# Quantifier Elimination for Real Closed Fields

Maarten Boonekamp (6243150)

22th of January 2021

A Bachelor Thesis Mathematics (7,5 ECTS)
under guidance of Jaap van Oosten

Universiteit Utrecht

# 1 Introduction

As the title already suggests, in this paper we will prove that the theory of real closed fields admits quantifier elimination. This is an important result in model theory, especially when it comes to proving results in real algebraic geometry. Model-theoretic proofs of theorems such as "Hilbert's 17th Problem" and the "Real Nullstellensatz" are much shorter than proofs using geometrical techniques.

In order to understand the results mentioned above, we will first determine what rings, fields and especially real closed fields are, from a model-theoretic point of view(Section 2). Then in section 3, we will introduce the compactness theorem and explain its name with the help of what are called "types". After this, as an aside, we will look at the space of types over algebraically closed fields(section 4) and study some properties of the theory of algebraically closed fields(section 5.1). This in order to become familiar with the application of model theory to fields.

In section 5.2 some properties of the theory of real closed fields will be discussed, including the "Tarski-Seidenberg Theorem", which will be proven in section 6. We will continue with a number of simple corollaries of this theorem in section 7, one of which is "Hilbert's 17th Problem", and section 8 will give a proof the "Real Nullstellensatz". Lastly, we will see in section 9 an example of how model theory and quantifier elimination can be applied to real algebraic geometry when we talk about "semi-algebraic" functions, which are known as "definable" functions in model theory.

For the research on the subject of this paper in general, I mainly have made use of two publications that give a good introduction to model theory and the model theory of fields: "A Course in Model Theory" by Martin Ziegler and Katrin Tent [6] and "Model Theory of Fields" by David Marker, Margit Messner and Anand Pillay [3]. For basic notions I also have used "Sets, Models and Proofs" by Ieke Moerdijk and Jaap van Oosten [4]. Furthermore, "Real Algebraic Geometry" by Jacek Bochnak, Michel Coste and Marie-Francoise [2] and the lecture notes "Rings and Galois Theory" by Frits Beukers [1] were a useful resource of information for section 2 and 9, respectively. Similarly, the notes of a seminar by Victoria Noquez [5] were of good use in section 6. For the historical background of the "Tarski-Seidenberg" theorem, I used an article by Lou van den Dries [7].

This thesis is written for readers who have a mathematical background and already are familiar with basic concepts in Model Theory such as for example logical sentences and theories. However, I tried to make the thesis as self-contained as possible, without affecting the readability too much. I hope readers will get a good picture of how model theory, and in specific quantifier elimination, can be used in order to prove algebraic results.

# 2 Theories of rings

In this section I will shortly introduce the concept of rings, fields and ideals and their corresponding theories. If you are already familiar with this, you can skip this section.
Firstly, let $L_R$ be the language of rings, consisting of the constants 1 and 0, and the two-place function symbols + and $\cdot$. Then the $L_R$-theory of commutative rings, $T_{rings}$, consists of axioms saying that R is an Abelian group under adding (1-4), that multiplication is associative (5) and distributive (6-7) and that there is a unit for multiplication (8). Furthermore, we have that multiplication is commutative (9). This results in te following $L_R$-sentences:

1. $\forall xy(x + y = y + x)$
2. $\forall xyz(x + (y + z) = (x + y) + z)$
3. $\forall x(0 + x) = (x + 0)$
4. $\forall x\exists y(x + y = y + x = 0)$
5. $\forall xyz(x(yz) = (xy)z)$
6. $\forall xyz(x(y + z) = xy + xz)$
7. $\forall xyz((y + z)x = yx + zx)$
8. $\forall x(1 \cdot x = x \cdot 1 = x)$
9. $\forall xy(xy = yx)$

There exist rings for which (9) is not true. These are called non-commutative. However, in this thesis I will not study these rings. Therefore, from now on, when I write $T_{rings}$ I mean the theory of commutative rings.

**Definition 2.1.** A ring is a model of $T_{rings}$.

Secondly, a ring $F$ is called a field if it is a model of $T_{fields}$, which is equal to $T_{rings}$ together with two axioms that state that 0 is not equal to 1 and that every non-zero element of $F$ has an multiplicative inverse:

1. $\neg(1 = 0)$
2. $\forall x(\neg(x = 0) \rightarrow \exists y(x \cdot y = 1))$

A special kind of field are the algebraically closed fields. In such a field F, every polynomial with coefficients from F has a root in F. Therefore, the theory of algebraic closed fields, $T_{ACF}$, is given by $T_{fields}$, enriched with the axioms $\forall a_0 \cdots a_{n-1}\exists x(x^n + a_{n-1}x^{n-1} + \cdots + a_0 = 0)$ for every $n \in \mathbb{N}$. An example of an algebraically closed field is $\mathbb{C}$.

In addition to that, we have real closed fields. These fields have characteristic 0 and are ordered. Furthermore, every polynomial of odd order has a root. In mathematical notation, this means that $T_{RCF}$ is equal to $T_{fields}$ extended by:

1. $\forall xyz(y = x^2 \wedge y + z^2 = 0 \rightarrow y = 0)$
2. $\forall y(\exists x(y = x^2) \vee \exists z(y + z^2 = 0))$
3. $\forall xyz(x^2 + y^2 + z^2 = 0 \rightarrow x = y = z = 0)$
4. for every odd $n$ the sentence:$\forall a_0 \cdots a_{n-1} \exists x(x^n + a_{n-1}x^{n-1} + \cdots + a_0 = 0)$

Also, with regard to the remainder of this thesis it is useful to introduce the notion of an ideal. An ideal $I$ is a subset of a ring $R$ that has the following properties:

1. $0 \in I$
2. $\forall x, y \in I, x - y \in I$
3. $\forall r \in R, \forall x \in I, xr \in I$

$I$ is called a prime ideal if furthermore $\neg(I = R)$ and $\forall x, y \in I, xy \in I \implies x \in I$ or y $\in I$.

For ideals of polynomial rings, we have some interesting notions and theorems, that we will use in other sections of this thesis:

**Theorem 1** (Hilbert's Basis Theorem)**.** For a field $R$, let $R[X_1, \ldots, X_n]$ denote the ring of polynomials in $n$ variables with coefficients in $R$. Its ideals are finitely generated.

**Definition 2.2.** Let $R$ be a ring. For an ideal $I \subseteq R[X_1, \ldots, X_n]$, by $Z(I)$ we denote the zero set of $I$, i.e. the set of elements $x$ on which all functions $f$ of $I$ are zero. In addition, the ideal that consists of the functions $f$ that are zero on $Z(I)$, is denoted by $I(Z(I))$.

**Definition 2.3.** Let $R$ be a ring and $I \subseteq R$ an ideal. Then its radical $\sqrt{I}$ is defined to be the set: $\{a \in R | a^n \in I \text{ for some } n \in \mathbb{Z}^+\}$

**Theorem 2** (Nullstellensatz for algebraically closed fields)**.** Let $A$ be an algebraically closed field. Also, let $I$ be an ideal of $A[X_1, \ldots, X_n]$. Then $I(Z(I)) = \sqrt{I}$

Furthermore, we will make use of some other field-theoretical terminology:

**Definition 2.4.** Field-theoretically, an element $a$ is algebraic over a field $K$, if there exists a polynomial $p(x)$ in $K[X]$ such that $p(a) = 0$.

**Definition 2.5.** An extension of a field $K$ is called algebraic if all its elements are algebraic over $K$

**Definition 2.6.** An algebraic extension field of $K$ that is real closed, is called a real closure of $K$.

For the rational numbers $\mathbb{Q}$ we have for example that a real closure is the field of real numbers $\mathbb{R}$. In fact, with the ordering $<$ on $\mathbb{Q}$, we see that $\mathbb{R}$ is the only real closure, up to isomorphism:

**Theorem 3.** Every ordered field has a real closure. Moreover, it is unique up to isomorphism.

Lastly, for the proof of the Tarski-Seidenberg Theorem, we will consider the quotient field of a ring. Intuitively, this means:

**Definition 2.7.** The quotient field is the smallest field into which a ring can be embedded.

Its elements are the equivalence classes defined by for example the relation $R$: $\frac{a}{b} R \frac{c}{d} \Leftrightarrow ad = bc$. So for instance, the quotient field of the ring of natural numbers $\mathbb{N}$ is $\mathbb{Q}$.

# 3 Types

With regard to the subject of types, it is useful to make the following clear. In a language $L$, formulas can have free variables and bounded variables. There is an exact definition of what it means if a variable is free, but this definition involves a whole list of other definitions(see for example [4]). So I will explain what it means intuitively. For example, let the language $L$ contain a two-place function symbol $f$. Then consider the $L$-formula $\forall x f(x, y) = 0$. Here we have that the variable $x$ is bounded and that the variable $y$ is free. The idea behind it, is that the formula does not say anything about $x$. In fact, we could replace the $x's$ by $z$ and the formula would have the same meaning. However, for the variable $y$, this is different. The formula $\forall x f(x, y) = 0$ may be true if we substitute $y$ for one element, but may be false if we substitute it with another. So in a way, the formula states a property of $y$. A formula in which all variables are bound is called a sentence.

With this in mind, we consider the following definitions:

**Definition 3.1.** Let $L$ be a language and $M$ an $L$-structure. Then elements $a_1, \ldots, a_n \in M$ are said to satisfy an $L$-formula $\phi(x_1, \ldots, x_n)$ if $M \models \phi(a_1, \ldots, a_n)$.

Additionally:

**Definition 3.2.** Let $L$ be a language and $M$ an $L$-structure. We say that elements $a_1, \ldots, a_n$ of $M$ realise a set $q$ of $L$-formules in the free variables $x_1, \ldots, x_n$, if every $\phi(x_1, \ldots, x_n) \in q$, is satified by $a_1, \ldots, a_n$. This is sometimes denoted as $M \models q(a_1, \ldots, a_n)$.

And lastly:

**Definition 3.3.** Let $L$ be a language and $M$ an $L$-structure. A set of $L$-formulas $q$ is called finitely satisfiable if for every finite subset $r$ of $q$ we have elements $a_1, \ldots, a_n \in M$ such that $M \models r(a_1, \ldots, a_n)$.

Then, in the definition of an $n-type$, we have (amongst other criteria) that a set of formulas in $n$ free variables has to be finitely satisfied by an $n$-tuple of elements from $M$:

**Definition 3.4.** For a language $L$, let $M$ be a model of an $L$-theory $T$. An $n$-type over a model $M$ is a set $p$ of $L$-formulas in the free variables $x_1, \ldots, x_n$ that is maximal with respect to the property: for every finite subset $\phi_1, \ldots, \phi_k$ of $p$, there exist $a_1, \ldots, a_n \in M$ such that $M \models \phi_1(a_1, \ldots, a_n) \wedge \ldots \wedge \phi_k(a_1, \ldots, a_n)$.

As I will show later in this section, the space of n-types over a model $M$, written $S_n(M)$, has interesting properties if we endow it with a topology. For this purpose, first I define $[\phi]$ to be the set of types that contain the $L$-formula $\phi(x_1, \ldots, x_n)$. Then the collection of these sets is basis of a topology on $S_n(M)$. This follows from the facts that $[\phi_1 \vee \phi_2] = [\phi_1] \cup [\phi_2]$ ; $[\phi_1 \wedge \phi_2] = [\phi_1] \cap [\phi_2]$ ; $\emptyset = [\bot]$ ; $S_n(M) = [\top]$. Additionally, note that $[\phi]^c = [\neg\phi]$, so the complements

of these sets are also open. Now with the help of the following fundamental theorem, it can be shown that $S_n(M)$ is compact with this topology. But first define:

**Definition 3.5.** Let $L$ be a language. An $L$-theory $T$ is called consistent, if there is an $L$-structure $M$ such that all these sentences are true in $M$. $M$ is called a model of $T$. Moreover, let $T'$ be an set of the form $T \cup q$, in which $q$ denotes a set of $L$-formulas in the free variables $x_1, \ldots, x_n$. Then $T'$ is called consistent if there is a a model $M$ of $T$ such that there are $a_1, \ldots, a_n \in M$ such that $M \models q(a_1, \ldots, a_n)$.

**Theorem 4** (Compactness Theorem)**.** If every finite subset of a theory T is consistent, then so is T.

Before I show that $S_n(M)$ is compact, consider the following Lemma:

**Lemma 1.** Let $L$ be a language, $T$ an $L$-theory and $M$ a model of $T$. In addition, let $p$ be a set of $L$-formulas in the free variables $x_1, \ldots, x_n$. Then $p \in S_n(M)$ if and only if $p$ is a maximal set with respect to the property that $T \cup p$ is consistent.

*Proof.* Let $p$ be a set of $L$-formulas in $x_1, \ldots, x_n$ that is maximal with respect to the property that $T' = T \cup p$ is consistent. Because $T'$ consistent, there is a model $M$ of $T$ in which $a_1, \ldots, a_n$ can be found such that $M \models p(a_1, \ldots, a_n)$. Hence $p$ is finitely satisfied in $M$. Moreover, $p$ is maximal with respect to this last property, for suppose there is a set $q$ of $L$-formulas in the free variables $x_1, \ldots, x_n$, such that $q$ is finitely satisfied in $M$ and $p \subseteq q$, $p \neq q$. Then every finite subset of $T \cup q$ is consistent. So by the Compactness Theorem, $T \cup q$ must be consistent. However, this contradicts our assumption that $p$ was already maximal with respect to that property. Therefore $p$ is maximal finitely satisfiable in $M$, so $p \in S_n(M)$.

For the implication in the other direction, let $p \in S_n(M)$. Then $p$ is maximal finitely satisfiable in $M$. Since $M$ is a model of $T$, it then follows that every subset of $T \cup p$ is consistent. Therefore by the Compactness Theorem, $T \cup p$ is consistent. Furthermore, $p$ is maximal. For suppose not. Then there is a maximal set $q$ of $L$-formulas in $x_1, \ldots, x_n$ such that $T \cup q$ is consistent and $p \subseteq q$, $p \neq q$. But then it follows from the last paragraph that $q \in S_n(M)$. This contradicts the assumption that $p$ was already a type over $M$. We conclude from this that $p$ is a maximal set such that $T \cup p$ is consistent. □

To show compactness, suppose $\cup_{i \in I}[\phi_i] = S_n(M)$ for some index set $I$. Then $\{[\phi_i] | i \in I\}$ is an open covering of $S_n(M)$. Now note that $\emptyset = (\cup_{i \in I}[\phi_i])^c = \cap_{i \in I}([\phi_i])^c = \cap_{i \in I}[\neg\phi_i]$. Therefore, there is no type $p$ such that $\{\neg\phi_i | i \in I\} \subseteq p$. This means that $T \cup \{\neg\phi_i | i \in I\}$ is inconsistent. For suppose it would be consistent, then by Lemma 1, $\{\neg\phi_i | i \in I\}$ cannot be maximal with respect to this property. There would be a maximal set $p$ of $L$-formulas such that $T \cup p$ is consistent and $\{\neg\phi_i | i \in I\} \subseteq p$, $\{\neg\phi_i | i \in I\} \neq p$. By Lemma 1 $p$ must be a

type, contradiciton. So $T \cup \{\neg\phi_i | i \in I\}$ is inconsistent.

Now, because of Theorem 4, the finite subset $T_* \cup \{\neg\phi_i | i \in J\}$ is inconsistent as well, for some finite $T_* \subseteq T$ and finite $J \subseteq I$. It follows by Lemma 1 that there is no type q with $\{\neg\phi_i | i \in J\} \subseteq q$. Hence $\cap_{i\in J}[\neg\phi_i] = \emptyset$. Thus $S_n(T) = \emptyset^c = (\cap_{i\in J}[\neg\phi_i])^c = \cup_{i\in J}[\phi_i]$. Since J is finite, we conclude that $S_n(M)$ is compact.

Moreover, $S_n(M)$ is Hausdorff. To see this, let $p$ and $q$ be different types. Then there is an $L-$formula $\phi_1$ such that it is contained in only one of these types. Without loss of generality, assume $\phi_1 \in p$. Then $p \in [\phi_1]$. Since $q$ is a maximal subset, $\neg\phi_1 \in q$, so $q \in [\neg\phi_1]$. Note that $[\phi_1]$ and $[\neg\phi_1]$ are disjunct, for suppose the opposite is true. Then there is a type r such that $\phi_1 \in r$ and $\neg\phi_1 \in r$. Because r is a type over $M$, both $M \models \phi_1(a_1, \ldots, a_n)$ and $M \models \neg\phi_1(a_1, \ldots, a_n)$. This leads to a contradiction. So $S_n(M)$ is Hausdorff.

# 4 The space of types over Algebraically Closed Fields

In a way, types are related to the ideals of a field. There is a bijection between the space of $n$-types over an agebraically closed field $K$ and the space of prime ideals of $K[X_1, \ldots, X_n]$. Moreover this bijection is continuous in one direction, as we will see in this section. But first define:

**Definition 4.1.** By $Spec(R)$ we denote the set of prime ideals of a commutative ring $R$.

We endow this space by a topology such that sets of the form:
$V_I = \{P \in Spec(R) | I \subseteq P\}$ are closed, where $I$ is an ideal of $R$.

Let $K$ be an algebraically closed field. First, let us look at what the elements of the prime ideals in $Spec(K[X_1, \ldots, X_n])$ are. As a result of the Nullstellensatz, we have that $I(Z(P)) = \sqrt{P}$ for a prime ideal $P \in Spec(K[X_1, \ldots, X_n])$. Moreover, note that $\sqrt{P} = P$ because of the properties of a prime ideal. So $I(Z(P)) = P$. This means that the elements of $P$ are polynomials that have at least one common zero. It also means that if one such common zero is $a$ and $f(a) = 0$ for a polynomial $f$, we have that $f \in P$.

With this in mind, we define the function $h : S_n(K) \to Spec(K[X_1, \ldots, X_n])$ in the following way: to each type $p$, we assign the ideal $I$ that consists of the functions $f$ for which the $L_R$-formula $f(v_1, \ldots, v_n) = 0$ is contained in $p$.

It can be shown that this map is bijective. For suppose we have an ideal $I \in Spec(K[X_1, \ldots, X_n])$. Then there are $v_1, \ldots, v_n$ such that $f(v_1, \ldots, v_n) = 0$ for every $f \in I$. Then consider the type $p$ that consists of consequences of the formulas "$f(v_1, \ldots, v_n) = 0$" for all the $f \in I$ and "$\neg(f(v_1, \ldots, v_n) = 0$" for all the $f \notin I$. This is a type because firstly, every finite subset is obviously satisfied. Secondly, it is maximal because it contains the consequences of all satisfiable atomic formulas and satisfiable negations of atomic formulas. Since $T_{ACF}$ has quantifier elimination, this means that all satisfiable formulas are in $p$. This will become clear in later sections. For now, note that $h(p) = I$. This means that $h$ is surjective. Moreover, $p$ is uniquely determined from $I$. For let there be another type $q$ that contains all formulas from $p$. Then because types are maximal, $q$ must be equal to $p$. So $h$ is injective.

Lastly, assume we have a closed subset $V$ of $Spec(K[X_1, \ldots, X_n])$. Then each ideal $P$ in this subset determines a type $p = h^{-1}(P)$. Then for an $f \in P$, the formula $\phi_P := f(v_1, \ldots, v_n) = 0$ is an element of $p$. Therefore, $p \in [\phi_P]$. This means that $h^{-1}(V)$ can be written as the union $\bigcup_{P \in V} [\phi_P]$, which is closed. This leads to the conclusion that $h$ is continuous.

The results above have some interesting corollaries. For example, we see now

that $Spec(K[X_1, \ldots, X_n])$ is also compact, which can be of importance in the study of algebraic geometry. Furthermore, the fact that
$|S_n(K)| = |Spec(K[X_1, \ldots, X_n])|$ will be useful in the next chapter.

# 5 Properties of $T_{ACF}$ and $T_{RCF}$

## 5.1 $T_{ACF}$

**Definition 5.1.** Let $L$ be a language. An $L$-theory $T$ is said to admit quantifier elimination, if for every $L$-formula $\phi$ in the free variables $x_1, \ldots, x_n$, there exists a quantifier free $L$-formula $\psi$ in at most the free variables $x_1, \ldots, x_n$ that is equivalent to $\phi$, modulo $T$. Equivalent modulo $T$ means that
$T \models \forall x_1, \ldots, x_n(\phi \leftrightarrow \psi)$ holds.

**Theorem 5** (Tarski)**.** The theory of algebraically closed fields admits quantifier elimination.

This theorem has important implications. But first we define:

**Definition 5.2.** Let $L$ be a language. An $L$-theory $T$ is called model complete if for every two models of $T$, with the property that one model is a substructure of the other, the former model is an elementary substructure of the latter.

**Definition 5.3.** Let $L$ be a language. An $L$-theory $T$ is called complete if for every $L$-sentence $\psi$, we have that either $T \models \psi$ or $T \models \neg\psi$.

Equivalently, we have:

**Proposition 1.** Let $L$ be a language and $T$ an $L$-theory. Then $T$ is complete if in all models the same $L$-sentences are true.

As a result of quantifier elimination, $T_{ACF}$ is model complete. For suppose we have two models $M_1$ and $M_2$ such that $M_1$ is a substructure of $M_2$. $M_1$ is an elementary substructure of $M_2$, if for all $L_R$-formulas $\phi$ in the free variables $x_1, \ldots, x_n$ we have the following: for elements $a_1, \ldots, a_n$ of $M_1$,
$M_1 \models \phi(a_1, \ldots, a_n) \iff M_2 \models \phi(a_1, \ldots, a_n)$. Since quantifier free formulas do not change meaning under substructure, this is true for quantifier free formulas. Now as a result of quantifier elimination, it then holds for all formulas. So $M_1$ is an elementary substructure of $M_2$.

However, $T_{ACF}$ is not complete. Namely, if one takes a pair of models, they do not necessarily satisfy the same $L_R$-sentences. To see this, consider the characteristic of fields. The characteristic is the number of times one has to add 1 to get 0. So if the characteristic of a field is three, the $L_R$-sentence $1 + 1 + 1 = 0$ is true for this field. The characteristic is not the same for every algebraically closed field, since it is not determined by $T_{ACF}$. Therefore, for a pair of models there may be an $L_R$-sentence that is true in only one of them.

Nevertheless, we can define theories that determine the characteristic:

**Definition 5.4.** By $T^0_{ACF}$ and $T^p_{ACF}$, we denote the theories that consist of $T_{ACF}$ with infinitely many extra $L_R$-sentences that together say that the characteristic is 0 (not 1, not 2, etc.) or one extra $L_R$-sentence that states that the characteristic is a prime number p, respectively.

We can prove that these theories are complete by analysing the $L$-sentences. Shortly saying, for a language $L$ every $L$-sentence is "built" from other sentences, called atomic sentences. As a result, properties of the latter sentences give us information about all sentences. More formally we write:

**Definition 5.5.** An atomic formula is an L-sentence of the form $s_1 = s_2$, $R(s_1, \ldots, s_n)$ or $\bot$.

We will demonstrate that $T^p_{ACF}$ is complete. We want to show that for every $L_R$-sentence $\psi$ either $T^p_{ACF} \models \psi$ or $T^p_{ACF} \models \neg\psi$. Since $T_{ACF}$ has quantifier elimination, $T^p_{ACF}$ has. As a result, we only have to look at the quantifier free $L_R$-sentences to determine whether $T^p_{ACF}$ is complete. These quantifier free sentences, are conjunctions, disjunctions, implications and negations of atomic sentences. Therefore, it is sufficient to look only at the atomic sentences.

The atomic sentences are of the form $s = t$ in which the $s$ and $t$ are closed terms. Since these terms are closed and we only have the constants 1 and 0 in $L_R$, $s$ and $t$ must be elements of $\mathbb{Z}$. Then as a result of the characteristic $p$, $T^p_{ACF} \models (t = s)$ if and only if the difference between $t$ and $s$ is a multiple of $p$. And $T^p_{ACF} \models \neg(t = s)$ if and only if the difference between $s$ and $t$ is not a multiple of $p$. It follows that $T^p_{ACF}$ is complete. In a similar way we have that $T^0_{ACF}$ is complete.

For complete theories we have the following notion:

**Definition 5.6.** If for any model $M$ of a complete theory $T$, $|S_n(M)| = |M|$, then $T$ is called $\omega$-stable.

It easily follows that the complete theories $T_{ACF}$ with characteristic 0 or prime $p$ , $T^0_{ACF}$ or $T^p_{ACF}$, are $\omega$-stable. On the one hand because of the fact that by Hilbert's basis theorem, for an algebraically closed field $K$ the ideals of $K[X_1, \ldots, X_n]$ are finitely generated and therefore $|Spec(K[X_1, \ldots, X_n])| = |K[X_1, \ldots, X_n]|$. And since polynomials are finitely generated as well, from the coefficients of $K$, $|K| = |K[X_1, \ldots, X_n]| = |Spec(K[X_1, \ldots, X_n])|$ On the other hand because, as we saw, $|S_n(K)| = |Spec(K[X_1, \ldots, X_n])|$, so $|S_n(K)| = |K|$.

Finally, we have the concept of a strongly minimal theory. We first define:

**Definition 5.7.** Let $L$ be a language. A subset $A$ of a set $B$ is definable, if there exists an $L$-formula $\phi$, such that $A = \{x \in B | B \models \phi(x)\}$.

A strongly minimal theory then, is what we call a theory with the property that for every definable subset $A$ of a model $B$, either $A$ is finite or $B - A$ is finite.

**Proposition 2.** In $L_R$, $T_{ACF}$ is strongly minimal.

*Proof.* Let $K$ be an algebraically field and $\phi(y)$ an $L_R$-formula that defines a subset $A$ of $K$. Note that as a result of quantifier elimination, $\phi(y)$ must have a quantifier free equivalent that is the finite disjunction of finite conjunctions of atomic formulas. These formulas are of the form $f(y) = 0$ or $\neg(f(y) = 0)$, since the $L_R$-terms with free variables must be polynomials(see also section 6). The formula $f(y) = 0$ defines a finite subset, since a polynomial has a finite number of roots. For the same reason, the set defined by $\neg(f(y) = 0)$ is cofinite. So if all conjunctions have a formula of the form $f(y) = 0$ in it, $A$ must be finite. If not, there is a conjunction that exists only of formulas of the form $\neg(f(y) = 0)$, in which case $A$ is cofinite. □

## 5.2 $T_{RCF}$

First of all, $T_{RCF}$ is not $\omega$-stable. For let $R$ be the field of real algebraic numbers over $\mathbb{Q}$. This field is a model of $T_{RCF}$. The elements of this field are the zeros of polynomials with coefficients in $\mathbb{Q}$. Since $\mathbb{Q}$ is countable, there are countably many of such polynomials. Moreover, a polynomial has finitely many zeros. As a result, $R$ must be countable.
However, $S_n(R)$ is uncountable. Namely, for a real number $r \in \mathbb{R}$, consider the type that contains: $A := \{q < x < p | q, p \in \mathbb{Q} : q < r < p\}$(Note that where I write $a < b$ in $L_R$-formulas, I mean that $(\exists z(z \neq 0 \wedge a + z^2 = b))$; $L_R$ does not have a relation symbol $<$). This type exists, because $A$ is finitely satisfiable. Moreover, for every real number $r$ we have a different type of this kind. The set of real numbers $\mathbb{R}$ is uncountable, so there are uncountably many of such types over $R$. Therefore, $S_n(R)$ is uncountable. So we see that $|S_n(R)| > |R|$

As opposed to $T_{ACF}$, $T_{RCF}$ does not have quantifier elimination in $L_R$. For assume by contradiction that it does. Let $\psi(x) := \exists z(z \neq 0 \wedge x + z^2 = 2)$. The formula $\psi(x)$ must have a quantifier free equivalent $\phi(x)$. Every quantifier free $L$-formula defines a finite or cofinite subset(this follows from the fact that these formulas are a finite combination of atomic formulas and disjunctions, conjunctions, negations and implications). So $\phi(x)$ must define a subset that is finite or cofinite. The formula $\psi(x)$ defines a subset that is neither finite nor cofinite, since in for example $\mathbb{R}$, there are infinitely many numbers bigger than 2 and there are infinitely many smaller. But $\phi(x)$ and $\psi(x)$ are equivalent, modulo $T_{ACF}$. Contradiction.

So in $L_R$, $T_{RCF}$ does not have quantifier elimination. In addition, the theory is model complete, but not complete. But that is not all there is to it. It depends on the language one uses. One can make a theory have quantifier elimination by admitting an extra relation symbol $R$ to the language $L_R$ for every $L_R$-sentence $\phi$ and expanding the theory by axioms saying: $\forall x_1, \ldots, x_n(R(x_1, \ldots, x_n) \leftrightarrow \phi(x_1, \ldots, x_n))$. This is called the Morleyisation of T. This theory is still not complete however.

Also, one can consider the theory of real closed fields when it is axiomated in the language of ordered rings, $L_{OR}$. This language has a two-place relation symbol $<$ for order. For this case, we have an important result:

**Theorem 6** (Tarski-Seidenberg)**.** The $L_{OR}$-theory $T_{RCF}$ has quantifier elimination.

In the next section I will give a proof of this theorem and background information. Furthermore, I will show in section 7 that with the language $L_{OR}$, $T_{RCF}$ is model complete and complete.

For theories like $T_{RCF}$, with a dense linear order $<$ without end-points, there is a similar notion to that of strong minimality for $T_{ACF}$, after we have defined what an interval is:

**Definition 5.8.** Let $L$ be a language with a relation symbol $<$ for order. Let $T$ be an $L$-theory that contains $L$-sentences that state that $<$ is a dense lineair order without end-points. Furthermore, let $M$ be a model of $T$. An interval in $M$ is a subset $I$ of $M$ such that for $a, b \in I$, $a \leq c \leq b \rightarrow c \in I$. An interval $I$ is called open, if there exist $q \in M$ and $r \in M$, such that $I$ consists of all elements $s \in M$ that satisfy $q < s < r$.

**Definition 5.9.** Let $L$ be a language with a relation symbol $<$ for order. An $L$-theory that contains $L$-sentences that state that $<$ is a dense lineair order without end-points, is o-minimal if in all models, every definable set can be written as the finite union of points and open intervals.

**Proposition 3.** The $L_{OR}$-theory $T_{RCF}$ is o-minimal.

*Proof.* Let $J$ be a real closed field an $B \subseteq J$ defined by $\psi(x)$.$T_{RCF}$ has quantifier elimination. Then in the same way as we did for a definable subset of an algebraically closed field, we can write $\psi$ in the form:
$\psi(x) = \vee_{i=1}^{n}(\wedge_{j=1}^{m} f_{ij}(x) = 0 \wedge \wedge_{k=1}^{p} g_{ik}(x) > 0)$. Formulas of the form $f_{ij}(x) = 0$ define a finite number of points, since the $f_{ij}$ are polynomials. Formulas of the form $g_{ik}(x) > 0$ define a finite number of open intervals, because of the fact that polynomials are continuous. So $\psi$ defines a finite union of finite intersections of intervals and points. As a result, $T_{RCF}$ is o-minimal. □

# 6 Tarski-Seidenberg

In 1940 Tarski proved that the theory of real closed fields admits quantifier elimination. Actually, he proved it twice. The second time he gave a proof in "A decision method for elementary algebra and geometry" in 1948. It was Seidenberg however who made the theorem accessible to more mathematicians, when he gave a different proof in 1954. Therefore, the theorem is widely known as the Tarski-Seidenberg theorem. However, the name Tarski's theorem is also commonly found. To complicate things even more, an important consequence of this theorem is regularly called the Tarski-Seidenberg Theorem as well. Namely the fact that a projection of a semi-algebraic set is also semi-algebraic, which will become clear in the next section. For now, we will concentrate on quantifier elimination. First consider the following definitions and lemma:

**Definition 6.1.** A basic formula is an atomic formula or its negation.

and:

**Definition 6.2.** A formula of the form $\phi = \exists x \beta(x)$ where $\beta$ is a quantifier free formula that is a conjunction of basic formulas, is named a primitive existential formula.

It is known that a theory has quantifier elimination if every primitive existential formula $\exists x \beta(x)$ is equivalent to a quantifier free formula. In fact, for a real closed field $R$ in de language $L_{OR}$, it is sufficient to prove that sentences of the form: $\exists x(\wedge_{i=1}^{m} f_i(x) = 0 \wedge_{j=1}^{k} g_j(x) > 0)$, with the $f_i$ and $g_j$ polynomials, are equivalent to a quantifier free one.
To see this, first note that $\beta(x)$ is of the form $\wedge_{q=1}^{s} \gamma_q(x) \wedge_{r=1}^{t} \neg\delta_r(x)$, with $\gamma$ and $\delta$ atomic formulas. Since we only have the two-place relationsymbol $<$ in $L_{OR}$, $\gamma_q$ and $\delta_r$ are either of the form $a = b$ or $a < b$. So the negations of the $\delta_r$ are either $\neg(a = b) \leftrightarrow (a < b \vee b < a)$ or $\neg(a < b) = (a = b \vee b < a)$, in which the $a$ and $b$ represent $L_{OR}$-terms. Note that actually, these negations are disjunctions of atomic formulas. As a result, $\beta$ can be written as: $\vee_{e=1}^{g} \chi_e$ in which the $\chi_e$ are conjuntions of atomic formulas.
It is easy to see that $\exists x \beta(x) \leftrightarrow \exists x \vee_{e=1}^{g} \chi_e(x) \leftrightarrow \vee_{e=1}^{g} (\exists x \chi_e(x))$. Now, as the $\chi_e(x)$ are atomic, they are of the form $a(x) = b(x) \leftrightarrow a(x) - b(x) = 0$ or $a(x) < b(x) \leftrightarrow b(x) - a(x) > 0$. In $L_{OR}$, the terms $a(x)$ and $b(x)$ must be polynomials. Therefore, we only have to prove that sentences of the form $\exists x(\wedge_{i=1}^{m} f_i(x) = 0 \wedge_{j=1}^{k} g_j(x) > 0)$ have a quantifier free equivalent, in order to show quantifier elimination.

We can show that a such formula has a quantifier-free equivalent, with the use of the following Lemma:

**Lemma 2.** Let $L$ be a language. For an $L$-theory $T$, a formula $\phi(x_1, \ldots, x_n)$ is equivalent to a quantifier free formula, if for every two models $M_1$ and $M_2$ that have a common substructure $R$, we have: $M_1 \models \phi(a_1, \ldots, a_n) \implies M_2 \models \phi(a_1, \ldots, a_n)$ for all $a_1, \ldots, a_n \in R$.

*Proof.* Let $T$ be an $L$-theory. Assume that we have a formula $\phi(x_1, \ldots, x_n)$ such that for any two models $M_1$ and $M_1$ with a common substructure $R$, $M_1 \models \phi(a_1, \ldots, a_n) \implies M_2 \models \phi(a_1, \ldots, a_n)$ holds for all $a_1, \ldots, a_n \in R$.

To start with, note that if $\phi(x)$ or $\neg\phi(x)$ is inconsistent with $T$, it is already clear that $\phi(x)$ must be equivalent to a quantifier free formula. Because, if $T \vdash \forall x \neg\phi(x)$, it follows that $T \vdash \forall x(\phi(x) \leftrightarrow \neg(c = c))$ for a constant $c$. In a similar way, if $T \vdash \forall x \phi(x)$, $T \vdash \forall x(\phi(x) \leftrightarrow (c = c))$.

From now on, assume $\phi(x)$ and $\neg\phi(x)$ are consistent with $T$. Let $\Gamma(x)$ be the set of quantifier free formulas $\psi(x)$ that follow from $\phi(x)$, modulo $T$. For new constant symbols $b_1, \ldots, b_n$, denote $(b_1, \ldots, b_n) = b$, we will prove the claim that $T \cup \Gamma(b) \vdash \phi(b)$, from which it will follow that $\phi(x)$ is equivalent to a quantifier free formula. But first we will prove our claim:

Suppose by contradiction that not $T \cup \Gamma(b) \vdash \phi(b)$. Then $T \cup \Gamma(b) \cup \{\neg\phi(b)\}$ is consistent and so there is a model $M_1$.

Consider the substructure $A$ of $M_1$ that is generated by $b$ and also consider $\Delta = T \cup Diag(A) \cup \{\phi(b)\}$, where $Diag(A)$ is the set of basic formulas for elements from $A$, that are true in $A$. If $\Delta$ by contradiction is not consistent, $Diag(A) \cup \{\phi(b)\}$ must be inconsistent, since $\phi(x)$ is consistent with $T$. It follows from the compactness theorem then, that there are $\chi_1(b), \ldots, \chi_k(b) \in Diag(A)$, such that $\{\chi_1(b), \ldots, \chi_k(b)\} \cup \{\phi(b)\}$ is inconsistent. So $\wedge_{i=1}^{k} \chi_i(b)$ implies $\neg\phi(b)$, modulo $T$. Now because $b_1, \ldots, b_n$ were newly added, they do not appear in $T$. Therefore, this means that for all $x$, $\wedge_{i=1}^{k} \chi_i(x)$ implies $\neg\phi(x)$. But then there is a $\chi(x)_i$ such that modulo $T$, for all $x$, $\phi(x) \rightarrow \neg\chi_i(x)$. As a result, $\neg\chi_i(x) \in \Gamma$. Because $\neg\chi_i(x)$ is quantifier free and $A$ a substructure of $M_1$, also $A \models \neg\chi_i(b)$. However, this contradicts the fact that $\chi_i(b) \in Diag(A)$. So $\Delta$ is consistent.

Since $\Delta$ is consistent and $Diag(A) \subset \Delta$, there is a model $M_2$ of $\Delta$ such that $A$ is a substructure of $M_2$. Now since both $M_1$ and $M_2$ are models of $T$ with a common substructure $A$, it follows with our assumption that $M_2 \models \neg\phi(b)$, because $M_1 \models \neg\phi(b)$. This is a contradiction with the fact that $M_2$ is a model of $\Delta$, so $T \cup \Gamma(b) \vdash \phi(b)$.

Now since $T \cup \Gamma(b) \vdash \phi(b)$, we have that $T \cup \Gamma(b) \cup \neg\phi(b)$ is inconsistent. The formula $\neg\phi(x)$ is consistent with $T$, so in particular, $\Gamma(b) \cup \phi(b)$ is inconsistent. By the compactness theorem and because $A$ is generated by $b$, there must be $\psi_1(b), \ldots, \psi_n(b) \in \Gamma(b)$ that are inconsistent with $\neg\phi(b)$. As a result, $\wedge_{i=1}^{n} \psi(b) \rightarrow \phi(b)$, modulo $T$. Because of the $b_1, \ldots, b_n$ being new constants, for all $x$ $\wedge_{i=1}^{n} \psi(x) \rightarrow \phi(x)$, modulo $T$. Also note that all $\psi_i(x)$ already follow from $\phi(x)$. Therefore $\wedge_{i=1}^{n} \psi(x)$ and $\phi(x)$ are equivalent. Since all $\psi_i(x)$ are quantifier free, $\phi(x)$ is equivalent to a quantifier free formula, modulo $T$. □

Now we can prove quantifier elimination for real closed fields:

**Theorem 7** (Tarski-Seidenberg)**.** The $L_{OR}$-theory $T_{RCF}$ has quantifier elimination.

*Proof.* Let $M_1$ and $M_2$ be two models of $T_{RCF}$ with a common substructure $R$. Let furthermore $\psi(a_1, \ldots, a_n) := \exists b \phi(b, a_1, \ldots, a_n)$ with $a_1, \ldots, a_n \in R$ be a primitive existential. As already pointed out, we need to prove that $\chi(a_1, \ldots, a_n) := \exists b(\wedge_{i=1}^{m} f_i(b) = 0 \wedge_{j=1}^{k} g_j(b) > 0)$ is equivalent to a quantifier free formula. As a consequence of lemma 2, we have to show that $M_2 \models \chi(a_1, \ldots, a_n)$ if $M_1 \models \chi(a_1, \ldots, a_n)$.

Assume $M_1 \models \chi(a_1, \ldots, a_n)$ for $a_1, \ldots, a_n \in R$. As shown before, this means there are functions $f_i$ and $g_j$ and a $b$ in $M_1$ such that $\wedge_{i=1}^{m}(f_i(b) = 0) \wedge \wedge_{j=1}^{k}(g_j(b) > 0)$ holds. Note that because $R$ is a substructure of $M_1$ and $M_2$, these functions have the same interpretation in $R$ and $M_2$ as in $M_1$.

Firstly, if one of the $f_i(x)$, say $f_h(x)$, is not zero for all $x$ in $M_2$, $b$ is a root of $f_l$. Then $b$ is algebraic over $R$ and its quotient field as well, so it is contained in the real closure $C$ of the latter. Note that by Theorem 3, this real closure is unique because the quotient field is ordered. This means $C \subseteq M_1$ and $C \subseteq M_2$. As a consequence of the fact that $C \subseteq M_2$, also $b \in M_2$ and therefore $M_2 \models \exists b \wedge_{i=1}^{m} (f_i(b) = 0) \wedge \wedge_{j=1}^{k}(g_j(b) > 0)$. So $\chi(a_1, \ldots, a_n)$ and therefore $\psi(a_1, \ldots, a_n)$ must be equivalent to a quantifier free formula.

Now assume that all $f_i(x)$ are zero for all $x \in M_2$. Then $\chi(a_1, \ldots, a_n)$ reduces to $\exists b \wedge_{j=1}^{k} (g_j(b) > 0)$. Since $C$ is real closed, polynomials of uneven order have roots in $C$. So these $g_j(x)$ can be factorized in factors of the form $(x - q)$ and $(x^2 + rx + s)$, where the latter factor remains either positive or negative for all $x$. It depends on the former kind of factors whether $g_j(x) > 0$ or not. Note that the $q$ in these factors are a root of the polynomial in $C$. Since $M_1 \models \exists b \wedge_{j=1}^{k} (g_j(b) > 0)$, we can find $\alpha$ and $\beta$ in terms of these roots, such that $\wedge_{j=1}^{k}(g_j(c) > 0)$ for all $\alpha < c < \beta$. Because $\alpha < \frac{\alpha+\beta}{2} < \beta$ and $\frac{\alpha+\beta}{2} \in C \subseteq M_2$, this means $M_2 \models \chi(a_1, \ldots, a_n)$. □

# 7 Applications of quantifier elimination for real closed fields

As I already mentioned, there is a well-known corollary of quantifier elimination in real algebraic geometry, which is known as the Tarski-Seidenberg theorem as well. Furthermore, we will see that $T_{RCF}$ is complete. In addition to this, there are several theorems in real algebraic geometry that were hard to prove geometrically, but are less complicated now we can use quantifier elimination. For instance, we can give a proof of Hilbert's 17th problem and of a nullstellensatz for real closed fields. These applications will be discussed in this section and in section 8, respectively.

In order to prove the first corollary define:

**Definition 7.1.** Let $L$ be a language and $T$ an $L$-theory. Then an $L$-structure that is a substructure of all models of $T$ is called a prime structure.

**Corollary 1.** The $L_{OR}$-theory $T_{RCF}$ is complete.

*Proof.* As a consequence of quantifier elimination, all $L_{OR}$-sentences are equivalent to a quantifier free $L_{OR}$-sentence, modulo $T_{RCF}$. Since in a substructure the same quantifier free sentences are true as in the structure of which it is a substructure, we see that in $T_{RCF}$, a substructure and a structure satisfy the same $L_{OR}$-sentences. Moreover, it can be shown that the set of rational numbers $\mathbb{Q}$ (up to isomorphism) together with the ordering $<$ is a prime structure for this theory. Hence all models of $T_{RCF}$ satisfy the same $L_{OR}$-sentences. So $T_{RCF}$ is complete. □

We also see that, just like $T_{ACF}$, $T_{RCF}$ is model complete, as a direct consequence of quantifier elimination. Additionally, we have a similar notion for real closed fields to that of what are called constructible sets for algebraic closed fields, which we call semi-algebraic sets.

**Definition 7.2.** A subset A of $R^n$, where $R$ is real closed, is called semi-algebraic, if it can be written as a combination of sets of the form $\{x|f(x) > 0\}$ for polynomials $f$ with coefficients in $R$.

We can make a so-called projection of these subsets. A projection of a set $A \subseteq R^{n+1}$ is a set $\Pi(A) \subseteq R^n$, such that for every $n$-tuple $y$ in $\Pi(A)$, there is an $x \in R$ such that the $(n+1)$-tuple $(x, y)$ is an element of $A$. In case $n = 2$, you could compare it with the shadow that a cube casts on the wall.

**Corollary 2.** The projection of a semi-algebraic set is again a semi-algebraic set.

*Proof.* First of all, note that every semi-algebraic set is defined by a quantifier free formula and that every quantifier free formula defines a semi-algebraic set. This is a fact, because the atomic formulas define sets of the form $\{x|0 < f(x)\}$

and $\{x|0 = f(x)\}$. Then the conjunctions, disjunctions and negations of the atomic formulas respectively translate into intersections, unions and complements of these sets. So the quantifier free formulas define the semi-algebraic sets.
Assume we have such a semi-algebraic subset $A \subseteq R^n$ that is defined by the quantifier free formula $\phi(x)$, in which the $x$ stands for the $n$-tuple $x_1, \ldots, x_n$. Then the formula $\exists x_1 \phi(x)$ defines the projection of $A$. As a consequence of quantifier elimination, there is a quantifier free equivalent $\psi(x_2, \ldots, x_n)$ of this formula. Since this $\psi$ is quantifier free, it defines a semi-algebraic subset of $R^{n-1}$. Because of the equivalence, this must be exactly the projection. □

The next theorem, Hilbert's 17th problem, is about rational, positive semi-definite functions $f$. This means that $f$ can be written as a fraction of polynomials in $R[X_1, \ldots, X_n]$, were $R$ is a field, and that $f(x_1, \ldots, x_n) \geq 0$ for all $x_1, \ldots, x_n$ in $R$, respectively. Abraham Robinson was the first to prove this result using quantifier elimination, after Emil Artin had proved it with the help of earlier techniques.

**Theorem 8** (Hilbert's 17th problem)**.** Let $R$ be a real closed field. Then every rational positive semi-definite function $f \in R(X_1, \ldots, X_n)$ can be written as $f = g_1^2 + \ldots + g_k^2$, where $g_i \in R(X_1, \ldots, X_n)$ for $i \in \{1, \ldots, k\}$.

*Proof.* Assume that $f$ is positive semi-definite and by contradiction that $f$ cannot be written as a sum of squares. Then there is an algebraic lemma (see [3], 2.10) that says that there is a possible ordering $>$ of the quotient field $R(X_1, \ldots, X_n)$ such that there is an element $a := (a_1, \ldots, a_n)$ for which $f(x_1, \ldots, x_n) < 0$. Therefore, the $L_{OR}$-sentence $\exists a(f(a) < 0)$ is true in $(R(X_1, \ldots, X_n), <)$. Since every ordered field has a unique real closure by Theorem 3, $(R(X_1, \ldots, X_n), <)$ has one. Call it $A$.Then $\exists a(f(a) < 0)$ holds in $A$. As a result of completeness, this sentence (or at least a quantifier-free equivalent) is true in $R$. However, $f$ is positive semi-definite. We reach a contradiction and therefore we conclude that $f$ is a sum of squares. □

# 8 Real Nullstellensatz

The Real Nullstellensatz is another theorem that can be proven with the use of quantifier elimination, instead of using geometrical techniques only. In order to prove this result, we consider the following lemma's and definitions from algebraic geometry. I will not prove these lemma's in this thesis, but for a proof, see for example [5].

**Definition 8.1.** Let $R$ be a ring. An ideal $I \subseteq R$ is called real if for every sum of squares $\sum_{i=1}^{n} a_i^2 \in I$, $a_1, \ldots, a_n \in I$.

**Lemma 3.** Let $R$ be a real closed field. Let I be a real ideal of $R[X_1, \ldots, X_n]$. Then $I$ can be written as the intersection of prime ideals P. Moreover, all of these prime ideals P are real.

**Definition 8.2.** Let $R$ be a real closed field and $I \subseteq R[X_1, \ldots, X_n]$ an ideal. By $R[X_1, \ldots, X_n]/I$ we denote the quotient ring of which the elements are equivalence classes determined by the following relation: $aRb \iff a - b \in I$ for $a, b \in R[X_1, \ldots, X_n]$.

Note that with definition 8.2, the elements of $I$ form one equivalence class. This class can be seen as the zero-element of the ring $R[X_1, \ldots, X_n]/I$. With regard to this quotient ring, we need one other lemma:

**Lemma 4.** Let R be a real closed field. If $P$ is a real prime ideal of $R[X_1, \ldots, X_n]$, then the quotient ring $R[X_1, \ldots, X_n]/P$ is a field and has an ordering.

**Theorem 9** (Real Nullstellensatz)**.** Let $R$ be a real closed field. If an ideal $I$ of $R[X_1, \ldots, X_n]$ is real, $I(Z(I)) = I$.

*Proof.* It is obvious that $I \subseteq I(Z(I))$. So assume $f \in I(Z(I))$. We will show that $f \in I$. Since by Lemma 3 we have that $I = \cap_{i \in L} P_i$ for some index set $L$ and real prime ideals $P_i$, it will be sufficient to show that $f \in P_i$ for all $i \in L$.

Consider one such $P_i$, say $P_k$. We will prove that $f \in P_k$. To start with, note that by Hilbert's Basis Theorem, $P_k$ is finitely generated. This means that there are polynomials $g_1, \ldots, g_n$ for some $n \in \mathbb{N}$, such that $P_k = \langle g_1, \ldots, g_n \rangle$. Suppose we have an $x \in R^n$, such that $g_1(x) = \ldots = g_n(x) = 0$. It then follows that $x \in Z(I)$ because one of the elements of $P_k$ must be in $I = \cap_{i \in L} P_i$. Otherwise, $I = \emptyset$ and the result would follow immediately. Now since $x \in Z(I)$, we must have that $f(x) = 0$. Hence we see that:

$$R \models \forall x (g_1(x) = \ldots = g_n(x) = 0 \rightarrow f(x) = 0) \tag{1}$$

Now consider the real closed extension $R'$ of $R[X_1, \ldots, X_n]/P_k$. By Lemma 4 and Theorem 3, this $R'$ exists. As a result of the completeness of $T_{RCF}$(which follows from quantifier elimination), we see that also:

$$R' \models \forall x (g_1(x) = \ldots = g_n(x) = 0 \rightarrow f(x) = 0) \tag{2}$$

Note that there are elements $a$ in $R'^n$ of the form $(X_1/P_K, \ldots, X_n/P_k)$. In particular, we have that the image of these elements by polynomials $h$ over $R'$ can be viewed of as equivalence classes of the form
$h(a) = h(X_1/P_K, \ldots, X_n/P_k) = h(X_1, \ldots, X_n)/P_k$. Since $g_1, \ldots, g_n \in P_k$, we find that $\wedge_{j=1}^n g_j(X_1, \ldots, X_n)/P_k = 0$ and so $\wedge_{j=1}^n g_j(X_1/P_K, \ldots, X_n/P_k) = 0$. Then it follows by (2) that also
$f(X_1/P_k, \ldots, X_n/P_k) = 0$, which means that $f(X_1, \ldots, X_n)/P_k = 0$. Therefore, we can conclude that $f \in P_k$.

It is clear that the same goes for all $P_i, i \in L$. As a result $f \in I$ and therefore $I(V(I)) \subset I$. Together with the fact that $I \subseteq I(V(I))$ this leads to the conclusion that $I(V(I)) = I$. $\square$

# 9 Semi-algebraic sets and functions

As we saw in corollary 2, there is a connection between semi-algebraic subsets and definable subsets of real closed fields. As a matter of fact, they are the same. This fact enables us to reformulate definitions and results from real algebraic geometry. For example, we have in real algebraic geometry:

**Definition 9.1.** Let $R$ be a real closed field and let $X \subseteq R^m$ and $Y \subseteq R^n$ be semi-algebraic subsets. A semi-algebraic function is a function $f : X \to Y$, such that its graph $G = \{(x, y) \in X \times Y | f(x) = y\} \subseteq R^{m+n}$ is semi-algebraic.

This becomes:

**Definition 9.2.** Let $L$ be a language. Let $R$ be a real closed field and $X \subseteq R^m$ and $Y \subseteq R^n$ be definable subsets. A semi-algebraic function is a function $f : X \to Y$, such that there is an L-formula $\phi(x, y)$ that defines its graph $G = \{(x, y) \in X \times Y | f(x) = y\} \subseteq R^{m+n}$.

In other words, the graph of f has to be a definable set. That is why a semi-algebraic function is called a definable function, from a model-theoretic point of view. With this in mind, we can prove some interesting properties of definable functions. For example:

**Proposition 4.** Let $R$ be a real closed field and let $X \subseteq R^m$ and $Y \subseteq R^n$ be semi-algebraic subsets.The image $f(A)$ of a semi-algebraic subset $A \subseteq X$ by a definable mapping $f : X \to Y$ is semi-algebraic.

*Proof.* Let $G$ be the image of $f$. Note that the image $f(A)$ is the composition of a finite number of projections of the set $(A \times Y) \cap G$, which is the intersection of two semi-algebraic sets and hence semi-algebraic. By corollary 2, $f(A)$ is semi-algebraic. □

In fact, we applied quantifier elimination here, by using corollary 2. We could also have used quantifier elimination directly.

## 9.1 Continuity of semi-algebraic functions

**Definition 9.3.** Let $R$ be a real closed field. Futhermore, for $u \in R^n$ define the opertation $||$ by: $|u| = a \equiv u_1^2 + \ldots + u_n^2 = a^2 \wedge (0 < a \vee a = 0)$. Then a function $f : R \to R$ is called continuous at $x$ if:

$\forall \epsilon > 0 \exists \delta > 0 \ \forall y(|x - y| < \delta \to |f(x) - f(y)| < \epsilon)$.

**Theorem 10.** Let $f$ be a definable mapping from $R$ to $R$ for a real closed field $R$. Then every open $U \subseteq R$, contains an $x$ such that $f$ is continuous at $x$.

*Proof.* Since $T_{RCF}$ is complete in $L_{OR}$, every two models satisfy the same $L_{OR}$-sentences. Since the result to be proven can be expressed in such a sentence, it is sufficient to prove the result only for the real numbers $\mathbb{R}$. Let $U \subseteq \mathbb{R}$. We

divide the proof in two cases:

First, assume there is an open subset $V$ of $U$, such that $f(V)$ is finite. Then there must be an open interval contained in $V$ on which $f$ is constant. Clearly, in this case $f$ is continuous at an $x$ in this interval.

Now assume for all open subsets $V \subseteq U$, $f(V)$ is infinite. With opens $V_i$ contained in $U$, we can construct a chain $V_0 \supseteq V_1 \supseteq \ldots$ that has the following properties:

1. $Clos(V_{n+1}) \subseteq V_n$
2. $V_{n+1}$ is an open subinterval of $V_n \cap f^{-1}(a,b)$. Here $(a,b)$ is an interval contained in $f(V_n)$ that has a length smaller than $\frac{1}{n}$. Such an interval exists, because by Proposition 4 $f(V_n)$ is semi-algebraic and therefore definable. Since $T_{RCF}$ is o-minimal, $f(V_n)$ is a finite union of open intervals and points. Moreover, $f(V_n)$ is infinite, so it contains at least one open interval. Within this interval, we can find an open subinterval $(a,b)$ such that $b - a < \frac{1}{n}$.

Now let $x \in \cap_{i \in \mathbb{N}} V_i$ let $0 < \epsilon \in \mathbb{R}$. Choose a natural number $n$, such that $\frac{1}{n} < \epsilon$. Then the length of the interval $(a,b)$ is smaller then $\epsilon$. Since $x \in V_n \cap f^{-1}(a,b)$ and $f(V_n \cap f^{-1}(a,b)) \subseteq (a,b)$, we have that
$\exists \delta > 0 \; \forall y(|x-y| < \delta \rightarrow |f(x) - f(y)| < \epsilon)$. So $f$ is continuous at $x$. □

**Corollary 3.** Let $R$ be a real closed field and let $f : R \rightarrow R$ be a definable function. It is continuous at all but a finite number of points.

*Proof.* Suppose by contradiction that there are infinitely many points at which $f$ is not continuous. Then if we let $\phi(x)$ be the $L_{OR}$-formula that describes continuity of $f$ at $x$, the set
$C := \{x \in R | \neg\phi(x)\}$ is infinite. Note that $C$ is definable. It then follows from the o-minimality of $T_{RCF}$ that $C$ contains an interval $D$. But by Theorem 10, there is an $x \in D \subseteq C$ such that $f$ is continuous at $x$. Contradiction. □

## 9.2 Curve Selection

With the help of Corollary 3, we can prove a result called "Curve Selection" for the real numbers. We first consider the following Theorem about definable Skolem functions:

**Theorem 11.** In $L_{OR}$, $T_{RCF}$ has definable Skolem functions. If we let $R$ be a real closed field, this means that for every definable subset $A \subseteq R^{n+m}$, there exists a definable mapping $f : R^n \rightarrow R^m$ with the following property: for every $x \in R^n$, if there exists an $y \in R^m$ such that $(x,y) \in A$, then $(x, f(x)) \in A$.

*Proof.* To start with, assume $m = 1$. Suppose $A$ is a definable subset of $R^{n+1}$. Let $x \in R^n$. Then let $B_x$ be the subset of $R$ that contains all the $y \in R$ for

which $(x, y) \in A$. Note that, since $A$ is definable, $B_x$ is definable as well. As a consequence of the o-minimality of real closed fields, $B_x$ must be the finite union of points and open intervals. Now we can divide the proof into 5 cases. We define $f(x)$ to be equal to:

1. 0 if $B_x$ is empty or equal to $R$.
2. $c$ if $B_x$ has a least element $c$.
3. $d - 2$ if $B_x$ has a leftmost interval $(-\infty, d)$.
4. $d + 2$ if $B_x$ has a leftmost interval $(d, \infty)$.
5. $\frac{d+e}{2}$ if $B_x$ has a leftmost interval $(d, e)$

These values of $f(x)$ can all be described by $L_{OR}$-sentences and so its graph can. Furthermore, $f$ has the required property. Hence $f$ is a definable Skolem function.
Now note that we can find a Skolem function for all $m \in \mathbb{N}$ by induction. For example, if $m = 2$, let $B \subseteq R^{n+2}$ be definable. Then there is a definable Skolem function $g : R^{n+1} \to R$ for $B$ (note that this is not already the desired function). Moreover, we have that the projection $\Pi(B) \subseteq R^{n+1}$ is definable and so there is a definable Skolem function $f : R^n \to R$ for $\Pi(B)$. Then the function $h : R^n \to R^2$ defined by $h(x) = (f(x), g(x, f(x)))$ is a definable Skolem function for $B$.
This way, we can find definable Skolem functions for every $m$, so we can conclude that $T_{RCF}$ has definable Skolem functions. □

**Definition 9.4.** Let $R$ be a real closed field. Also, let $f : R \to R^n$ be a function. Then $a$ is called the limit of $f$ when $x$ goes to $b$ if:

$$\forall \epsilon > 0 \exists \delta > 0 (|x - b| < \delta \to |f(x) - a| < \epsilon)$$

.

**Definition 9.5.** Let $R$ be a real closed field. The closure $Clos(Y)$ of a subset $Y \subseteq R_n$ is defined by: $Clos(Y) = \{x \in R^n | \forall \delta > 0 \exists y \in Y | x - y| < \delta\}$.

**Theorem 12** (Curve Selection)**.** Let $\mathbb{R} = R$ and $Y \subseteq R^n$ be definable. In addition, let an element $c$ be contained in the closure of $Y$. Then there exist an $\epsilon > 0$ and a continuous function $f : (0, \epsilon) \to R^n$ with the property that its image is contained in $Y$ and that its limit when $x$ goes to 0 is $c$.

*Proof.* This can be proven with the help of Skolem functions. Consider the definable subset $A := \{(\delta, y) \in R^{n+1} | y \in Y \text{ and } |y - c| < \delta\}$. Then by Theorem 11, there is a definable function $f : R \mapsto R^n$, such that for all $\delta \in R$, if there exists an $y \in R^n$ for which $(\delta, y) \in A$, $(\delta, f(\delta)) \in A$. Since $c$ is in the closure of $Y$, this $y$ exists for every $\delta \in R$, so there is an interval $(0, \beta) \in R$ on which $|f(\delta) - c| < \delta$. Therefore, the limit of $f$ when $x$ goes to 0, is equal to $c$. Moreover, as a result of Corollary 3, there is an $\epsilon > 0$ for which $f$ is continuous on $(0, \epsilon)$. □

## 9.3 Extensions of Semi-algebraic Sets

Quantifier elimination allows us to say things about real closed extensions of real closed fields, since quantifier elimination implies model completeness. Moreover, we can talk about extensions of semi-algebraic sets:

**Definition 9.6.** Let $R$ be a real closed field and $K$ a real closed extension of $R$. Let furthermore $A$ be a semi-algebraic subset of $R$ defined by the $L_{OR}$-formula $\phi(x)$. Then the subset $A_K$ of $K$ defined by $\phi(x)$ is called the extension of $A$ to $K$.

Note that because it is definable, the extension $A_K$ is semi-algebraic as well.

**Proposition 5.** Let $R$ be a real closed field and $K$ a real closed extension of $R$. For two semi-algebraic sets $A \subseteq R^m$ and $B \subseteq R^n$, let
$f : A \to B$ be a definable function with graph $G$. Then $G_K$ is the graph of a definable mapping $f_K : A_K \to B_K$.

*Proof.* Let $\phi(x)$ and $\psi(y)$ be $L_{OR}$-formulas that define $A$ and $B$, respectively. In addition, let $\chi(x, y)$ be the formula that defines $G$. Since $G$ is a graph, the following sentences are true in $R$:

1. $\forall x(\phi(x) \leftrightarrow \exists y \chi(x, y))$
2. $\forall x(\forall y, \chi(x, y) \to \psi(y))$
3. $\forall x(\forall y \forall y'(\chi(x, y) \wedge \chi(x, y') \to y = y'))$

As a result of quantifier elimination, these sentences hold in the real closed extension $K$ of $R$, as well. Only in $K$, $\phi(x), \psi(y)$ and $\chi(x, y)$ respectively define $A_K, B_K$ and $G_K$. Therefore, the foregoing sentences now express the fact that $G_K$ is the graph of a definable mapping $f_K$ from $A_K$ to $B_K$. □

The mapping $f_K$ is known as the extension of $f$ to $K$. With the help of quantifier elimination, we can show that $f_K$ has many of the properties that $f$ has. For instance, when it comes to injectivity:

**Proposition 6.** Let $R$ be a real closed field and $A \subseteq R^m$ and $B \subseteq R^n$ semi-algebraic. Let $f : A \to B$ be a definable mapping. The mapping $f_K$ is injective if and only if $f$ is injective.

*Proof.* Let again $\phi(x)$ and $\psi(y)$ be $L_{OR}$-formulas that define $A$ and $B$, respectively. And let $\chi(x, y)$ be the formula that defines $G$. Injectivity of $f$ or $f_K$ is expressed by the $L_{OR}$-sentence:

$$\forall x \forall x' \forall y((\phi(x) \wedge \phi(x') \wedge \psi(y)) \to ((\chi(x, y) \wedge \chi(x', y)) \to x = x')$$

By quantifier elimination, this sentence holds in both $R$ and $K$ if it is true in one of both. □

Lastly, in a similar way it is possible to show for example that the extension $f_K$ of a mapping $f$ is continuous if and only if $f$ is. The same goes for differentiability. Simply by finding an $L_{OR}$-sentence that expresses the property in terms of $\phi(x)$, $\psi(y)$ and $\chi(x, y)$. And then concluding that this sentence is true in both $R$ and $K$ by quantifier elimination.

# References


[1] Frits Beukers. *Rings and Galois Theory*. 2016. URL: https://webspace.science.uu.nl/~beuke106/ringengalois/dic.pdf (visited on 01/19/2021).

[2] Jacek Bochnak, Michel Coste, and Marie-Francoise Roy. *Real Algebraic Geometry* -. Berlin Heidelberg: Springer Science Business Media, 2013. ISBN: 978-3-662-03718-8.

[3] David Marker, Margit Messmer, and Anand Pillay. *Model Theory of Fields* -. Cambridge: Cambridge University Press, 2017. ISBN: 978-1-107-16807-7.

[4] Ieke Moerdijk and Jaap van Oosten. *Sets, Models and Proofs* -. Berlin, Heidelberg: Springer, 2018. ISBN: 978-3-319-92414-4.

[5] Victoria Noquez. *Model Theory of Real Closed Fields, Louise Hay Logic Seminar, UIC*. 2012. URL: http://homepages.math.uic.edu/~noquez/pdfs/omintalk.pdf (visited on 01/19/2021).

[6] Katrin Tent and Martin Ziegler. *A Course in Model Theory* -. Cambridge: Cambridge University Press, 2012. ISBN: 978-0-521-76324-0.

[7] Lou Van den Dries. "Alfred Tarski's Elimination Theory for Real Closed Fields." In: *The Journal of Symbolic Logic* 53.1 (1988), pp. 7–19.