mathcal{F}_j$'s.

Define the formulas

$$\Psi_j(\bar{X}, \bar{Y}, \bar{Z}^1, \dots, \bar{Z}^j) := \left(\bigwedge_{1 \leq i \leq j} \Psi(\bar{X}, \bar{Y}, \bar{Z}^i) \right) \wedge \left(\bigwedge_{i(1) \neq i(2)} Z^{i(1)} \neq Z^{i(2)} \right)$$

for $1 \leq j \leq e$, where $\bar{Z}^i = (Z_1^i, \dots, Z_k^i)$ for $1 \leq i \leq j$. Define $\mathcal{H}_j := \Psi_j(\bar{x}, k^{n+jk})$. Fix $\bar{y} \in \mathcal{F}_j$. Suppose that the points $(\bar{x}, \bar{y}, \bar{z}^1), \dots, (\bar{x}, \bar{y}, \bar{z}^j)$ are in the set defined by Ψ . Then, the points $(\bar{x}, \bar{y}, \bar{z}^{\sigma(1)}, \bar{z}^{\sigma(2)}, \dots, \bar{z}^{\sigma(j)})$ are in the set defined by Ψ_j for each permutation σ on $\{1, \dots, j\}$. Hence, for each $\bar{y} \in \mathcal{F}_j$ there are $j!$ elements in $\mathcal{H}_j$. Now, fix $\bar{y} \in \mathcal{F}_{j+t}$ for some $t \in \{0, \dots, e-j\}$. Suppose that the points $(\bar{x}, \bar{y}, \bar{z}^1), \dots, (\bar{x}, \bar{y}, \bar{z}^{j+t})$ satisfy Ψ .

Let $i(1), \dots, i(j)$ be j distinct elements in $\{1, \dots, j+t\}$. Then $(\bar{x}, \bar{y}, \bar{z}^{i(1)}, \bar{z}^{i(2)}, \dots, \bar{z}^{i(j)})$ satisfy Ψ_j . There are $j! \binom{j+t}{j} = \frac{(j+t)!}{t!}$ many such elements. So,

$$\text{card}(\mathcal{H}_j) = \sum_{t=0}^{e-j} \frac{(j+t)!}{t!} \text{card}(\mathcal{F}_{j+t}) \quad (5.6)$$

Claim 5.3. $\text{card}(\mathcal{F}) = r_1 \text{card}(\mathcal{H}_1) + \dots + r_e \text{card}(\mathcal{H}_e)$ for rationals $r_1, \dots, r_e$ depending only on e .

Proof. We will show that each $\text{card}(\mathcal{F}_i)$ is given by a linear combination of $\text{card}(\mathcal{H}_j)$'s over rational numbers. This will be done by induction on i . Fix e . We have $\text{card}(\mathcal{F}_e) = \frac{\text{card}(\mathcal{H}_e)}{e!}$. Now, suppose that for some $j \in \{1, \dots, e\}$ and for all $i \in \{0, \dots, j-1\}$ there are rational numbers $s_{e-i}, \dots, s_e$ such that

$$\text{card}(\mathcal{F}_{e-i}) = s_{e-i} \text{card}(\mathcal{H}_{e-i}) + s_{e-i+1} \text{card}(\mathcal{H}_{e-i+1}) + \dots + s_e \text{card}(\mathcal{H}_e) \quad (5.7)$$

Now,

$$\text{card}(\mathcal{H}_{e-j}) = \sum_{t=0}^{e-(e-j)} \frac{(e-j+t)!}{t!} \text{card}(\mathcal{F}_{e-j+t}) \quad (5.8)$$

$$= \sum_{t=0}^j \frac{(e-j+t)!}{t!} \text{card}(\mathcal{F}_{e-j+t}) \quad (5.9)$$

$$= (e-j)! \text{card}(\mathcal{F}_{e-j}) + \sum_{t=1}^{(j+1)} \frac{(e-j+t)!}{t!} \text{card}(\mathcal{F}_{e-j+t}) \quad (5.10)$$

So,

$$\text{card}(\mathcal{F}_{e-j}) = \frac{1}{(e-j)!} \left(\text{card}(\mathcal{H}_{e-(j+1)}) - \sum_{t=1}^{(j+1)} \frac{(e-j+t)!}{t!} \text{card}(\mathcal{F}_{e-j+t}) \right) \quad (5.11)$$

Combining the last equation and Equation 5.7 we see that there are rational numbers $r'_{e-(j+1)}, \dots, r'_e$ such that

$$\text{card}(\mathcal{F}_{e-j}) = r'_{e-j} \text{card}(\mathcal{H}_{e-j}) + r'_{e-(j+1)} \text{card}(\mathcal{H}_{e-j+1}) + \dots + r'_e \text{card}(\mathcal{H}_e) \quad (5.12)$$

By induction, we see that each $\mathcal{F}_i$ is given by $\sum_{j=i}^e r_{ij} \text{card}(\mathcal{H}_j)$ for some rational numbers $r_{ii}, \dots, r_{ie}$. Since $\mathcal{F}$ is given by the sum $\sum_{i=1}^e \mathcal{F}_i$, put r_i to be the rational number $\sum_{j=i}^e r_{ij}$ for each i . $\square$

Claim 5.4. $\text{card}(\mathcal{H}_j) \leq \kappa \cdot \text{card}(\mathcal{G})$ for some constant κ .

Proof. Fix $j \in \{1, \dots, e\}$. By Equation 5.2 and Equation 5.6 we have

$$\begin{aligned} e! \cdot \text{card}(\mathcal{G}) &= e! \cdot \text{card}(\mathcal{F}_1) + e! \cdot 2 \cdot \text{card}(\mathcal{F}_2) + \dots + e! \cdot e \cdot \text{card}(\mathcal{F}_e) \\ &\geq \frac{(j)!}{0!} \text{card}(\mathcal{F}_j) + \frac{(j+1)!}{1!} \text{card}(\mathcal{F}_{j+1}) + \dots + \frac{e!}{(e-j)!} \text{card}(\mathcal{F}_e) \quad \square \end{aligned}$$

If we apply Lemma 5.1 to the quantifier-free formula Ψ_j , we get some d_j and μ_j such that $|\text{card}(\mathcal{H}_j) - \mu_j q_j^{d_j}| \leq A_j q_j^{d_j-1/2}$. Moreover, by Claim 5.4 for some constant κ we have $\text{card}(\mathcal{H}_j) \leq \kappa \text{card}(\mathcal{G})$. Using (5.3) we get $\text{card}(\mathcal{H}_j) \leq \kappa A_j q^{d-1/2} + \kappa \mu_j q^d$. When q is large enough, we must have $d_j \leq d$. Therefore, by allowing μ_j to be zero, we get $|\text{card}(\mathcal{H}_j) - \mu_j q^d| \leq A_j q^{d-1/2}$

Combining Equation 5.12 and the above result gives

$$\begin{aligned} |\text{card}(\mathcal{F}) - (r_1 \mu_1 + \dots + r_e \mu_e) q^d| &\leq \sum_j |r_j \text{card}(\mathcal{H}_j) - r_j \mu_j q^d| \\ &\leq \sum_j |r_j| A_j q^{d-1/2} \end{aligned}$$

Put $C := \sum_j |r_j| A_j$, and $\mu := \sum_i r_i \mu_i$. By Equation 5.5, for large enough cardinalities q , we have $\mu > 0$. $\square$

Here is a result of the main theorem that will be used in the next chapter.

Proposition 5.3. *[3, Proposition 3.8] Let φ be an $\mathcal{L}_{ring}$ -formula. Let C be the constant and D be the finite set given by Theorem 5.2. Then for any $(d, \mu) \in D$, there is a formula $\varphi_{d,\mu}(\bar{X})$ such that for any field k and for any $\bar{x} \in k^m$*

$$k \models \varphi_{d,\mu}(\bar{x}) \Leftrightarrow |\text{card}(\phi(\bar{x}, k^n)) - \mu q^d| \leq C q^{d-(1/2)}$$

6. APPLICATIONS

In this chapter, we will investigate the meaning of the constants d and μ found in the main theorem. The constant d will be interpreted as the algebraic dimension of a definable set, that is the dimension of the closure of the definable set in the Zariski topology. The constant μ will correspond to a probability measure on the set of definable subsets of a given definable set. Lastly, we will present a result about this measure μ found by Halupczok [19].

6.1. Interpretation of the Numbers d and μ : Dimension and Measure

Define *algebraic dimension* of a subset S of F^n , where F is an infinite field, to be the dimension of the closure of S in $\tilde{F}^n$ with respect to Zariski Topology. Algebraic dimension of S will be denoted by $\text{algd}(\text{dim}(S))$.

Let F be a pseudo-finite field. Suppose that $S \subseteq F^n$ is definable by the formula $\varphi(x, F^n)$ for some $x \in F^m$. By Theorem 5.2, there is a finite set D of pairs (d, μ) determined by the formula φ . Moreover, by Proposition 5.3 for each (d, μ) there is a formula $\varphi_{d,\mu}(X)$ such that the set of $x \in F^m$ that satisfies $|\text{card}(\varphi(x, F^n)) - \mu q^d| \leq Cq^{d-(1/2)}$ is definable by $\varphi_{d,\mu}(X)$. Since $\varphi_{d,\mu}(X)$ is independent of the field it is defined over, by Claim 5.2, there is exactly one pair (d, μ) such that $k \models \varphi_{d,\mu}(x)$ for finite fields k of sufficiently large cardinality. This means that there is exactly one pair (d, μ) such that $k \models \varphi_{d,\mu}(x)$ for almost all finite fields k . Since F is pseudo-finite $F \models \varphi_{d,\mu}(x)$. By a similar argument, any two formulas defining S will have the same pair (d, μ) , so the pair (d, μ) is independent of the defining formula, too.

Proposition 6.1. [3, Proposition 4.9] *If d is the unique natural number discussed above for the definable set S , then $d = \text{algd}(\text{dim}(S))$.*

Proof. If S is the set of F -points of an F -algebraic set $V \subseteq \tilde{F}^n$, then by the proof of Proposition 5.1, $d = \dim(V^*)$. F is a pseudo-finite field, therefore $V(F)$ is dense in V^*

in the Zariski topology, that is, . Since $S = V(F)$, we have $\text{alldim}(S) = d$.

Apply Proposition 3.3 on S to get $e, k \in \mathbb{N}$ with $S = \pi(V(F))$ for $V \subseteq \tilde{F}^{n+k}$ an F -algebraic set, π is a projection from $\tilde{F}^{n+k}$ to $\tilde{F}^k$ with $\pi : (x_1, \dots, x_n, y_1, \dots, y_k) \mapsto (x_1, \dots, x_n)$, and $V_{\bar{x}} := \{\bar{y} \in \tilde{F}^k : (\bar{x}, \bar{y}) \in V\}$ is of cardinality less than or equal to e for all $\bar{x} \in \tilde{F}^n$. Since for each $\bar{x} \in S$ we have $\text{card}(V_{\bar{x}}) \leq e$ and $V(F) = \bigcup_{\bar{x} \in S} V_{\bar{x}} \cap F^k$, we have $\text{card}(V(F)) \leq \text{card}(S)$. Therefore, if the pair (d', μ') is the pair associated to the set $V(F)$, then for large values of q we get $d = d'$.

The set $V(F)$ is dense in V^* in Zariski topology, therefore $\dim(V^*) = \dim(V(F)) = d$. Also, the set $\pi(V(F))$ is dense in $\pi(V^*)$. Since each $V_{\bar{x}}^*$ also have $\leq e$ elements, we get $\text{alldim}(S) = \dim(V^*) = d$. $\square$

Let F be a pseudo-finite field, S a subset of F^m definable by $\phi(x, F^n)$ and (d, μ) be the pair given by Theorem 5.2. Define $\text{Def}(S)$ to be the boolean algebra of definable subsets of S . Define $\mu_S : \text{Def}(S) \rightarrow [1, 0]$ as follows: Put $\mu_S(\emptyset) = 0$, and for any nonempty $A \subseteq S$ with corresponding pair (e, ν) , $\mu_S(A) = 0$ if $e < d$, and $\mu_S(A) = \nu/\mu$ if $e = d$.

Proposition 6.2. *[3, p.130] μ_S is a finitely additive probability measure on S taking only rational values.*

Proof. For any A with corresponding pair (e, ν) , ν and μ are rational, therefore ν/μ is also rational. Moreover, for finite fields of sufficiently large cardinality q , $\text{card}(A)$ is close to νq^d and $\text{card}(S)$ is close to μq^d . Hence $\nu \leq \mu$ and so $\nu/\mu \leq 1$.

Now suppose that A and B are disjoint subsets of S with defining formulas ψ_1, ψ_2 and corresponding pairs (d, ν_1) and (d, ν_2) , respectively. Then $|\text{card}(\psi_1(x, k^n)) - \nu_1 q^d| < C_1 q^{d-(1/2)}$ and $|\text{card}(\psi_2(x, k^n)) - \nu_2 q^d| < C_2 q^{d-(1/2)}$ for a finite field k of cardinality q .

Therefore

$$\begin{aligned}
& |\text{card}(\psi_1(x, k^n)) + \text{card}(\psi_2(x, k^n)) - \nu_1 q^d - \nu_2 q^d| \\
& \leq |\text{card}(\psi_1(x, k^n)) - \nu_1 q^d| + |\text{card}(\psi_2(x, k^n)) - \nu_2 q^d| \\
& < C_1 q^{d-(1/2)} + C_2 q^{d-(1/2)} \\
& = (C_1 + C_2) q^{d-(1/2)}
\end{aligned}$$

for any q . Hence $(d, \nu_1 + \nu_2)$ is the pair corresponding to the formula $\psi_1 \vee \psi_2$ defining $A \cup B$. Thus

$$\mu_S(A \cup B) = \frac{\nu_1 + \nu_2}{\mu} = \frac{\nu_1}{\mu} + \frac{\nu_2}{\mu} = \mu_S(A) + \mu_S(B)$$

□

In [19] Halupczok proved that the measure defined in [3] can be extended to perfect PAC-fields with bounded Galois groups that are not $\hat{\mathbb{Z}}$ if the Galois group is pro-cyclic. However, such an extension is not possible when the Galois group is not pro-cyclic.

REFERENCES

1. Hasse, H., “Zur Theorie der abstrakten elliptischen Funktionenkörper III. Die Struktur des Meromorphismenrings. Die Riemannsche Vermutung.”, *Journal für die reine und angewandte Mathematik*, Vol. 175, pp. 193–208, 1936.
2. Lang, S. and A. Weil, “Number of Points of Varieties in Finite Fields”, *American Journal of Mathematics*, Vol. 76, pp. 819–827, 1954.
3. Chatzidakis, Z., L. van den Dries and A. Macintyre, “Definable Sets over Finite Fields”, *Journal für die reine und angewandte Mathematik*, Vol. 427, pp. 107–135, 1992.
4. Hrushovski, E., *Stable Group Theory and Approximate Subgroups*, 2011, <http://arxiv.org/abs/0909.2190v4>, accessed at August 2013.
5. Tao, T., *Reading Seminar: “Stable Group Theory and Approximate Subgroups”*, by Ehud Hrushovski, 2009, <http://terrytao.wordpress.com/2009/10/15/reading-seminar-stable-group-theory-and-approximate-subgroups-by-ehud-hrushovski/>, accessed at August 2013.
6. Lang, S., *Introduction to Algebraic Geometry*, Addison-Wesley Publishing Company, Inc., Reading, MA, USA, 1973.
7. Fried, M. D. and M. Jarden, *Field Arithmetic*, Springer-Verlag, Berlin, 2008.
8. Hulek, K., *Elementary Algebraic Geometry*, American Mathematical Society, Providence, RI, USA, 2003.
9. Hungerford, T. W., *Algebra*, Graduate Texts in Mathematics, Springer, New York, NY, USA, 1974.

10. Lang, S., *Algebra*, Graduate Texts in Mathematics, Springer, New York, NY, USA, 2002.
11. Marker, D., *Model Theory: An Introduction*, Springer-Verlag New York, Inc., New York, NY, USA, 2002.
12. Tent, K. and M. Ziegler, *A Course in Model Theory*, Cambridge University Press, Cambridge, UK, 2012.
13. Hrushovski, E., “Pseudo-Finite Fields and Related Structures”, *Quaderni di Matematica*, Vol. 11, pp. 151–212, 2002.
14. Chatzidakis, Z., “Model Theory of Finite Fields and Pseudo-Finite Fields”, *Annals of Pure and Applied Logic*, Vol. 88, pp. 95–108, 1997.
15. van den Dries, L., “A Remark on Ax’s Theorem on Solvability Modulo Primes”, *Mathematische Zeitschrift*, Vol. 208, pp. 65–70, 1991.
16. Ax, J., “The Elementary Theory of Finite Fields”, *Annals of Mathematics*, Vol. 88, pp. 239–271, 1968.
17. Seidenberg, A., “Constructions in algebra”, *Transactions of the American Mathematical Society*, Vol. 197, pp. 273–313, 1974.
18. Schmidt, K. and L. van den Dries, “Bounds in the Theory of Polynomial Rings over Fields. A Nonstandard Approach.”, *Inventiones Mathematicae*, Vol. 76, pp. 77–92, 1984.
19. Halupczok, I., “A Measure for Perfect PAC fields with Pro-cyclic Galois Group”, *Journal of Algebra*, Vol. 310, pp. 371–395, 2007.