Quantifier Elimination and Applications

David Marker
Mathematical Logic

Fall 2015

1 Quantifier Elimination

In model theory we try to understand structures by studying their definable sets. Recall that if $\mathcal{M}$ is an $\mathcal{L}$-structure, then $X \subseteq M^n$ is *definable* if there is an $\mathcal{L}$-formula $\phi(v_1, \ldots, v_n, w_1, \ldots, w_m)$ and $b_1, \ldots, b_m \in M$ such that

$$X = \{\overline{a} \in M^n : \mathcal{M} \models \phi(\overline{a}, \overline{b})\}.$$

The study of definable sets is often complicated by quantifiers. For example, in the structure $(\mathbb{N}, +, \cdot, <, 0, 1)$ the quantifier-free definable sets are defined by polynomial equations and inequalities. Even if we use only existential quantifiers the definable sets become complicated. By the Matijasevič–Robinson–Davis–Putnam solution to Hilbert's 10th problem every recursively enumerable subset of $\mathbb{N}$ is defined by a formula

$$\exists v_1 \ldots \exists v_n \; p(x, v_1, \ldots, v_n) = 0$$

for some polynomial $p \in \mathbb{N}[X, Y_1, \ldots, Y_n]$. As we allow more alternations of quantifiers, we get even more complicated definable sets.

Not surprisingly, it will be easiest to study definable sets that are defined by quantifier-free formulas. Sometimes formulas with quantifiers can be shown to be equivalent to formulas without quantifiers. Here are two well-known examples. Let $\phi(a, b, c)$ be the formula

$$\exists x \; ax^2 + bx + c = 0.$$

By the quadratic formula,

$$\mathbb{R} \models \phi(a, b, c) \leftrightarrow [(a \neq 0 \wedge b^2 - 4ac \geq 0) \vee (a = 0 \wedge (b \neq 0 \vee c = 0))],$$

whereas in the complex numbers

$$\mathbb{C} \models \phi(a, b, c) \leftrightarrow (a \neq 0 \vee b \neq 0 \vee c = 0).$$

In either case, ϕ is equivalent to a quantifier-free formula. However, ϕ is not equivalent to a quantifier-free formula over the rational numbers $\mathbb{Q}$.

For a second example, let $\phi(a,b,c,d)$ be the formula

$$\exists x \exists y \exists u \exists v \ (xa + yc = 1 \wedge xb + yd = 0 \wedge ua + vc = 0 \wedge ub + vd = 1).$$

The formula $\phi(a,b,c,d)$ asserts that the matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is invertible. By the determinant test,

$$F \models \phi(a,b,c,d) \leftrightarrow ad - bc \neq 0$$

for any field F.

Definition 1.1 We say that a theory T has *quantifier elimination* if for every formula ϕ there is a quantifier-free formula ψ such that

$$T \models \phi \leftrightarrow \psi.$$

Our goal in this section is to give a very useful model theoretic test for elimination of quantifiers. In the next section we will show that this method can be applied to the theory of algebraically closed fields and develop some rich consequences. We begin by introducing some preliminary tools.

Diagrams

We begin by giving a way to construct $\mathcal{L}$-embeddings.

Definition 1.2 Suppose that $\mathcal{M}$ is an $\mathcal{L}$-structure. Let $\mathcal{L}_M$ be the language where we add to $\mathcal{L}$ constant symbols m for each element of M. The *atomic diagram* of $\mathcal{M}$ is $\{\phi(m_1,\ldots,m_n) : \phi$ is either an atomic $\mathcal{L}$-formula or the negation of an atomic $\mathcal{L}$-formula and $\mathcal{M} \models \phi(m_1,\ldots,m_n)\}$. We let $\mathrm{Diag}(\mathcal{M})$ denote the atomic diagram of $\mathcal{M}$

Lemma 1.3 *Suppose that $\mathcal{N}$ is an $\mathcal{L}_M$-structure and $\mathcal{N} \models \mathrm{Diag}(\mathcal{M})$; then, viewing $\mathcal{N}$ as an $\mathcal{L}$-structure, there is an $\mathcal{L}$-embedding of $\mathcal{M}$ into $\mathcal{N}$.*

Proof Let $j : M \to N$ be defined by $j(m) = m^{\mathcal{N}}$; that is, $j(m)$ is the interpretation of this constant symbol m in $\mathcal{N}$. If m_1, m_2 are distinct elements of M, then $m_1 \neq m_2 \in \mathrm{Diag}(\mathcal{M})$; thus, $j(m_1) \neq j(m_2)$ so j is an embedding. If f is a function symbol of $\mathcal{L}$ and $f^{\mathcal{M}}(m_1,\ldots,m_n) = m_{n+1}$, then $f(m_1,\ldots,m_n) = m_{n+1}$ is a formula in $\mathrm{Diag}(\mathcal{M})$ and $f^{\mathcal{N}}(j(m_1),\ldots,j(m_n)) = j(m_{n+1})$. If R is a relation symbol and $\overline{m} \in R^{\mathcal{M}}$, then $R(m_1,\ldots,m_n) \in \mathrm{Diag}(\mathcal{M})$ and $(j(m_1),\ldots,j(m_n)) \in R^{\mathcal{N}}$. Hence, j is an $\mathcal{L}$-embedding.

Quantifier Elimination Tests

Theorem 1.4 *Suppose that $\mathcal{L}$ contains a constant symbol c, T is an $\mathcal{L}$-theory, and $\phi(\overline{v})$ is an $\mathcal{L}$-formula. The following are equivalent:*

i) There is a quantifier-free $\mathcal{L}$-formula $\psi(\overline{v})$ such that $T \models \forall \overline{v}\ (\phi(\overline{v}) \leftrightarrow \psi(\overline{v}))$.

ii) If $\mathcal{M}$ and $\mathcal{N}$ are models of T, $\mathcal{A}$ is an $\mathcal{L}$-structure, $\mathcal{A} \subseteq \mathcal{M}$, and $\mathcal{A} \subseteq \mathcal{N}$, then $\mathcal{M} \models \phi(\overline{a})$ if and only if $\mathcal{N} \models \phi(\overline{a})$ for all $\overline{a} \in \mathcal{A}$.

Proof i)⇒ ii) Suppose that $T \models \forall \overline{v}\ (\phi(\overline{v}) \leftrightarrow \psi(\overline{v}))$, where ψ is quantifier-free. Let $\overline{a} \in \mathcal{A}$, where $\mathcal{A}$ is a common substructure of $\mathcal{M}$ and $\mathcal{N}$ and the latter two structures are models of T. In Proposition **??**, we saw that quantifier-free formulas are preserved under substructure and extension. Thus

$$\begin{aligned} \mathcal{M} \models \phi(\overline{a}) &\Leftrightarrow \mathcal{M} \models \psi(\overline{a}) \\ &\Leftrightarrow \mathcal{A} \models \psi(\overline{a}) \ \text{ (because } \mathcal{A} \subseteq \mathcal{M}) \\ &\Leftrightarrow \mathcal{N} \models \psi(\overline{a}) \ \text{ (because } \mathcal{A} \subseteq \mathcal{N}) \\ &\Leftrightarrow \mathcal{N} \models \phi(\overline{a}). \end{aligned}$$

ii) ⇒ i) First, if $T \models \forall \overline{v}\ \phi(\overline{v})$, then $T \models \forall \overline{v}\ (\phi(\overline{v}) \leftrightarrow c = c)$. Second, if $T \models \forall \overline{v}\ \neg\phi(\overline{v})$, then $T \models \forall \overline{v}\ (\phi(\overline{v}) \leftrightarrow c \neq c)$.

Thus, we may assume that both $T \cup \{\phi(\overline{v})\}$ and $T \cup \{\neg\phi(\overline{v})\}$ are satisfiable.

Let $\Gamma(\overline{v}) = \{\psi(\overline{v}) : \psi \text{ is quantifier-free and } T \models \forall \overline{v}\ (\phi(\overline{v}) \rightarrow \psi(\overline{v}))\}$. Let $d_1, \ldots, d_m$ be new constant symbols. We will show that $T \cup \Gamma(\overline{d}) \models \phi(\overline{d})$. Then, by compactness, there are $\psi_1, \ldots, \psi_n \in \Gamma$ such that

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

Thus

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

and $\bigwedge_{i=1}^{n} \psi_i(\overline{v})$ is quantifier-free. We need only prove the following claim.

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

Suppose not. Let $\mathcal{M} \models T \cup \Gamma(\overline{d}) \cup \{\neg\phi(\overline{d})\}$. Let $\mathcal{A}$ be the substructure of $\mathcal{M}$ generated by $\overline{d}$.

Let $\Sigma = T \cup \mathrm{Diag}(\mathcal{A}) \cup \phi(\overline{d})$. If Σ is unsatisfiable, then there are quantifier-free formulas $\psi_1(\overline{d}), \ldots, \psi_n(\overline{d}) \in \mathrm{Diag}(\mathcal{A})$ such that

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

But then

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

so $\bigvee_{i=1}^{n} \neg\psi_i(\overline{v}) \in \Gamma$ and $\mathcal{A} \models \bigvee_{i=1}^{n} \neg\psi_i(\overline{d})$, a contradiction. Thus, Σ is satisfiable.

Let $\mathcal{N} \models \Sigma$. Then $\mathcal{N} \models \phi(\overline{d})$. Because $\Sigma \supseteq \text{Diag}(\mathcal{A})$, $\mathcal{A} \subseteq \mathcal{N}$, by Lemma 1.3 i). But $\mathcal{M} \models \neg\phi(\overline{d})$; thus, by ii), $\mathcal{N} \models \neg\phi(\overline{d})$, a contradiction.

The proof above can easily be adapted to the case where $\mathcal{L}$ contains no constant symbols. In this case, there are no quantifier-free sentences, but for each sentence we can find a quantifier-free formula $\psi(v_1)$ such that $T \models \phi \leftrightarrow \psi(v_1)$.

The next lemma shows that we can prove quantifier elimination by getting rid of one existential quantifier at a time.

Lemma 1.5 *Let T be an $\mathcal{L}$-theory. Suppose that for every quantifier-free $\mathcal{L}$-formula $\theta(\overline{v}, w)$ there is a quantifier-free formula $\psi(\overline{v})$ such that $T \models \exists w\, \theta(\overline{v}, w) \leftrightarrow \psi(\overline{v})$. Then, T has quantifier elimination.*

Proof Let $\phi(\overline{v})$ be an $\mathcal{L}$-formula. We wish to show that $T \models \forall\overline{v}\ (\phi(\overline{v}) \leftrightarrow \psi(\overline{v}))$ for some quantifier-free formula $\phi(\overline{v})$. We prove this by induction on the complexity of $\phi(\overline{v})$.

If ϕ is quantifier-free, there is nothing to prove. Suppose that for $i = 0, 1$, $T \models \forall\overline{v}\ (\theta_i(\overline{v}) \leftrightarrow \psi_i(\overline{v}))$, where ψ_i is quantifier free.

If $\phi(\overline{v}) = \neg\theta_0(\overline{v})$, then $T \models \forall\overline{v}\ (\phi(\overline{v}) \leftrightarrow \neg\psi_0(\overline{v}))$.

If $\phi(\overline{v}) = \theta_0(\overline{v}) \wedge \theta_1(\overline{v})$, then $T \models \forall v\ (\phi(\overline{v}) \leftrightarrow (\psi_0(\overline{v}) \wedge \psi_1(\overline{v})))$.

In either case, ϕ is equivalent to a quantifier-free formula.

Suppose that $T \models \forall\overline{v}(\theta(\overline{v}, w) \leftrightarrow \psi_0(\overline{v}, w))$, where ψ_0 is quantifier-free and $\phi(\overline{v}) = \exists w\theta(\overline{v}, w)$. Then $T \models \forall\overline{v}\ (\phi(\overline{v}) \leftrightarrow \exists w\ \psi_0(\overline{v}, w))$. By our assumptions, there is a quantifier-free $\psi(\overline{v})$ such that $T \models \forall\overline{v}\ (\exists w\ \psi_0(\overline{v}, w) \leftrightarrow \psi(\overline{v}))$. But then $T \models \forall\overline{v}\ (\phi(\overline{v}) \leftrightarrow \psi(\overline{v}))$.

Combining Theorem 1.4 and Lemma 1.5 gives us the following simple, yet useful, test for quantifier elimination.

Corollary 1.6 *Let T be an $\mathcal{L}$-theory. Suppose that for all quantifier-free formulas $\phi(\overline{v}, w)$, if $\mathcal{M}, \mathcal{N} \models T$, $\mathcal{A}$ is a common substructure of $\mathcal{M}$ and $\mathcal{N}$, $\overline{a} \in A$, and there is $b \in M$ such that $\mathcal{M} \models \phi(\overline{a}, b)$, then there is $c \in N$ such that $\mathcal{N} \models \phi(\overline{a}, c)$. Then, T has quantifier elimination.*

Theories with Quantifier Elimination

We conclude with several observations about theories with quantifier elimination.

Definition 1.7 An $\mathcal{L}$-theory T is *model-complete* $\mathcal{M} \prec \mathcal{N}$ whenever $\mathcal{M} \subseteq \mathcal{N}$ and $\mathcal{M}, \mathcal{N} \models T$.

Stated in terms of embeddings: T is model-complete if and only if all embeddings are elementary.

Proposition 1.8 *If T has quantifier elimination, then T is model-complete.*

Proof Suppose that $\mathcal{M} \subseteq \mathcal{N}$ are models of T. We must show that $\mathcal{M}$ is an elementary submodel. Let $\phi(\overline{v})$ be an $\mathcal{L}$-formula, and let $\overline{a} \in M$. There is a quantifier-free formula $\psi(\overline{v})$ such that $\mathcal{M} \models \forall \overline{v}\ (\phi(\overline{v}) \leftrightarrow \psi(\overline{v}))$. Because quantifier-free formulas are preserved under substructures and extensions, $\mathcal{M} \models \psi(\overline{a})$ if and only if $\mathcal{N} \models \psi(\overline{a})$. Thus

$$\mathcal{M} \models \phi(\overline{a}) \Leftrightarrow \mathcal{M} \models \psi(\overline{a}) \Leftrightarrow \mathcal{N} \models \psi(\overline{a}) \Leftrightarrow \mathcal{N} \models \phi(\overline{a}).$$

There are model-complete theories that do not have quantifier elimination, but model completeness implies that we can eliminate all but the last existential quantifiers.

Proposition 1.9 *If T is model complete, then for any formula $\phi(\overline{v})$, there is a quantifier free formula $\psi(\overline{v}, \overline{w})$ such that*

$$T \models \forall \overline{v}\, [\phi(\overline{v}) \leftrightarrow \exists \overline{w}\ \psi(\overline{v}, \overline{w})]\,.$$

Let us just point out the following test for completeness of model-complete theories.

Proposition 1.10 *Let T be a model-complete theory. Suppose that there is $\mathcal{M}_0 \models T$ such that $\mathcal{M}_0$ embeds into every model of T. Then, T is complete.*

Proof If $\mathcal{M} \models T$, the embedding of $\mathcal{M}_0$ into $\mathcal{M}$ is elementary. In particular $\mathcal{M}_0 \equiv \mathcal{M}$. Thus, any two models of T are elementarily equivalent.

We will use Proposition 1.10 below in cases where Vaught's test does not apply.

We have provided a number of proofs of quantifier elimination without explicitly explaining how to take an arbitrary formula and produce a quantifier free one. In all of these cases, one can give explicit effective procedures. After the fact, the following lemma tells us that there is an algorithm to eliminate quantifiers.

Proposition 1.11 *Suppose that T is a decidable theory with quantifier elimination. Then, there is an algorithm which when given a formula ϕ as input will output a quantifier-free formula ψ such that $T \models \phi \leftrightarrow \psi$.*

Proof Given input $\phi(\overline{v})$ we search for a quantifier-free formula $\psi(\overline{v})$ such that $T \models \forall \overline{v}\ (\phi(\overline{v}) \leftrightarrow \psi(\overline{v}))$. Because T is decidable this is an effective search. Because T has quantifier elimination, we will eventually find ψ.

2 Algebraically Closed Fields

We now return to the theory of algebraically closed fields. In Proposition **??**, we proved that the theory of algebraically closed fields of a fixed characteristic is complete. We begin this section by showing that algebraically closed fields have quantifier elimination. For convenience we will formulate ACF in the language $\mathcal{L} = \{+, -, \cdot, 0, 1\}$. We add $-$ to the language, so that substructures are integral domains. Without $-$ we would have weaker structures that are a bit more cumbersome to deal with.

Theorem 2.1 ACF *has quantifier elimination.*

Proof

Suppose K and L are algebraically closed fields and A is an integral domain with $A \subseteq K \cap L$. By Corollary 5.6, we need to show that if $\phi(v, \overline{w})$ is a quantifier free formula, $\overline{a} \in A$, $b \in K$ and $K \models \phi(b, \overline{a})$, then there is $c \in L$ such that $L \models \phi(c, \overline{a})$.

Let F be the algebraic closure of the fraction field of A . We, may without loss of generality, assume that $F \subseteq K \cap L$. It will be enough to show that , $\overline{a} \in F$, and $K \models \phi(b, \overline{a})$ for some $b \in K$, then there is $c \in F$ such that $F \models \phi(c, \overline{a})$, for then, by Proposition 1.8, $L \models \phi(c, \overline{a})$.

We first note that ϕ can be put in disjunctive normal form, namely there are atomic or negated atomic formulas $\theta_{i,j}(\overline{v}, w)$ such that:

$$\phi(\overline{v}, w) \leftrightarrow \bigvee_{i=1}^{n} \bigwedge_{j=1}^{m} \theta_{i,j}(\overline{v}, w).$$

Because $K \models \phi(\overline{a}, b)$, $K \models \bigwedge_{j=1}^{m} \theta_{i,j}(\overline{a}, b)$ for some i. Thus, without loss of generality, we may assume that ϕ is a conjunction of atomic and negated atomic formulas. In our language atomic formulas $\theta(v_1, \ldots, v_n)$ are of the form $p(\overline{v}) = 0$, where $p \in \mathbb{Z}[X_1, \ldots, X_n]$. If $p(X, \overline{Y}) \in \mathbb{Z}[X, \overline{Y}]$, we can view $p(X, \overline{a})$ as a polynomial in $F[X]$. Thus, there are polynomials $p_1, \ldots, p_n, q_1, \ldots, q_m \in F[X]$ such that $\phi(v, \overline{a})$ is equivalent to

$$\bigwedge_{i=1}^{n} p_i(v) = 0 \wedge \bigwedge_{i=1}^{m} q_i(v) \neq 0.$$

If any of the polynomials p_i are nonzero, then b is algebraic over F. In this case, because F is algebraically closed, $b \in F$. Thus, we may assume that $\phi(v, \overline{a})$ is equivalent to

$$\bigwedge_{i=1}^{m} q_i(v) \neq 0.$$

But $q_i(X) = 0$ has only finitely many solutions for each $i \leq m$. Thus, there are only finitely many elements of F that do not satisfy F. Because algebraically closed fields are infinite, there is a $c \in F$ such that $F \models \phi(c, \overline{a})$.

Corollary 2.2 ACF *is model-complete and* ACF_p *is complete where* $p = 0$ *or* p *is prime.*

Proof Model-completeness is an immediate consequence of quantifier elimination.

The completeness of ACF_p was proved in Proposition **??**, but it also follows from quantifier elimination. Suppose that $K, L \models ACF_p$. Let ϕ be any sentence in the language of rings. By quantifier elimination, there is a quantifier-free sentence ψ such that

$$\mathrm{ACF} \models \phi \leftrightarrow \psi.$$

Because quantifier-free sentences are preserved under extension and substructure,

$$K \models \psi \Leftrightarrow \mathbb{F}_p \models \psi \Leftrightarrow L \models \psi,$$

where $\mathbb{F}_p$ is the p-element field if $p > 0$ and the rationals if $p = 0$. Thus,

$$K \models \phi \Leftrightarrow K \models \psi \Leftrightarrow L \models \psi \Leftrightarrow L \models \phi.$$

Thus $K \equiv L$ and ACF_p is complete.

Definable Sets and Constructible Sets

Quantifier elimination has a geometric interpretation. We begin by looking at the sets defined by quantifier free formulas.

Lemma 2.3 *Let* K *be a field. The subsets of* K^n *defined by atomic formulas are exactly those of the form* $V(p) = \{x \text{ for some } p \in K[\overline{X}]$. *A subset of* K^n *is quantifier-free definable if and only if it is a Boolean combination of Zariski closed subsets.*

Proof If $\phi(\overline{x}, \overline{y})$ is an atomic $\mathcal{L}_{\mathrm{r}}$-formula, then there is $q(\overline{X}, \overline{Y}) \in \mathbb{Z}[\overline{X}, \overline{Y}]$ such that $\phi(\overline{x}, \overline{y})$ is equivalent to $q(\overline{x}, \overline{y}) = 0$. If $X = \{\overline{x} : \phi(\overline{x}, \overline{a})\}$, then $X = V(q(\overline{X}, \overline{a}))$ and $q(\overline{X}, \overline{a}) \in K[\overline{X}]$. On the other hand, if $p \in K[\overline{X}]$, there is $q \in \mathbb{Z}[\overline{X}, \overline{Y}]$ and $\overline{a} \in K^m$ such that $p(\overline{X}) = q(\overline{X}, \overline{a})$. Then, $V(p)$ is defined by the quantifier-free formula $q(\overline{X}, \overline{a}) = 0$.

If $X \subseteq K^n$ is a finite Boolean combination of Zariski closed sets we call X *constructible*. If K is algebraically closed, the constructible sets have much stronger closure properties.

Corollary 2.4 *Let* K *be an algebraically closed field.*

i) $X \subseteq K^n$ *is constructible if and only if it is definable.*

ii) **(Chevalley's Theorem)** *The image of a constructible set under a polynomial map is constructible.*

Proof i) By Lemma 2.3, the constructible sets are exactly the quantifier-free definable sets, but by quantifier elimination every definable set is quantifier-free definable.

ii) Let $X \subseteq K^n$ be constructible and $p : K^n \to K^m$ be a polynomial map. Then, the image of $X = \{y \in K^m : \exists x \in K^n \ p(x) = y\}$. This set is definable and hence constructible.

Quantifier elimination has very strong consequences for definable subsets of K.

Corollary 2.5 *If K is an algebraically closed field and $X \subseteq K$ is definable, then either X or $K \setminus X$ is finite.*

Proof By quantifier elimination X is a finite Boolean combination of sets of the form $V(p)$, where $p \in K[X]$. But $V(p)$ is either finite or (if $p = 0$) all of K.

We say that a theory T is *strongly minimal* if for any $\mathcal{M} \models T$ and any definable $X \subseteq M$ either X or $M \setminus X$ is finite. This is a very powerful assumption. For example, it can be shown that any strongly minimal theory in a countable language is κ-categorical for every uncountable κ.

The model-completeness of algebraically closed fields can be used to give a proof of the Nullstellensatz.

Theorem 2.6 (Hilbert's Nullstellensatz) *Let K be an algebraically closed field. Suppose that I and J are radical ideals in $K[X_1, \ldots, X_n]$ and $I \subset J$. Then $V(J) \subset V(I)$. Thus $X \mapsto I(X)$ is a bijective correspondence between Zariski closed sets and radical ideals.*

Proof Let $p \in J \setminus I$. By Primary Decomposition, there is a prime ideal $P \supseteq I$ such that $p \notin P$. We will show that there is $x \in V(P) \subseteq V(I)$ such that $p(x) \neq 0$. Thus $V(I) \neq V(J)$. Because P is prime, $K[\overline{X}]/P$ is a domain and we can take F, the algebraic closure of its fraction field.

Let $q_1, \ldots, q_m \in K[X_1, \ldots, X_n]$ generate I. Let a_i be the element X_i/P in F. Because each $q_i \in P$ and $p \notin P$,

$$F \models \bigwedge_{i=1}^{m} q_i(\overline{a}) = 0 \wedge p(\overline{a}) \neq 0.$$

Thus

$$F \models \exists \overline{w} \ \bigwedge_{i=1}^{m} q_i(\overline{w}) = 0 \wedge p(\overline{w}) \neq 0$$

and by model-completeness

$$K \models \exists \overline{w} \ \bigwedge_{i=1}^{m} q_i(\overline{w}) = 0 \wedge p(\overline{w}) \neq 0.$$

Thus there is $\overline{b} \in K^n$ such that $q_1(\overline{b}) = \ldots = q_m(\overline{b}) = 0$ and $p(\overline{b}) \neq 0$. But then $\overline{b} \in V(P) \setminus V(J)$.

The next corollary is a simple consequence of model completeness.

Corollary 2.7 *Suppose $K \subseteq L$ are algebraically closed fields, V and W are varieties defined over K and $f : V \to W$ is a polynomial isomorphism defined over L. Then there is an isomorphism defined over K.*

Proof Suppose $f : V \to W$ is a polynomial isomorphism defined over L and f and f^{-1} both have degree at most d. As in the proof of Ax's Theorem we can write down an $\mathcal{L}$-formula Ψ with parameters from K saying that for some choice of coefficients there is a polynomial bijection from V between V and W where the polynomials have degree at most d. Since $L \models \Psi$, by model completeness, $K \models \Psi$. Thus we can choose an isomorphism defined over K.

Quantifier elimination gives us a powerful tool for analyzing definability in algebraically closed fields. For example, we will give the following characterization of definable functions.

Definition 2.8 Let $X \subseteq K^n$. We say that $f : X \to K$ is *quasirational* if either

i) K has characteristic zero and for some rational function $q(\overline{X}) \in K(X_1, \ldots, X_n)$, $f(\overline{x}) = q(\overline{x})$ on X, or

ii) K has characteristic $p > 0$ and for some rational function $q(\overline{X}) \in K(\overline{X})$, $f(\overline{x}) = q(\overline{x})^{\frac{1}{p^n}}$.

Rational functions are easily seen to be definable. In algebraically closed fields of characteristic p, the formula $x = y^p$ defines the function $x \mapsto x^{\frac{1}{p}}$, because every element has a unique p^{th}-root. Thus, every quasirational function is definable.

Proposition 2.9 *If $X \subseteq K^n$ is constructible and $f : X \to K$ is definable, then there are constructible sets $X_1, \ldots, X_m$ and quasirational functions $\rho_1, \ldots, \rho_m$ such that $\bigcup X_i = X$ and $f|X_i = \rho_i|X_i$.*

Proof Let $\Gamma(v_1, \ldots, v_n) = \{f(\overline{v}) \neq \rho(\overline{v}) : \rho \text{ a quasirational function}\} \cup \{\overline{v} \in X\} \cup \text{ACF} \cup \text{Diag}(K)$.

Claim Γ is not satisfiable.

Suppose that Γ is consistent. Let $L \models \text{ACF} + \text{Diag}(K)$ with $b_1, \ldots, b_n \in L$ such that for all $\gamma(\overline{v}) \in \Gamma$, $L \models \gamma(\overline{b})$.

Let K_0 be the subfield of L generated by K and $\overline{b}$. Then, K_0 is the closure of $B = \{b_1, \ldots, b_n\}$ under the rational functions of K. Let K_1 be the closure of B under all quasirational functions. If K has characteristic 0, then $K_0 = K_1$. If K has characteristic $p > 0$, $K_1 = \bigcup K_0^{\frac{1}{p^n}}$, the perfect closure of K_0.

By model-completeness, $K \prec L$, thus f^L, the interpretation of f in L, is a function from X^L to L, extending f. Because $L \models \Gamma(\overline{b})$, $f(\overline{b})$ is not in K_1. Because K_1 is perfect there is an automorphism α of L fixing K_1 pointwise such that $\alpha(f^L(\overline{b})) \neq f^L(\overline{b})$. But f^L is definable with parameters from K; thus, any automorphism of L which fixes K and fixes $\overline{a}$ must fix $f(\overline{a})$, a contradiction. Thus Γ is unsatisfiable.

Thus, by compactness, there are quasirational functions $\rho_1, \ldots, \rho_m$ such that

$$K \models \forall x \in X \bigwedge f(\overline{x}) = \rho_i(\overline{x}).$$

Let $X_i = \{\overline{x} \in X : f(\overline{x}) = \rho_i(\overline{x})\}$. Each X_i is definable.

We end by stating two more far reaching definability results for algebraically closed. They are a bit more involved–and ideally best understood using the model theoretic tool of *ω-stability* that we will not discuss in these lectures.

Let K be algebraically closed.

Theorem 2.10 (Elimination of Imaginaries) *Suppose $X \subseteq K^n$ is definable and E is a deifnable equivalence relation on X. There is a definable $f : X \to K^m$ for some m such that xEy if and only if $f(x) = f(y)$.*

This is related to the existence of fields of definitions. It is a useful tool for viewing projective, quasiprojective or abstract varieties (at least in the style of Weil) as constructible objects.

Theorem 2.11 *Let $G \subseteq K^n$ be a definable group. Then G is definably isomorphic to an algebraic group.*

Combining these we could conclude that if G is an algebraic group and H is a normal algebraic subgroup, then G/H is an algebraic group.

3 Real Closed Fields and o-minimality

In this section, we will concentrate on the field of real numbers. Unlike algebraically closed fields, the theory of the real numbers does not have quantifier elimination in $\mathcal{L}_{\mathrm{r}} = \{+, 1, \cdot, 0, 1\}$, the language of rings. The proof of Corollary 2.5 shows that any field with quantifier elimination is strongly minimal, whereas in $\mathbb{R}$, if $\phi(x)$ is the formula $\exists z\ z^2 = x$, then ϕ defines an infinite coinfinite definable set. In fact, algebraically closed fields are the only infinite fields with quantifier elimination.

In fact, the ordering is the only obstruction to quantifier elimination. We will eventually analyze the real numbers in the language $\mathcal{L}_{\mathrm{or}} = \{+, -, \cdots, <, 0, 1\}$ and show that we have quantifier elimination in this language. Because the ordering $x < y$ is definable in the real field by the formula

$$\exists\ z\ (z \neq 0 \wedge x + z^2 = y),$$

any subset of $\mathbb{R}^n$ definable using an $\mathcal{L}_{\mathrm{or}}$-formula is already definable using an $\mathcal{L}_{\mathrm{r}}$-formula). We will see that quantifier elimination in $\mathcal{L}_{\mathrm{or}}$ leads us to a good geometric understanding of the definable sets.

We begin by reviewing some of the necessary algebraic background on ordered fields. All of the algebraic results stated in this chapter are due to Artin and Schreier. These results are all proved in the appendix

Definition 3.1 We say that a field F is *orderable* if there is a linear order $<$ of F making $(F, <)$ an ordered field.

Although there are unique orderings of the fields $\mathbb{R}$ and $\mathbb{Q}$, orderable fields may have many possible orderings. The field of rational functions $\mathbb{Q}(X)$ has $2^{\aleph_0}$ distinct orderings. To see this, let x be any real number transcendental over $\mathbb{Q}$. The evaluation map $f(X) \mapsto f(x)$ is a field isomorphism between $\mathbb{Q}(X)$ and $\mathbb{Q}(x)$, the subfield of $\mathbb{R}$ generated by x. We can lift the ordering of the reals to an ordering $\mathbb{Q}(X)$ by $f(X) < g(X)$ if and only if $f(x) < g(x)$. Because $X < q$ if and only if $x < q$, choosing a different transcendental real would yield a different ordering. These are not the only orderings. We can also order $\mathbb{Q}(X)$ by making X infinite or infinitesimally close to a rational.

There is a purely algebraic characterization of the orderable fields.

Definition 3.2 We say that F is *formally real* if -1 is not a sum of squares.

In any ordered field all squares are nonnegative. Thus, every orderable field is formally real. The following result shows that the converse is also true.

Theorem 3.3 *If F is a formally real field, then F is orderable. Indeed, if $a \in F$ and $-a$ is not a sum of squares of elements of F, then there is an ordering of F where a is positive.*

Because the field of complex numbers is the only proper algebraic extension of the real field, the real numbers have no proper formally real algebraic extensions. Fields with this property will play a key role.

Definition 3.4 A field F is *real closed* if it is formally real with no proper formally real algebraic extensions.

Although it is not obvious at first that real closed fields form an elementary class, the next theorem allows us to axiomatize the real closed fields.

Theorem 3.5 *Let F be a formally real field. The following are equivalent.*
i) F is real closed.
ii) $F(i)$ is algebraically closed (where $i^2 = -1$).
iii) For any $a \in F$, either a or $-a$ is a square and every polynomial of odd degree has a root.

Corollary 3.6 *The class of real closed fields is an elementary class of $\mathcal{L}_{\mathrm{r}}$-structures.*

Proof We can axiomatize real closed fields by:
i) axioms for fields
ii) for each $n \geq 1$, the axiom

$$\forall x_1 \ldots \forall x_n \; x_1^2 + \ldots + x_n^2 + 1 \neq 0$$

iii) $\forall x \exists y \; (y^2 = x \vee y^2 + x = 0)$
iv) for each $n \geq 0$, the axiom

$$\forall x_0 \ldots \forall x_{2n} \exists y \; y^{2n+1} + \sum_{i=0}^{2n} x_i y^i = 0.$$

Although we can axiomatize real closed fields in the language of rings, we already noticed that we do not have quantifier elimination in this language. Instead, we will study real closed fields in $\mathcal{L}_{\mathrm{or}}$, the language of ordered rings. If F is a real closed field and $0 \neq a \in F$, then exactly one of a and $-a$ is a square. This allows us to order F by

$$x < y \;\text{ if and only if } y - x \text{ is a nonzero square.}$$

It is easy to check that this is an ordering and it is the only possible ordering of F.

Definition 3.7 We let RCF be the $\mathcal{L}_{\mathrm{or}}$-theory axiomatized by the axioms above for real closed fields and the axioms for ordered fields.

The models of RCF are exactly real closed fields with their canonical ordering. Because the ordering is defined by the $\mathcal{L}_{\mathrm{r}}$-formula

$$\exists z \; (z \neq 0 \wedge x + z^2 = y),$$

the next result tells us that using the ordering does not change the definable sets.

Proposition 3.8 *If F is a real closed field and $X \subseteq F^n$ is definable by an $\mathcal{L}_{\mathrm{or}}$-formula, then X is definable by an $\mathcal{L}_{\mathrm{r}}$-formula.*

Proof Replace all instances of $t_i < t_j$ by $\exists v\ (v \neq 0 \wedge v^2 + t_i = t_j)$, where t_i and t_j are terms occurring in the definition of X.

The next result suggests another possible axiomatization of RCF.

Theorem 3.9 *An ordered field F is real closed if and only if whenever $p(X) \in F[X]$, $a, b \in X$, $a < b$, and $p(a)p(b) < 0$, there is $c \in F$ such that $a < c < b$ and $p(c) = 0$.*

Definition 3.10 If F is a formally real field, a *real closure* of F is a real closed algebraic extension of F.

By Zorn's Lemma, every formally real field F has a maximal formally real algebraic extension. This maximal extension is a real closure of F.

The real closure of a formally real field may not be unique. Let $F = \mathbb{Q}(X)$, $F_0 = F(\sqrt{X})$, and $F_1 = F(\sqrt{-X})$. By Theorem 3.3, F_0 and F_1 are formally real. Let R_i be a real closure of F_i. There is no isomorphism between R_0 and R_1 fixing F because X is a square in R_0 but not in R_1. Thus, some work needs to be done to show that any ordered field $(F, <)$ has a real closure where the canonical order extends the ordering of F.

Lemma 3.11 *If $(F, <)$ is an ordered field, $0 < x \in F$, and x is not a square in F, then we can extend the ordering of F to $F(\sqrt{x})$.*

Proof We can extend the ordering to $F(\sqrt{x})$ by $0 < a + b\sqrt{x}$ if and only if
i) $b = 0$ and $a > 0$, or
ii) $b > 0$ and ($a > 0$ or $x > \frac{a^2}{b^2}$), or
iii) $b < 0$ and ($a < 0$ and $x < \frac{a^2}{b^2}$).

Corollary 3.12 *If $(F, <)$ is an ordered field, there is a real closure R of F such that the canonical ordering of R extends the ordering on F.*

Proof

By successive applications of Lemma 3.11, we can find an ordered field $(L, <)$ extending $(F, <)$ such that every positive element of F has a square root in L. We now apply Zorn's Lemma to find a maximal formally real algebraic extension R of L. Because every positive element of F is a square in R, the canonical ordering of R extends the ordering of F.

Although a formally real field may have nonisomorphic real closures, if $(F, <$) is an ordered field there will be a unique real closure compatible with the ordering of F.

Theorem 3.13 *If $(F, <)$ is an ordered field, and R_1 and R_2 are real closures of F where the canonical ordering extends the ordering of F, then there is a unique field isomorphism $\phi : R_1 \to R_2$ that is the identity on F.*

Note that because the ordering of a real closed field is definable in $\mathcal{L}_{\mathrm{r}}$, ϕ also preserves the ordering. We often say that any ordered field $(F, <)$ has a unique real closure. By this we mean that there is a unique real closure that extends the given ordering.

Quantifier Elimination for Real Closed Fields

We are now ready to prove quantifier elimination.

Theorem 3.14 *The theory* RCF *admits elimination of quantifiers in $\mathcal{L}_{\mathrm{or}}$.*

Proof We use the quantifier elimination tests of §5. Suppose K and L are real closed ordered fields and A is a common substructure. Then A is an ordered integral domain. We extend the ordering on A to its fraction field to obtain an ordered subfield $F_0 \subseteq K \cap L$. Let F be the real closure of F_0. By uniqueness of real closures, F is isomorphic, as an ordered field, to the algebraic closure of F_0 inside K and L. Without loss of generality we may assume $F \subseteq K \cap L$.

It suffices then to show that if $\phi(v, \overline{w})$ is a quantifier-free formula, $\overline{a} \in F$, $b \in K$ and $K \models \phi(b, \overline{a})$, then there is $b' \in F$ such that $F \models \phi(b', \overline{a})$.

Note that

$$p(X) \neq 0 \leftrightarrow \ (p(\overline{X}) > 0 \vee -p(\overline{X}) > 0)$$

and

$$p(\overline{X}) \not> 0 \leftrightarrow (p(\overline{X}) = 0 \vee -p(\overline{X}) > 0).$$

With this in mind, we may assume that ϕ is a disjunction of conjunctions of formulas of the form $p(v, \overline{w}) = 0$ or $p(v, \overline{w}) > 0$. As in Theorem 2.1, we may assume that there are polynomials $p_1, \ldots, p_n$ and $q_1, \ldots, q_m \in F[X]$ such that

$$\phi(v, \overline{a}) \leftrightarrow \bigwedge_{i=1}^{n} p_i(v) = 0 \wedge \bigwedge_{i=1}^{m} q_i(v) > 0.$$

If any of the polynomials $p_i(X)$ is nonzero, then b is algebraic over F. Because F has no proper formally real algebraic extensions, in this case $b \in F$. Thus, we may assume that

$$\phi(v, \overline{a}) \leftrightarrow \bigwedge_{i=1}^{m} q_i(v) > 0.$$

The polynomial $q_i(X)$ can only change signs at zeros of q_i and all zeros of q_i are in F. Thus, we can find $c_i, d_i \in F$ such that $c_i < b < d_i$ and $q_i(x) > 0$ for all $x \in (c_i, d_i)$. Let $c = \max(c_1, \ldots, c_m)$ and $d = \min(d_1, \ldots, d_m)$. Then, $c < d$ and $\bigwedge_{i=1}^{m} q_i(x) > 0$ whenever $c < x < d$. Thus, we can find $b' \in F$ such that $F \models \phi(b', \overline{a})$.

Corollary 3.15 RCF *is complete, model complete and decidable. Thus* RCF *is the theory of* $(\mathbb{R}, +, \cdot, <)$ *and* RCF *is decidable.*

Proof By quantifier elimination, RCF is model complete.

Every real closed field has characteristic zero; thus, the rational numbers are embedded in every real closed field. Therefore, $\mathbb{R}_{\text{alg}}$, the field of real algebraic numbers (i.e., the real closure of the rational numbers) is a subfield of any real closed field. Thus, for any real closed field R, $\mathbb{R}_{\text{alg}} \prec R$, so $R \equiv \mathbb{R}_{\text{alg}}$.

In particular, $R \equiv \mathbb{R}_{\text{alg}} \equiv \mathbb{R}$.

Because RCF is complete and recursively axiomatized, it is decidable.

Semialgebraic Sets

Quantifier elimination for real closed fields has a geometric interpretation.

Definition 3.16 Let F be an ordered field. We say that $X \subseteq F^n$ is *semialgebraic* if it is a Boolean combination of sets of the form $\{\overline{x} : p(\overline{x}) > 0\}$, where $p(\overline{X}) \in F[X_1, \ldots, X_n]$.

By quantifier elimination, the semialgebraic sets are exactly the definable sets. The next corollary is a geometric restatement of quantifier elimination. It is analogous to Chevalley's Theorem (2.4) for algebraically closed fields.

Corollary 3.17 (Tarski–Seidenberg Theorem) *The semialgebraic sets are closed under projection.*

The next corollary is a typical application of quantifier elimination.

Corollary 3.18 *If* $F \models RCF$ *and* $A \subseteq F^n$ *is semialgebraic, then the closure (in the Euclidean topology) of* A *is semialgebraic.*

Proof We repeat the main idea of Lemma **??**. Let d be the definable function

$$d(x_1, \ldots, x_n, y_1, \ldots, y_n) = z \text{ if and only if } z \geq 0 \wedge z^2 = \sum_{i=1}^{n} (x_i - y_i)^2.$$

The closure of A is

$$\{\overline{x} : \forall \epsilon > 0 \; \exists \overline{y} \in A \; d(\overline{x}, \overline{y}) < \epsilon\}.$$

Because this set is definable, it is semialgebraic.

We say that a function is semialgebraic if its graph is semialgebraic. The next result shows how we can use the completeness of RCF to transfer results from $\mathbb{R}$ to other real closed fields.

Corollary 3.19 *Let* F *be a real closed field. If* $X \subseteq F^n$ *is semialgebraic, closed and bounded, and* f *is a continuous semialgebraic function, then* $f(X)$ *is closed and bounded.*

Proof If $F = \mathbb{R}$, then X is closed and bounded if and only if X is compact. Because the continuous image of a compact set is compact, the continuous image of a closed and bounded set is closed and bounded.

In general, there are $\overline{a}, \overline{b} \in F$ and formulas ϕ and ψ such that $\phi(\overline{x}, \overline{a})$ defines X and $\psi(\overline{x}, y, \overline{b})$ defines $f(\overline{x}) = y$. There is a sentence Φ asserting:

> $\forall \overline{u}, \overline{w}$ [if $\psi(\overline{x}, y, \overline{w})$ defines a continuous function with domain $\phi(\overline{x}, \overline{u})$ and $\phi(\overline{x}, \overline{u})$ is a closed and bounded set, then the range of the function is closed and bounded].

By the remarks above, $\mathbb{R} \models \Phi$. Therefore, by the completeness of RCF, $F \models \Phi$ and the range of f is closed and bounded.

Model-completeness has several important applications. A typical application is Abraham Robinson's simple proof of Artin's positive solution to Hilbert's 17th problem.

Definition 3.20 Let F be a real closed field and $f(\overline{X}) \in F(X_1, \ldots, X_n)$ be a rational function. We say that f is *positive semidefinite* if $f(\overline{a}) \geq 0$ for all $\overline{a} \in F^n$.

Theorem 3.21 (Hilbert's 17th Problem) *If f is a positive semidefinite rational function over a real closed field F, then f is a sum of squares of rational functions.*

Proof Suppose that $f(X_1, \ldots, X_n)$ is a positive semidefinite rational function over F that is not a sum of squares. By Theorem 3.3, there is an ordering of $F(\overline{X})$ so that f is negative. Let R be the real closure of $F(\overline{X})$ extending this order. Then

$$R \models \exists \overline{v}\ f(\overline{v}) < 0$$

because $f(\overline{X}) < 0$ in R. By model-completeness

$$F \models \exists \overline{v}\ f(\overline{v}) < 0,$$

contradicting the fact that f is positive semidefinite.

We will show that quantifier elimination gives us a powerful tool for understanding the definable subsets of a real closed field.

Definition 3.22 Let $\mathcal{L} \supseteq \{<\}$. Let T be an $\mathcal{L}$-theory extending the theory of linear orders. We say that T is *o-minimal* if for all $\mathcal{M} \models T$ if $X \subseteq M$ is definable, then X is a finite union of points and intervals with endpoints in $M \cup \{\pm\infty\}$.

We can think of o-minimality as an analog of strong minimality for ordered structures. Strong minimality says that the only definable subsets in dimension one can be defined using only equality–i.e., the ones that can be defined in any structure. O-minimality says the only sets that can be defined in one dimension are the ones definable in any ordered structure.

Corollary 3.23 RCF*is an o-minimal theory.*

Proof Let $R \models \text{RCF}$. We need to show that every definable subset of R is a finite union of points and intervals with endpoints in $R \cup \{\pm\infty\}$. By quantifier elimination, very definable subset of R is a finite Boolean combination of sets of the form

$$\{x \in R : p(x) = 0\}$$

and

$$\{x \in R : q(x) > 0\}.$$

Solution sets to nontrivial equations are finite, whereas sets of the second form are finite unions of intervals. Thus, any definable set is a finite union of points and intervals.

Next we will show that definable functions in one variable are piecewise continuous. The first step is to prove a lemma about $\mathbb{R}$ that we will transfer to all real closed fields.

Lemma 3.24 *If $f : \mathbb{R} \to \mathbb{R}$ is semialgebraic, then for any open interval $U \subseteq \mathbb{R}$ there is a point $x \in U$ such that f is continuous at x.*

Proof
case 1: There is an open set $V \subseteq U$ such that f has finite range on V.

Pick an element b in the range of f such that $\{x \in V : f(x) = b\}$ is infinite. By o-minimality, there is an open set $V_0 \subseteq V$ such that f is constantly b on V.

case 2: Otherwise.

We build a chain $U = V_0 \supset V_1 \supset V_2 \ldots$ of open subsets of U such that the closure $\overline{V}_{n+1}$ of V_{n+1} is contained in V_n. Given V_n, let X be the range of f on V_n. Because X is infinite, by o-minimality, X contains an interval (a, b) of length at most $\frac{1}{n}$. The set $Y = \{x \in V_n : f(x) \in (a, b)\}$ contains a suitable open interval V_{n+1}. Because $\mathbb{R}$ is locally compact,

$$\bigcap_{i=1}^{\infty} V_i = \bigcap_{i=1}^{\infty} \overline{V_i} \neq \emptyset.$$

If $x \in \bigcap_{i=1}^{\infty} V_i$, then f is continuous at x.

The proof above makes essential use of the completeness of the ordering of the reals. However, because the statement is first order, it is true for all real closed fields, by the completeness of RCF.

Corollary 3.25 *Let F be a real closed field and $f : F \to F$ is a semialgebraic function. Then, we can partition F into $I_1 \cup \ldots \cup I_m \cup X$, where X is finite and the I_j are pairwise disjoint open intervals with endpoints in $F \cup \{\pm\infty\}$ such that f is continuous on each I_j.*

Proof Let

$$D = \{x : F \models \exists \epsilon > 0 \ \forall \delta > 0 \ \exists y \ |x - y| < \delta \wedge |f(x) - f(y)| > \epsilon\}$$

be the set of points where f is discontinuous. Because D is definable, by o-minimality D is either finite or has a nonempty interior. By Corollary 3.23, D must be finite. Thus, $F \setminus D$ is a finite union of intervals on which F is continuous.

If F is real closed, then o-minimality tells us what the definable subsets of F look like. Definable subsets of F^n are also relatively simple.

Definition 3.26 We inductively define the collection of *cells* as follows.

- $X \subseteq F^n$ is a 0-cell if it is a single point.
- $X \subseteq F$ is a 1-cell if it is an interval (a, b), where $a \in F \cup \{-\infty\}$, $b \in F \cup \{+\infty\}$, and $a < b$.
- If $X \subseteq F^n$ is an n-cell and $f : X \to F$ is a continuous definable function, then $Y = \{(\overline{x}, f(\overline{x})) : \overline{x} \in X\}$ is an n-cell.
- Let $X \subseteq F^n$ be an n-cell. Suppose that f is either a continuous definable function from X to F or identically $-\infty$ and g is either a continuous definable function from X to F such that $f(\overline{x}) < g(\overline{x})$ for all $\overline{x} \in X$ or g is identically $+\infty$; then

$$Y = \{(\overline{x}, y) : \overline{x} \in X \wedge f(\overline{x}) < y < g(\overline{x})\}$$

is an $n + 1$-cell.

In a real closed field, every nonempty definable set is a finite disjoint union of cells. The proof relies on the following lemma.

Lemma 3.27 (Uniform Bounding) *Let $X \subseteq F^{n+1}$ be semialgebraic. There is a natural number N such that if $\overline{a} \in F^n$ and $X_{\overline{a}} = \{y : (\overline{a}, y) \in X\}$ is finite, then $|X_{\overline{a}}| < N$.*

Proof First, note that $X_{\overline{a}}$ is infinite if and only if there is an interval (c, d) such that $(c, d) \subseteq X_{\overline{a}}$. Thus $\{(\overline{a}, b) \in X : X_{\overline{a}} \text{ is finite}\}$ is definable. Without loss of generality, we may assume that for all $\overline{a} \in F^n$, $X_{\overline{a}}$ is finite. In particular, we may assume that

$$F \models \forall \overline{x} \forall c \forall d \neg [c < d \wedge \forall y (c < y < d \to y \in X_{\overline{a}})].$$

Consider the following set of sentences in the language of fields with constants added for each element of F and new constants $c_1, \ldots, c_n$. Let Γ be

$$\mathrm{RCF} + \mathrm{Diag}(F) + \left\{ \exists y_1, \ldots, y_m \ \left[\bigwedge_{i<j} y_i \neq y_j \wedge \bigwedge_{i=1}^{m} y_i \in X_{\overline{c}} \right] : m \in \omega \right\}$$

Suppose that Γ is satisfiable. Then, there is a real closed field $K \supseteq F$ and elements $\overline{c} \in K^n$ such that $X_{\overline{c}}$ is infinite. By model-completeness, $F \prec K$. Therefore

$$K \models \forall \overline{x} \forall c, d \ \neg [c < d \ \wedge \forall y \ (c < y < d \to y \in X_{\overline{a}})].$$

This contradicts the o-minimality of K. Thus, Γ is unsatisfiable and there is an N such that

$$\mathrm{RCF} + \mathrm{Diag}(F) \models \ \forall \overline{x} \ \neg \left(\exists y_1, \ldots, y_N \ \left[\bigwedge_{i<j} y_i \neq y_j \wedge \bigwedge_{i=1}^{N} y_i \in X_{\overline{x}} \right] \right).$$

In particular, for all $\overline{a} \in F^n, |X_{\overline{a}}| < N$.

We now state the Cell Decomposition Theorem and give the proof for subsets of F^2. In the exercises, we will outline the results needed for the general case.

Theorem 3.28 (Cell Decomposition) *Let $X \subseteq F^m$ be semialgebraic. There are finitely many pairwise disjoint cells $C_1, \ldots, C_n$ such that $X = C_1 \cup \ldots \cup C_n$.*

Proof (for $m = 2$) For each $a \in F$, let

$$C_a = \{x : \forall \epsilon > 0 \exists y, z \in (x - \epsilon, x + \epsilon)\ [(a, y) \in X \wedge (a, z) \notin X]\}.$$

We call C_a the *critical values* above a. By o-minimality, there are only finitely many critical values above a. By uniform bounding, there is a natural number N such that for all $a \in F$, $|C_a| \leq N$. We partition F into $A_0, A_1, \ldots, A_N$, where $A_n = \{a : |C_a| = n\}$.

For each $n \leq N$, we have a definable function $f_n : A_1 \cup \ldots \cup A_n \to F$ by $f_n(a) = n$th element of C_a. As above, $X_a = \{y : (a, y) \in X\}$.

For $n \leq N$ and $a \in A_n$, we define $P_a \in 2^{2n+1}$, the *pattern* of X above a, as follows.

If $n = 0$, then $P_a(0) = 1$ if and only if $X_a = F$. Suppose that $n > 0$.

$P_a(0) = 1$ if and only if $x \in X_a$ for all $x < f_1(a)$.

$P_a(2i - 1) = 1$ if and only if $f_i(a) \in X$.

For $i < n$, $P_a(2i) = 1$ if and only if $x \in X_a$ for all $x \in (f_i(a), f_{i+1}(a))$.

$P(2n) = 1$ if and only if $x \in X_a$ for all $x > f_n(a)$.

For each possible pattern $\sigma \in 2^{2n+1}$, let $A_{n,\sigma} = \{a \in A_n : P_a = \sigma\}$. Each $A_{n,\sigma}$ is semialgebraic. For each $A_{n,\sigma}$, we will give a decomposition of $\{(x, y) \in X : x \in A_{n,\sigma}\}$ into disjoint cells. Because the $A_{n,\sigma}$ partition F, this will suffice.

Fix one $A_{n,\sigma}$. By Corollary 3.25, we can partition $A_{n,\sigma} = C_1 \cup \ldots \cup C_l$, where each C_j is either an interval or a singleton and f_i is continuous on C_j for $i \leq n, j \leq l$. We can now give a decomposition of $\{(x, y) : x \in A_{n,\sigma}\}$ into cells such that each cell is either contained in X or disjoint from X.

For $j \leq l$, let $D_{j,0} = \{(x, y) : x \in C_j \text{ and } y < f(1)\}$.

For $j \leq l$ and $1 \leq i \leq n$, let $D_{j,2i-1} = \{(x, f_i(x)) : x \in C_j\}$.

For $j \leq l$ and $1 \leq i < n$, let $D_{j,2i} = \{(x, y) : x \in C_j, f_i(x) < y < f_{i+1}(x)\}$.

For $j \leq l$, let $D_{j,2n} = \{(x, y) : x \in C_j, y > f_n(x)\}$.

Clearly, each $D_{j,i}$ is a cell, $\bigcup D_{j,i} = \{(x, y) : x \in A_{n,\sigma}\}$, and each $D_{j,i}$ is either contained in X or disjoint from X. Thus, taking the $D_{j,i}$ that are contained in X, we get a partition of $\{(x, y) \in X : x \in A_{n,\sigma}\}$ into disjoint cells.

o-minimal Expansions of $\mathbb{R}$

The proofs above used very little about semialgebraic sets beyond o-minimality. Indeed, they would work in any o-minimal expansion of the real field. Indeed, there is a rich theory of definable sets in o-minimal expansions of the reals. We

will survey some of the results in this section. For full details, see van den Dries book *Tame topology and o-minimal structures.*

Let $\mathcal{R} = (\mathbb{R}, +, \cdot, <, \ldots)$ be an o-minimal expansion of the reals, i.e., a structure obtained by adding extra structure to the reals such that $\mathrm{Th}(\mathcal{R})$ is o-minimal. Below by "definable" we will mean definable in $\mathcal{R}$.

Theorem 3.29 *Assume $\mathcal{R}$ is an o-minima expansion of $\mathbb{R}$.*

i) Every definable subset of $\mathbb{R}^n$ is a finite union of cells.

ii) If $f : X \to \mathbb{R}^n$ is definable, there is a finite partition of X into cells $X_1, \cup, X_n$ such that $f|X_i$ is continuous for each i. Indeed, for any $r \geq 0$, we can choose the partition such that $f|X_i$ is $\mathcal{C}^r$ for each i.

An easy consequence of ii) is that definable sets have only finitely many connected components. Much more is true, for example:

- Definable bounded sets can be definably triangulated.
- Suppose $X \subseteq \mathbb{R}^{n+m}$ is definable. For $a \in \mathbb{R}^m$ let

$$X_a = \{\overline{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n : (\overline{x}, a) \in X\}.$$

There are only finitely many definable homeomorphism types for the sets X_a.

- (Curve selection) If $X \subseteq \mathbb{R}^n$ is definable and a is in the closure of X, then there is a continuous definable $f : (0,1) \to X$ such that

$$\lim_{x \to 1} f(x) = a.$$

- If G is a definable group, then G is definably isomorphic to a Lie group.
- If we assume in addition that all definable functions are majorized by polynomials, then many of the metric properties of semialgebraic sets and asymptotic properties of semialgebraic functions also generalize.

Of course, this leads to the question: are there interesting o-minimal expansions of $\mathbb{R}$?

$\mathbb{R}_{\mathrm{an}}$ and subanalytic sets

Most of the results on o-minimal structures mentioned above were proved before we knew of any interesting o-minimal structures other than the real field. The first new example of an o-minimal theory was given by van den Dries.

Let $\mathcal{L}_{\mathrm{an}} = \mathcal{L} \cup \{\widehat{f} : \text{for some open } U \supset [0,1]^n,\ f : U \to \mathbb{R} \text{ is analytic}\}$.

We define $\widehat{f} : \mathbb{R}^n \to \mathbb{R}$ by

$$\widehat{f}(x) = \begin{cases} f(x) & x \in [0,1]^n \\ 0 & \text{otherwise.} \end{cases}$$

We let $\mathbb{R}_{\mathrm{an}}$ be the resulting $\mathcal{L}_{\mathrm{an}}$-structure. Denef and van den Dries proved that $\mathbb{R}_{\mathrm{an}}$ is o-minimal and that $\mathbb{R}_{\mathrm{an}}$ has quantifier elimination if we add a function

$$D(x,y) = \begin{cases} x/y & \text{if } 0 \leq |x| \leq |y| \\ 0 & \text{otherwise} \end{cases}$$

to the language. Quantifier elimination is proven by using the Weierstrass preparation theorem to replace arbitrary analytic functions of several variables by analytic functions that are polynomial in one of the variables. Tarski's elimination procedure is then used to eliminate this variable.

Denef and van den Dries also showed that if $f : \mathbb{R} \to \mathbb{R}$ is definable in $\mathbb{R}_{\text{an}}$, then f is asymptotic to a rational function. In particular, although we can define the restriction of the exponential function to bounded intervals, we cannot define the exponential function globally. It is also impossible to define the sine function globally; for its zero set would violate o-minimality.

Although $\mathbb{R}_{\text{an}}$ may seem unnatural, the definable sets form an interesting class.

We say that $X \subseteq \mathbb{R}^n$ is *semi-analytic* if for all x in $\mathbb{R}^n$ there is an open neighborhood U of x such that $X \cap U$ is a finite Boolean combination of sets $\{\overline{x} \in U : f(\overline{x}) = 0\}$ and $\{\overline{x} \in U : g(\overline{x}) > 0\}$ where $f, g : U \to \mathbb{R}$ are analytic. We say that $X \subseteq \mathbb{R}^n$ is *subanalytic* if for all x in $\mathbb{R}^n$ there is an open U and $Y \subset \mathbb{R}^{n+m}$ a bounded semianalytic set such that $X \cap U$ is the projection of Y into U. It is well known that subanalytic sets share many of the nice properties of semialgebraic sets.

If $X \subset \mathbb{R}^n$ is bounded, then X is definable in $\mathbb{R}_{\text{an}}$ if and only if X is subanalytic. Indeed $Y \subseteq \mathbb{R}^n$ is definable in $\mathbb{R}_{\text{an}}$ if and only if it is the image of a bounded subanalytic set under a semialgebraic map. Most of the known properties of subanalytic sets generalize to sets defined in any polynomial bounded o-minimal theory.

Exponentiation

The big breakthrough in the subject came in 1991. While quantifier elimination for $\mathbb{R}_{\text{exp}}$ is impossible, Wilkie proved the next best thing.

Theorem 3.30 (Wilkie) *Let $\phi(x_1, \ldots, x_m)$ be an $\mathcal{L}_{\text{exp}}$ formula. Then there is $n \geq m$ and $f_1, \ldots, f_s \in \mathbb{Z}[x_1, \ldots, x_n, e^{x_1}, \ldots, e^{x_n}]$ such that $\phi(x_1, \ldots, x_n)$ is equivalent to*

$$\exists x_{m+1} \ldots \exists x_n \; f_1(x_1, \ldots, x_n, e^{x_1}, \ldots, e^{x_n}) = \ldots = f_s(x_1, \ldots, x_n, e^{x_1}, \ldots, e^{x_n}) = 0.$$

Thus every formula is equivalent to an existential formula (this property is equivalent to model completeness) and every definable set is the projection of an exponential variety.

Wilkie's proof depends heavily on the following special case of a theorem of Khovanski. Before Wilkie's theorem, Khovanski's result was the best evidence that $\mathbb{R}_{\text{exp}}$ is o-minimal; indeed Khovanski's theorem is also the crucial tool needed to deduce o-minimality from model completeness.

Theorem 3.31 (Khovanski) *If $f_1, \ldots, f_m : \mathbb{R}^n \to \mathbb{R}$ are exponential polynomials, then $\{x \in \mathbb{R}^n : f_1(x) = \ldots f_n(x) = 0\}$ has finitely many connected components.*

If $X \subseteq \mathbb{R}$ is definable in $\mathbb{R}_{\exp}$ then by Wilkie's Theorem there is an exponential variety $V \subseteq \mathbb{R}^n$ such that X is the projection of V. By Khovanski's Theorem V has finitely many connected components and X is a finite union of points and intervals. Thus $\mathbb{R}_{\exp}$ is o-minimal.

Using the o-minimality of $\mathbb{R}_{\exp}$ one can improve some of Khovanski's results on "fewnomials". From algebraic geometry we know that we can bound the number of connected components of a hypersurface in $\mathbb{R}^n$ uniformly in the degree of the defining polynomial. Khovanski showed that it is also possible to bound the number of connected component uniformly in the number of monomials in the defining polynomial. We will sketch the simplest case of this. Let $\mathcal{F}_{n,m}$ be the collection of polynomials in $\mathbb{R}[X_1, \ldots, X_n]$ with at most m monomials. For $p \in \mathcal{F}_{n,m}$ let

$$V^+(p) = \{\overline{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n : \bigwedge_{i=1}^{n} x_i \geq 0 \wedge p(\overline{x}) = 0\}.$$

We claim that there are only finitely many homeomorphism types of $V^+(p)$ for $p \in \mathcal{F}_{n,m}$. Let $\Phi_{m,n}(x_1, \ldots, x_n, r_{1,1} \ldots, r_{1,n}, \ldots, r_{m,1}, \ldots, r_{m,n}, a_1, \ldots, a_m)$ be the formula

$$\exists w_1 \ldots, w_m \ ((\bigwedge_{i=1^m} e^{w_i} = x_i) \wedge \sum_{i=1}^{m} a_i \prod_{j=1}^{n} e^{w_i r_{i,j}} = 0).$$

We see that Φ expresses

$$\sum_{i=1}^{m} a_i \prod_{j=1}^{n} x_j^{r_{i,j}} = 0.$$

Let $X_{\overline{r},\overline{a}}$ denote the set of $\overline{x} \in \mathbb{R}^n$ such that $\Phi(\overline{x}, \overline{r}, \overline{a})$ holds. By o-minimality, $\{X_{\overline{r},\overline{a}} : \overline{r} \in \mathbb{R}^{mn}, \overline{a} \in \mathbb{R}^m\}$ represents only finitely many homeomorphism types.

In addition to answering the question of o-minimality, some headway has been made on the problem of decidability. Making heavy use of Wilkie's methods and Khovanski's theorem, Macintyre and Wilkie have shown that if Schanuel's Conjecture in is true then the first order theory of $\mathbb{R}_{\exp}$ is decidable. Where Schanuel's Conjecture is the assertion that if $\lambda_1, \ldots, \lambda_n$ are complex numbers linearly independent over $\mathbb{Q}$, then the transcendence degree of the field

$$\mathbb{Q}(\lambda_1, \ldots, \lambda_n, e^{\lambda_1}, \ldots, e^{\lambda_n})$$

is at least n.

Miller provided an interesting counterpoint to Wilkie's theorem. Using ideas of Rosenlicht he showed that if $\mathcal{R}$ is any o-minimal expansion of the real field that contains a function that is not majorized by a polynomial, then exponentiation is definable in $\mathcal{R}$.

Let $\mathcal{L}_{\text{an,exp}}$ be $\mathcal{L}_{\text{an}} \cup \{e^x\}$ and let $\mathbb{R}_{\text{an,exp}}$ be the real numbers with both exponentiation and restricted analytic functions. Using the Denef-van den Dries

quantifier elimination for $\mathbb{R}_{\text{an}}$ and a mixture of model-theoretic and valuation theoretic ideas, van den Dries, Macintyre, and I were able to show that $\mathbb{R}_{\text{an,exp}}$ has quantifier elimination if we add log to the language. Using quantifier elimination and Hardy field style arguments (but avoiding the geometric type of arguments used by Khovanski) we were able to show that $\mathbb{R}_{\text{an,exp}}$ is o-minimal.

Since the language $\mathcal{L}_{\text{an,exp}}$ has size $2^{\aleph_0}$, one would not expect to give a simple axiomatization of the first order theory of $\mathbb{R}_{\text{an,exp}}$. Ressayre noticed that the model-theoretic analysis of $\mathbb{R}_{\text{an,exp}}$ uses very little global information about exponentiation. This observation leads to a "relative" axiomatization. The theory $Th(\mathbb{R}_{\text{an,exp}})$ is axiomatized by the theory of $\mathbb{R}_{\text{an}}$ and axioms asserting that exponentiation is an increasing homomorphism from the additive group onto the multiplicative group of positive elements that majorizes every polynomial.

Using this axiomatization and quantifier elimination one can show that any definable function is piecewise given by a composition of polynomials, exp, log, and restricted analytic functions on $[0,1]^n$. For example, the definable function $f(x) = e^{e^x} - e^{x^2} - 3x$ is eventually increasing and unbounded. Thus for some large enough $r \in \mathbb{R}$ there is a function $g : (r, +\infty) \to \mathbb{R}$ such that $f(g(x)) = x$ for $x > r$. The graph of g is the definable set $\{(x, y) : x > r \text{ and } e^{e^y} - e^{y^2} - 3y = x\}$. Thus g is a definable function and there is some way to express g explicitly as a composition of rational functions, exp, log, and restricted analytic functions. In most cases it is in no way clear how to get these explicit representations of an implicitly defined function. One important corollary is that every definable function is majorized by an iterated exponential.

A Real Algebra

We prove some of the algebraic facts needed in Section 3. All of these results are due to Artin and Schreier. See Lang's *Algebra* §XI for more details.

All fields are assumed to be of characteristic 0.

Definition A.1 A field K is *real* if -1 can not be expressed as a sum of squares of elements of K. In general, we let $\sum K^2$ be the sums of squares from K.

If F is orderable, then F is real because squares are nonnegative with respect to any ordering.

Lemma A.2 *Suppose that F is real and $a \in F \setminus \{0\}$. Then, at most one of a and $-a$ is a sum of squares.*

Proof If a and b are both sums of squares, then $\frac{a}{b} = \frac{a}{b^2}b$ is a sum of squares. Thus, if F is real, at least one of a and $-a$ is not in $\sum F^2$.

Lemma A.3 *If F is real and $-a \in F \setminus \sum F^2$, then $F(\sqrt{a})$ is real. Thus, if F is real and $a \in F$, then $F(\sqrt{a})$ is real or $F(\sqrt{-a})$ is real.*

Proof We may assume that $\sqrt{a} \notin F$. If $F(\sqrt{a})$ is not real, then there are $b_i, c_i \in F$ such that

$$-1 = \sum (b_i + c_i\sqrt{a})^2 = \sum (b_i^2 + 2c_i b_i \sqrt{a} + c_i^2 a).$$

Because $\sqrt{a}$ and 1 are a vector space basis for $F(\sqrt{a})$ over F,

$$-1 = \sum b_i^2 + a \sum c_i^2.$$

Thus

$$-a = \frac{1 + \sum b_i^2}{\sum c_i^2} = \frac{\left(\sum b_i^2\right)\left(\sum c_i^2\right) + \left(\sum c_i^2\right)}{\left(\sum c_i^2\right)^2}$$

and $-a \in \sum F^2$, a contradiction.

Lemma A.4 *If F is real, $f(X) \in F[X]$ is irreducible of odd degree n, and $f(\alpha) = 0$, then $F(\alpha)$ is real.*

Proof We proceed by induction on n. If $n = 1$, this is clear. Suppose, for purposes of contradiction, that $n > 1$ is odd, $f(X) \in F[X]$ is irreducible of degree n, $f(\alpha) = 0$, and $F(\alpha)$ is not real. There are polynomials g_i of degree at most $n-1$ such that $-1 = \sum g_i(\alpha)^2$. Because F is real, some g_i is nonconstant. Because $F(\alpha) \cong F[X]/(f)$, there is a polynomial $q(X) \in F[X]$ such that

$$1 = \sum g_i^2(X) + q(X)f(X).$$

The polynomial $\sum g_i^2(X)$ has a positive even degree at most $2n-2$. Thus, q has odd degree at most $n-2$. Let β be the root of an irreducible factor of q. By induction, $F(\beta)$ is real, but $-1 = \sum g_i^2(\beta)$, a contradiction.

Definition A.5 We say that a field R is *real closed* if and only if R is real and has no proper real algebraic extensions.

If R is real closed and $a \in R$, then, by Lemmas A.2 and A.3, either $a \in R^2$ or $-a \in R^2$. Thus, we can define an order on R by

$$a \geq 0 \Leftrightarrow a \in R^2.$$

Moreover, this is the only way to define an order on R because the squares must be nonnegative. Also, if R is real closed, every polynomial of odd degree has a root in R.

Lemma A.6 *Let F be a real field. There is $R \supseteq F$ a real closed algebraic extension. We call R a* real closure *of F.*

Proof Let $I = \{K \supseteq F : K \text{ real}, K/F \text{ algebraic}\}$. The union of any chain of real fields is real; thus, by Zorn's Lemma, there is a maximal $R \in I$. Clearly, R has no proper real algebraic extensions; thus, R is real closed.

Corollary A.7 *If F is any real field, then F is orderable. Indeed, if $a \in F$ and $-a \notin \sum F^2$, then there is an ordering of F, where $a > 0$.*

Proof By Lemma A.3, $F(\sqrt{a})$ is real. Let R be a real closure of F. We order F by restricting the ordering of R because a is a square in R, $a > 0$.

The following theorem is a version of the Fundamental Theorem of Algebra.

Theorem A.8 *Let R be a real field such that*
i) for all $a \in R$, either $\sqrt{a}$ or $\sqrt{-a} \in R$ and
ii) if $f(X) \in R[X]$ has odd degree, then f has a root in R.
If $i = \sqrt{-1}$, then $K = R(i)$ is algebraically closed.

Proof

Claim 1 Every element of K has a square root in K.

Let $a + bi \in K$. Note that $\frac{a+\sqrt{a^2+b^2}}{2}$ is nonnegative for any ordering of R. Thus, by i), there is $c \in R$ with

$$c^2 = \frac{a + \sqrt{a^2+b^2}}{2}.$$

If $d = \frac{b}{2c}$, then $(c+di)^2 = a + bi$.

Let $L \supseteq K$ be a finite Galois extension of R. We must show that $L = K$. Let $G = Gal(L/R)$ be the Galois group of L/R. Let H be the 2-Sylow subgroup of G.

Claim 2 $G = H$.

Let F be the fixed field of H. Then F/R must have odd degree. If $F = R(x)$, then the minimal polynomial of x over R has odd degree, but the only irreducible polynomials of odd degree are linear. Thus, $F = R$ and $G = H$.

Let $G_1 = Gal(L/K)$. If G_1 is nontrivial, then there is G_2 a subgroup of G_1 of index 2. Let F be the fixed field of G_2. Then, F/K has degree 2. But by Claim 1, K has no extensions of degree 2. Thus, G_1 is trivial and $L = K$.

Corollary A.9 *Suppose that R is real. Then R is real closed if and only if $R(i)$ is algebraically closed.*

Proof

($\Rightarrow$) By Theorem A.8.

($\Leftarrow$) $R(i)$ is the only algebraic extension of R, and it is not real.

Let $(R, <)$ be an ordered field. We say that R has the *intermediate value property* if for any polynomial $p(X) \in R[X]$ if $a < b$ and $p(a) < 0 < p(b)$, then there is $c \in (a, b)$ with $p(c) = 0$.

Lemma A.10 *If $(R, <)$ is an ordered field with the intermediate value property, then R is real closed.*

Proof Let $a > 0$ and let $p(X) = X^2 - a$. Then $p(0) < 0$, and $p(1 + a) > 0$; thus, there is $c \in R$ with $c^2 = a$.

Let

$$f(X) = X^n + \sum_{i=0}^{n-1} a_i X^i$$

where n is odd. For M large enough, $f(M) > 0$ and $f(-M) < 0$; thus, there is a c such that $f(c) = 0$.

By Theorem A.8, $R(i)$ is algebraically closed. Because R is real, it must be real closed.

Lemma A.11 *Suppose that R is real closed and $<$ is the unique ordering, then $(R, <)$ has the intermediate value property.*

Proof Suppose $f(X) \in R[X]$, $a < b$, and $f(a) < 0 < f(b)$. We may assume that $f(X)$ is irreducible (for some factor of f must change signs). Because $R(i)$ is algebraically closed, either $f(X)$ is linear, and hence has a root in (a, b), or

$$f(X) = X^2 + cX + d,$$

where $c^2 - 4d < 0$. But then

$$f(X) = \left(X + \frac{c}{2}\right)^2 + \left(d - \frac{c^2}{4}\right)$$

and $f(x) > 0$ for all x.

We summarize as follows.

Theorem A.12 *The following are equivalent.*

i) R is real closed.

ii) For all $a \in R$, either a or $-a$ has a square root in R and every polynomial of odd degree has a root in R.

iii) We can order R by $a \geq 0$ if and only if a is a square and, with respect to this ordering, R has the intermediate value property.

Finally, we consider the question of uniqueness of real closures. We first note that there are some subtleties. For example, there are nonisomorphic real closures of $F = \mathbf{Q}(\sqrt{2})$. The field of real algebraic numbers is one real closure of F. Because $a + b\sqrt{2} \mapsto a - b\sqrt{2}$ is an automorphism of F, $\sqrt{2}$ is not in $\sum F^2$. Thus, by Corollary B.5, $F(\sqrt{-2})$ is real. Let R be a real closure of F containing $F(\sqrt{-2})$. Then, R is not isomorphic to the real algebraic numbers over F.

This is an example of a more general phenomenon. It is proved by successive applications of Lemmas A.2 and A.3.

Lemma A.13 *If $(F, <)$ is an ordered field, then there is a real closure of F in which every positive element of F is a square.*

Because $\mathbb{Q}(\sqrt{2})$ has two distinct orderings, it has two nonisomorphic real closures. The field $\mathbb{Q}(t)$ of rational functions over $\mathbb{Q}$ has $2^{\aleph_0}$ orderings and hence $2^{\aleph_0}$ nonisomorphic real closures.

The next theorem shows that once we fix an ordering of F, there is a unique real closure that induces the ordering.

Theorem A.14 *Let $(F, <)$ be an ordered field. Let R_0 and R_1 be real closures of F such that $(R_i, <)$ is an ordered field extension of $(F, <)$. Then, R_0 is isomorphic to R_1 over F and the isomorphism is unique.*

The proof of Theorem A.14 uses Sturm's algorithm.

Definition A.15 Let R be a real closed field. A *Sturm sequence* is a finite sequence of polynomials $f_0, \ldots, f_n$ such that:

i) $f_1 = f_0'$;

ii) for all x and $0 \leq i \leq n-1$, it is not the case that $f_i(x) = f_{i+1}(x) = 0$;

iii) for all x and $1 \leq i \leq n-1$, if $f_i(x) = 0$, then $f_{i-1}(x)$ and $f_{i+1}(x)$ have opposite signs;

iv) f_n is a nonzero constant.

If $f_0, \ldots, f_n$ is a Sturm sequence and $x \in \mathbb{R}$, define $v(x)$ to be the number of sign changes in the sequence $f_0(x), \ldots, f_n(x)$.

Suppose that $f \in R[X]$ is nonconstant and does not have multiple roots. We define a Sturm sequence as follows:

$f_0 = f$;

$f_1 = f'$.

Given f_i nonconstant, use the Euclidean algorithm to write

$$f_i = g_i f_{i-1} - f_{i+1}$$

where the degree of f_{i+1} is less than the degree of f_{i-1}. We eventually reach a constant function f_n.

Lemma A.16 *If f has no multiple roots, then $f_0, \ldots, f_n$ is a Sturm sequence.*

Proof

iv) If $f_n = 0$, then $f_{n-1} | f_i$ for all i. But f has no multiple roots; thus f and f' have no common factors, a contradiction.

ii) If $f_i(x) = f_{i+1}(x) = 0$, then by induction $f_n(x) = 0$, contradicting iv).

iii) If $1 \leq i \leq n-1$ and $f_i(x) = 0$, then $f_{i-1}(x) = -f_{i+1}(x)$. Thus, $f_{i-1}(x)$ and $f_{i+1}(x)$ have opposite signs.

Theorem A.17 (Sturm's Algorithm) *Suppose that R is a real closed field, $a, b \in R$, and $a < b$. Let f be a polynomial without multiple roots. Let $f = f_0, \ldots, f_n$ be a Sturm sequence such that $f_i(a) \neq 0$ and $f_i(b) \neq 0$ for all i. Then, the number of roots of f in (a, b) is equal to $v(a) - v(b)$.*

Proof Let $z_1 < \ldots < z_m$ be all the roots of the polynomials $f_0, \ldots, f_n$ that are in the interval (a, b). Choose $c_1, \ldots, c_{m-1}$ with $z_i < c_i < z_{i+1}$. Let $a = c_0$ and $b = c_m$. For $0 \leq i \leq m-1$, let r_i be the number of roots of f in the interval (c_i, c_{i+1}). Clearly, $\sum r_i$ is the number of roots of f in the interval (a, b). On the other hand,

$$v(a) - v(b) = \sum_{i=0}^{m-1} (v(c_i) - v(c_{i+1})).$$

Thus, it suffices to show that if $c < z < d$ and z is the only root of any f_i in (c, d), then

$$v(d) = \begin{cases} v(c) - 1 & z \text{ is a root of } f \\ v(c) & \text{otherwise} \end{cases}.$$

If $f_i(b)$ and $f_i(c)$ have different signs, then $f_i(z) = 0$. We need only see what happens at those places.

If z is a root of f_i, $i > 0$, then $f_{i+1}(z)$ and $f_{i-1}(z)$ have opposite signs and f_{i+1} and f_{i-1} do not change signs on $[c, d]$. Thus, the sequences $f_{i-1}(c), f_i(c), f_{i+1}(c)$ and $f_{i-1}(d), f_i(d), f_{i+1}(d)$ each have one sign change. For example, if $f_{i-1}(z) > 0$ and $f_{i-1}(z) < 0$, then these sequences are either $+, +, -$ or $+, -, +$, and in either case both sequences have one sign change.

If z is a root of f_0, then, because $f'(z) \neq 0$, f is monotonic on (c, d). If f is increasing on (c, d), the sequence at c starts $-, +, \ldots$ and the sequence at d starts $+, +, \ldots$. Similarly, if f is decreasing, the sequence at c starts $+, -, \ldots$, and the sequence at b starts $-, -, \ldots$. In either case, the sequence at c has one more sign change than the sequence at d. Thus, $v(c) - v(d) = 1$, as desired.

Corollary A.18 *Suppose that $(F, <)$ is an ordered field. Let f be a nonconstant irreducible polynomial over F. If R_0 and R_1 are real closures of F compatible with the ordering, then f has the same number of roots in both R_0 and R_1.*

Proof Let $f_0, \ldots, f_n$ be the Sturm sequence from Lemma A.16. Note that each $f_i \in F[X]$. We can find $M \in F$ such that any root of f_i is in $(-M, M)$ (if $g(X) = X^n + \sum a_i X^i$, then any root of g has absolute value at most $1 + \sum |a_i|$, for example). Then, the number of roots of f in R_i is equal to $v(-M) - v(M)$, but $v(M)$ depends only on F.

Lemma A.19 *Suppose $(F, <)$ is an ordered field and R_0 and R_1 are real closures of F such that $(R_i, <)$ is an ordered field extension of $(F, <)$. If $\alpha \in R_0 \setminus F$, there is an ordered field embedding of $F(\alpha)$ into R_1 fixing F.*

Proof Let $f \in F[X]$ be the minimal polynomial of α over F. Let $\alpha_1 < \ldots < \alpha_n$ be all zeros of f in R_0. By Corollary B.18, f has exactly n zeros $\beta_1 < \ldots < \beta_n \in R_1$. Let

$$\sigma : F(\alpha_1, \ldots, \alpha_n) \to F(\beta_1, \ldots, \beta_n)$$

be the map obtained by sending α_i to β_i. We claim that σ is an ordered field isomorphism.

For $i = 1, \ldots, n-1$, let $\gamma_i = \sqrt{\alpha_{i+1} - \alpha_i} \in R_0$. By the Primitive Element Theorem, there is $a \in F$ such that

$$F(a) = F(\alpha_1, \ldots, \alpha_n, \gamma_1, \ldots, \gamma_{n-1}).$$

Let $g \in F[X]$ be the minimal polynomial of a over F. By Corollary B.18, g has a zero $b \in R_1$ and there is a field isomorphism $\phi : F(a) \to F(b)$. Because $F(a)$ contains n zeros of F, so does $F(b)$. Thus $\beta_1, \ldots, \beta_n \in F(b)$ and for each i there is a j such that $\phi(\alpha_i) = \beta_j$. But

$$\phi(\gamma_i)^2 = \phi(\alpha_{i+1}) - \phi(\alpha_i).$$

Thus $\phi(\alpha_i) = \beta_i$ for $i = 1, \ldots, n$. We still must show that σ is order preserving. Suppose $c \in F(\alpha_1, \ldots, \alpha_n)$ and $c > 0$. There is $d \in R_0$ such that $d^2 = c$. Arguing as above, we can find a field embedding

$$\psi : F(\alpha_1, \ldots, \alpha_n, d) \subseteq R_1$$

fixing F. As above, $\psi(\alpha_i) = \beta_i$ and $\psi \supseteq \sigma$. Because

$$\psi(d)^2 = \psi(c) = \sigma(c),$$

we have $\sigma(c) > 0$. Thus σ is order preserving.

Proof of Theorem A.14 Let $\mathcal{P}$ be the set of all order preserving $\sigma : K \to R_1$ where $F \subseteq K \to R_0$ and $\sigma | F$ is the identity. By Zorn's Lemma, there is a maximal $\sigma : K \to R_1$ in $\mathcal{P}$. By identifying K and $\sigma(K)$ and applying the previous lemma, we see that $K = R_0$. A similar argument shows that $\sigma(K) = R_1$.

Uniqueness follows because the ith root of $f(X)$ in R_0 must be sent to the ith root of $f(X)$ in R_1.