Math 509 Lecture Notes: Model theory of the real numbers

Alex Kruckman

Final version: May 4, 2020

Contents

These are notes from a course taught at Wesleyan University in Spring 2020. Thanks to my students for feedback and corrections. Section 3 follows the general outline of real algebra from Marker's *Model Theory: An Introduction*, but with many proofs and details filled in. Section 4 follows closely the exposition and proofs in van den Dries's *Tame Topology and O-minimal Structures*.

1 Preliminaries: The language of model theory

1.1 Languages and structures

Definition 1.1. A first-order **language** $\mathcal{L}$ is a set of **function symbols** and **relation symbols**. Each symbol $s \in \mathcal{L}$ comes with an **arity** $\mathrm{ar}(s) \in \mathbb{N}$. A function symbol of arity 0 is called a **constant symbol**, and a relation symbol of arity 0 is called a **proposition symbol**. The words **unary**, **binary**, and **ternary** mean arity 1, 2, and 3, respectively, and we also write n**-ary** to mean arity n.

We are intentionally vague about what counts as a symbol; the name symbol is meant to suggest something that you could write down with a pencil on paper, but we have no intention of formalizing this notion. In practice, real-world symbols on paper can be encoded as mathematical objects (e.g. sets) in any way you like, and a symbol can be any mathematical object. In particular, a vocabulary may be uncountably infinite.

Example 1.2. The language of rings, $\mathcal{L}_r$, is $\{0, 1, +, -, \cdot\}$, where 0 and 1 are constant symbols, $+$ and $\cdot$ are binary function symbols, and $-$ is a unary function symbol for the additive inverse operation. The language of ordered rings, $\mathcal{L}_{or}$, is $\mathcal{L}_r \cup \{\leq\}$, where $\leq$ is a binary relation symbol.

In all further discussions, we always have in the background a language $\mathcal{L}$.

Definition 1.3. A **structure** is a set A, equipped with:

- For each function symbol $f \in \mathcal{L}$ with $\mathrm{ar}(f) = n$, a function $f^A \colon A^n \to A$. In the case of a constant symbol c, we have $c^A \colon A^0 \to A$. The set A^0 is a singleton $\{()\}$, and we identify c^A with the element $c^A(()) \in A$.
- For each relation symbol $R \in \mathcal{L}$ with $\mathrm{ar}(R) = n$, a relation $R^A \subseteq A^n$. In the case of a proposition symbol P, we have $P^A \subseteq A^0 = \{()\}$, and P^A is either $\{()\}$ ("true") or $\emptyset$ ("false").

Contrary to the common convention, we do not require structures to be non-empty in general. But note that if $\mathcal{L}$ contains any constant symbols, then any structure will be non-empty.

If $\mathcal{L}$ is listed as $(f_1, \ldots, f_n, R_1, \ldots, R_m)$, then we often use the following notation to describe a structure:

$$(A; f_1^A, \ldots, f_n^A, R_1^A, \ldots, R_m^A).$$

Example 1.4. Any ring, and in particular any field, is a structure in the language $\mathcal{L}_r$ in a natural way. It is tempting to include a unary function symbol $^{-1}$ in the language when we discuss a field K, but this would present some annoyances: the interpretation of $^{-1}$ would have to be a total function $K \to K$, but 0 has no multiplicative inverse. So we typically use the language of rings when discussing fields; we will see later that the multiplicative inverse function is a definable function in this context.

Note that there are many $\mathcal{L}_r$-structures which are not rings. Indeed, in an $\mathcal{L}_r$-structure, the symbols can be interpreted as arbitrary functions and relations. We will need to impose axioms to restrict ourselves to natural classes of structures.

1.2 Terms and evaluation

Definition 1.5. A **variable context** is a finite tuple $\overline{x} = (x_1, \ldots, x_k)$ of **variables**, with no repetitions. This includes the empty variable context ().

Just like with symbols, we will be intentionally vague about what counts as a variable. The important thing is that there are infinitely many of them, so we can introduce new variables to any context. We will abuse notation freely when concatenating variable contexts. For example, when $\overline{x} = (x_1, \ldots, x_k)$ is a variable context and y is another variable not in $\overline{x}$, we write $\overline{x}y$ for the variable context $(x_1, \ldots, x_k, y)$.

Definition 1.6. A **term** in context $\overline{x} = (x_1, \ldots, x_k)$ is one of the following:

- A variable x_i from $\overline{x}$.
- A constant symbol c in $\mathcal{L}$.
- A composite term $f(t_1, \ldots, t_n)$, where $f \in \mathcal{L}$ is a function symbol of arity n and t_i is a term of type in context $\overline{x}$, for all $1 \leq i \leq n$.

Note that the case of constant symbols is really a special case of the case of composite terms, when $n = 0$.

This is a recursive definition, so we obtain a corresponding method of proof by induction. To prove a claim about all terms in context $\overline{x}$, it suffices to check the base case (the claim holds for all variables in $\overline{x}$), and the inductive step (given that the claim holds for the terms $t_1, \ldots, t_n$, it holds for the composite term $f(t_1, \ldots, t_n)$). Sometimes it is useful to handle the constant symbols as a separate base case.

Example 1.7. In the language $\mathcal{L}_{or}$ of ordered rings, the following are terms in context (x, y):

$$x, \quad 0, \quad (x + 0) \cdot (-y), \quad ((x \cdot x) \cdot x) \cdot x$$

Note that we use the natural notation for our symbols when they differ from the formal syntax described above, for example writing $(x+0)$ instead of $+(x, 0)$. We use parentheses freely to avoid ambiguity. We will often want to abbreviate terms in natural ways, e.g. by writing $((x \cdot x) \cdot x) \cdot x$ as x^4. But for this to make sense, we need to be working in a context where associativity of $\cdot$ is assumed.

Definition 1.8. Given a variable context $\overline{x} = (x_1, \ldots, x_k)$ and a structure A, an **interpretation** of $\overline{x}$ in A is a tuple $\overline{a} = (a_1, \ldots, a_k) \in A^k$. We say the variable x_i is interpreted as the element a_i.

Note that there is a unique interpretation of the empty context (), namely the empty tuple $() \in A^0$.

Definition 1.9. Let A be a structure, let t be a term in context $\overline{x} = (x_1, \ldots, x_k)$, and let $\overline{a} = (a_1, \ldots, a_k)$ be an interpretation of $\overline{x}$ in A. Then we define the **evaluation** of t at $\overline{a}$, written $t^A(\overline{a})$, by induction on the complexity of t:

- If t is x_i, then $t^A(\overline{a}) = a_i$.
- If t is c, a constant symbol, then $t^A(\overline{a}) = c^A$.
- If t is $f(t_1, \ldots, t_n)$, a composite term, then we have elements $t_i^A(\overline{a}) \in A$ by induction for all $1 \leq i \leq n$. We define $t^A(\overline{a}) = f^A(t_1^A(\overline{a}), \ldots, t_n^A(\overline{a}))$.

Instead of fixing the interpretation $\overline{a}$ and letting the term t vary, we can fix the term t and let the interpretation $\overline{a}$ vary. Then t determines a function $t^A \colon A^k \to A$, by $\overline{a} \mapsto t^A(\overline{a})$.

We often write $t(\overline{x})$ to denote that the term t is in context $\overline{x}$. A term always comes with an associated context, even if it is not explicit in the notation. If t is a term in context $\overline{x}$, and y is a variable not in $\overline{x}$, then we can also view t as a term in context $\overline{x}y$. Indeed, the context just restricts which variables can be mentioned in t. Note that if $t(\overline{x})$ is a term in context $\overline{x}$, $t(\overline{x}, y)$ is the same term in context $\overline{x}y$, $\overline{a}$ is an interpretation of $\overline{x}$ in A, and b is any interpretation of y in A, then $t^A(\overline{a}) = t^A(\overline{a}, b)$. But the domains of the functions defined by these terms are different cartesian powers of A.

1.3 Formulas and satisfaction

Definition 1.10. An **atomic formula** in context $\overline{x}$ is one of the following:

- $(t_1 = t_2)$, where t_1 and t_2 are terms in context $\overline{x}$.
- $R(t_1, \ldots, t_n)$, where $R \in \mathcal{L}$ is a relation symbol of arity n and t_i is a term in context $\overline{x}$, for all $1 \leq i \leq n$.

Definition 1.11. A **formula** in context $\overline{x}$ is one of the following:

- An atomic formula in context $\overline{x}$.
- $\top$ or $\bot$.
- $(\psi \wedge \chi)$, $(\psi \vee \chi)$, or $\neg\psi$, where ψ and χ are formulas in context $\overline{x}$.
- $\exists y\, \psi$ or $\forall y\, \psi$, where ψ is a formula in context $\overline{x}y$ and y is a variable not in $\overline{x}$.

This is a definition by recursion (simultaneously across all contexts $\overline{x}$), so we obtain a corresponding method of proof by induction. To prove a claim about all formulas, it suffices to check the base cases (the claim holds for all atomic formulas, $\top$, and $\bot$), and the inductive steps (given that the claim holds for the formulas ψ and χ, it holds for the formulas $(\psi \wedge \chi)$, $(\psi \vee \chi)$, $\neg\psi$, $\exists y\, \psi$, and $\forall y\, \psi$).

Example 1.12. In the language $\mathcal{L}_{or}$ of ordered rings, the following are formulas:

$$\text{Context } (x,y,z)\colon ((x\cdot x)+(x+x))+1=0, \quad x+z\leq y, \quad \exists w\,(x\cdot w=1)$$
$$\text{Empty context } ()\colon \neg(0=1), \quad \forall x\,(x=0\vee\exists w\,(x\cdot w=1))$$

Note that we use the natural notation for our symbols when they differ from the formal syntax described above, for example writing $x\leq y$ instead of $\leq(x,y)$. In the context of rings, we could rewrite the first example as $x^2+2x+1=0$.

We will also employ the following standard shorthands:

- $(t_1\neq t_2)$ is shorthand for $\neg(t_1=t_2)$.
- $t_1<t_2$ is shorthand for $(t_1\leq t_2)\wedge(t_1\neq t_2)$, when the relation symbol $\leq$ is in the language.
- $(\psi\to\chi)$ is shorthand for $(\neg\psi\vee\chi)$.
- $(\psi\leftrightarrow\chi)$ is shorthand for $((\psi\to\chi)\wedge(\chi\to\psi))$.
- $\bigwedge_{i=1}^n\varphi_i$ and $\bigvee_{i=1}^n\varphi_i$ are shorthand for $(\ldots((\varphi_1\wedge\varphi_2)\wedge\varphi_3)\cdots\wedge\varphi_n)$ and $(\ldots((\varphi_1\vee\varphi_2)\vee\varphi_3)\cdots\vee\varphi_n)$, respectively. In the case $n=0$, the empty conjunction is $\top$ and the empty disjunction is $\bot$.

Definition 1.13. Let A be a structure, let φ be a formula in context $\overline{x}$, and let $\overline{a}$ be an interpretation of $\overline{x}$ in A. We define the relation $A\models\varphi(\overline{a})$, read A **satisfies** $\varphi(\overline{a})$ or φ is **true** of $\overline{a}$ in A, by induction on the complexity of φ:

- If φ is $(t_1=t_2)$, then $A\models\varphi(\overline{a})$ iff $t_1^A(\overline{a})=t_2^A(\overline{a})$.
- If φ is $R(t_1,\ldots,t_n)$, then $A\models\varphi(\overline{a})$ iff $(t_1^A(\overline{a}),\ldots,t_n^A(\overline{a}))\in R^A$.
- If φ is $\top$, then $A\models\varphi(\overline{a})$.
- If φ is $\bot$, then $A\not\models\varphi(\overline{a})$.
- If φ is $(\psi\wedge\chi)$, then $A\models\varphi(\overline{a})$ iff $A\models\psi(\overline{a})$ and $A\models\chi(\overline{a})$.
- If φ is $(\psi\vee\chi)$, then $A\models\varphi(\overline{a})$ iff $A\models\psi(\overline{a})$ or $A\models\chi(\overline{a})$.
- If φ is $\neg\psi$, then $A\models\varphi(\overline{a})$ iff $A\not\models\psi(\overline{a})$.
- If φ is $\exists y\,\psi$, then $A\models\varphi(\overline{a})$ iff there exists $b\in A$ such that $A\models\psi(\overline{a},b)$.
- If φ is $\forall y\,\psi$, then $A\models\varphi(\overline{a})$ iff for all $b\in A$, $A\models\psi(\overline{a},b)$.

Definition 1.14. Let φ be a formula in context $\overline{x}$. We write

$$\varphi(A)=\{\overline{a}\in A^n\mid A\models\varphi(\overline{a})\}$$

and call this a **definable set**.

If ψ is a formula in context $\overline{xy}=(x_1,\ldots,x_n,y_1,\ldots,y_m)$ and $\overline{b}\in A^m$, we write

$$\psi(A,\overline{b})=\{\overline{a}\in A^n\mid A\models\varphi(\overline{a},\overline{b})\}$$

and call this a **definable set with parameters** $\overline{b}$.

The inductive definition of satisfaction shows that the class of definable sets:

- Contains $\Delta = \{(a, a) \mid a \in A\} \subseteq A^2$ and $R^A \subseteq A^{\mathrm{ar}(R)}$ and all their preimages under term functions $(t_1^A, \ldots, t_k^A)\colon A^n \to A^k$ for all n and all k (definable by atomic formulas)
- Contains A^n and $\emptyset \subseteq A^n$ for all n (definable by $\top$ and $\bot$).
- Is closed under $\cap$, $\cup$, and relative complement in A^n for all n (defined by Boolean combinations).
- Is closed under coordinate projection and coprojection $A^{n+1} \to A^n$ for all n (defined by quantifiers). Here "coprojection" can be defined as the complement of the projection of the complement.

The class of sets definable with parameters is additionally closed under slices (fibers of coordinate projection maps).

We often write $\varphi(\overline{x})$ to denote that the formula φ is in context $\overline{x}$. A formula always comes with an associated context, even if it is not explicit in the notation. If φ is a formula in context $\overline{x}$, and y is a variable which does not appear anywhere in φ, then we can also view φ as a formula in context $\overline{x}y$. Note that if $\varphi(\overline{x})$ is a formula in context $\overline{x}$, $\varphi(\overline{x}, y)$ is the same formula in context $\overline{x}y$, $\overline{a}$ is an interpretation of $\overline{x}$ in A, and b is any interpretation of y in A, then $A \models \varphi(\overline{a})$ if and only if $A \models \varphi(\overline{a}, b)$. But the definable sets defined by these formulas are subsets of different cartesian powers of A.

Example 1.15. Consider the language $\mathcal{L} = \mathcal{L}_{or} \cup \{f\}$, where f is a unary function symbol. Consider $\mathbb{R}$ as an $\mathcal{L}$-structure, with the symbols in $\mathcal{L}_{or}$ interpreted in the standard ways, and with f interpreted as an arbitrary function $f^{\mathbb{R}}\colon \mathbb{R} \to \mathbb{R}$. Then the set $C(f) = \{a \in \mathbb{R} \mid f \text{ is continuous at } a\}$ is definable by the following formula in context x:

$$\forall \varepsilon \left((\varepsilon > 0) \to \exists \delta \left((\delta > 0) \wedge \forall z \left((x - \delta < z < x + \delta) \to (f(x) - \varepsilon < f(z) < f(x) + \varepsilon)\right)\right)\right).$$

Exercise 1. Show that $D(f) = \{a \in \mathbb{R} \mid f \text{ is differentiable at } a\}$ is definable.

The previous example shows that first-order formulas can be rather complicated and quite expressive. However, they are also limited. Crucially, they are finite, and quantifiers range only over elements of a structure, not over subsets, or functions, or natural numbers, etc.

Example 1.16. In $(\mathbb{R}; 0, 1, +, -, \cdot)$, every integer is definable, in the sense that $\{n\}$ is a definable set. Indeed, consider the formulas

$$x = 0 \quad \text{or} \quad x = \underbrace{1 + \cdots + 1}_{n \text{ times}} \quad \text{or} \quad x = -(\underbrace{1 + \cdots + 1}_{n \text{ times}}).$$

It follows that every finite set of integers is definable, by combining the above formulas with $\vee$. But it turns out that $\mathbb{Z} \subseteq \mathbb{R}$ is not definable! We will develop

tools for understanding definable sets and proving non-definability results like this.

Similarly, every rational number is definable, for example by

$$(\underbrace{1+\cdots+1}_{m \text{ times}}) \cdot x = (\underbrace{1+\cdots+1}_{n \text{ times}}),$$

but $\mathbb{Q} \subseteq \mathbb{R}$ is not definable.

1.4 Theories and models

Definition 1.17. A **sentence** is a formula in the empty variable context.

A sentence φ has no free variables: all variables in φ are bound by quantifiers. When A is a structure, there is a unique interpretation $() \in A^0$ of the empty context in A, so the satisfaction of a sentence φ does not depend on a choice of interpretation of variables. A sentence expresses a property of A, rather than a property of tuples from A. To put it another way, a sentence defines a subset of A^0, which is either $\{()\}$ ("true") or $\emptyset$ ("false"). We write $A \models \varphi$ or $A \not\models \varphi$, instead of $A \models \varphi(())$ or $A \not\models \varphi(())$.

Definition 1.18. A **theory** T is a set of sentences.

A structure M is a **model** of T, written $M \models T$, if $M \models \varphi$ for all $\varphi \in T$.

If φ is a sentence, the T **entails** φ, written $T \models \varphi$, if every model of T satisfies φ.

Example 1.19. The language of groups is $\mathcal{L}_g = \{\cdot, e, ^{-1}\}$, where $\cdot$ is a binary function symbol, e is a constant symbol, and $^{-1}$ is a unary function symbol. The theory of groups T_g consists of the following three sentences:

$$\forall x \forall y \forall z \, ((x \cdot y) \cdot z = x \cdot (y \cdot z))$$
$$\forall x \, ((x \cdot e = x) \wedge (e \cdot x = x))$$
$$\forall x \, ((x \cdot x^{-1} = e) \wedge (x^{-1} \cdot x = e))$$

Of course, an $\mathcal{L}_g$-structure G is a group if and only if $G \models T_g$.

We have

$$T_g \models (\forall x \, (x \cdot x = e) \to \forall x \forall y \, (x \cdot y = y \cdot x)),$$

since all groups of exponent 2 are abelian.

On the other hand, we have

$$T_g \not\models \forall x \forall y \, (x \cdot y = y \cdot x),$$

since there exists a group which is not abelian.

Definition 1.20. A theory T is **satisfiable** if T has a model.

A theory T is **complete** if it is satisfiable, and for every sentence φ, either $T \models \varphi$ or $T \models \neg\varphi$.

Note that T is satisfiable if and only if $T \not\models \bot$. Indeed, if T has a model M, then $M \not\models \bot$, so $T \not\models \bot$. On the other hand, if T has no models, then $T \models \bot$ vacuously. In fact, if T is not satisfiable, then for any sentence φ, $T \models \varphi$ vacuously.

Example 1.21. Let A be any structure. Then the **complete theory of** A is

$$\mathrm{Th}(A) = \{\varphi \mid A \models \varphi\}.$$

The name is justified, since for any sentence φ, either $A \models \varphi$ or $A \not\models \varphi$. In the first case, $\varphi \in \mathrm{Th}(A)$, while in the second case $A \models \neg\varphi$, so $\neg\varphi \in \mathrm{Th}(A)$.

Definition 1.22. Structure A and B are **elementarily equivalent**, written $A \equiv B$, if $\mathrm{Th}(A) = \mathrm{Th}(B)$, i.e., for all sentences φ, $A \models \varphi$ if and only if $B \models \varphi$.

1.5 Completeness and compactness

Give a structure M, it can be hard to understand the complete theory $\mathrm{Th}(M)$, and hence hard to understand the relation of elementarily equivalence between structures. For example, $\mathrm{Th}(\mathbb{N}; 0, 1, +, \times)$ contains answers to many open problems in number theory.

It can also be hard to understand the entailment relation $T \models \varphi$, since the definition of entailment quantifies over all models of T. Rather than considering all models, there is an easier way to show that T entails φ, namely to provide a proof.

Indeed, there is a system of formal proofs for first-order logic. We write $T \vdash \varphi$ if there exists a proof of φ from T, which consists of a finite list of formulas, ending with φ, such that for each formula in the list, it is either a sentence in T, or it follows from the previous formulas by some proof rule. The proof rules formalize basic steps of reasoning used in mathematical practice. Beyond this, the details of the proof system will not be important to us. But we will take for granted the following theorem, which connects proofs to entailment.

Theorem 1.23 (Gödel's Completeness Theorem)**.** *Let T be a theory, and let φ be a sentence. Then $T \models \varphi$ if and only if $T \vdash \varphi$.*

The implication from $T \vdash \varphi$ to $T \models \varphi$ is called **soundness**: it says that the proof rules don't do anything wrong. The reverse implication from $T \models \varphi$ to $T \vdash \varphi$ is called **completeness**: it says that the proof rules are strong enough to capture all entailments. The latter is harder to prove; the strategy is to assume that $T \nvdash \varphi$ and then to construct a model $M \models T$ such that $M \not\models \varphi$, witnessing $T \not\models \varphi$.

Going forward, we will not talk about formal proofs or the relation $\vdash$. The relevance of Gödel's Theorem for us is that the main tool in model theory, the compactness theorem, is a direct consequence.

Corollary 1.24 (Compactness Theorem)**.** *Suppose $T \models \varphi$. Then there is a finite subset $T' \subseteq_{fin} T$ such that $T' \models \varphi$.*

Proof. Suppose $T \models \varphi$. By completeness, $T \vdash \varphi$. But proofs are finite, so only finitely many sentences from T can be used in the proof. Thus there is a finite subset $T' \subseteq_{\text{fin}} T$ such that $T' \vdash \varphi$, and by soundness, $T' \models \varphi$. □

The compactness theorem is most often used in the following form, to construct models.

Corollary 1.25 (Compactness, Version 2)**.** *If every finite subset of a theory T is satisfiable, then T is satisfiable.*

Proof. Suppose T is not satisfiable. Then $T \models \bot$, so there is some finite subset $T' \subseteq_{\text{fin}} T$ such that $T' \models \bot$. But then T' is not satisfiable. □

Example 1.26. Let $T = \text{Th}(\mathbb{R}; 0, 1, +, -, \times, \leq)$. This is a complete theory in $\mathcal{L}_{or}$. Let $\mathcal{L}' = \mathcal{L}_{or} \cup \{\varepsilon\}$, where ε is a new constant symbol, and let

$$T' = T \cup \{0 < \varepsilon\} \cup \{\varepsilon \leq 1/n \mid n \in \mathbb{N}_+\}$$

Formally, we should express $\varepsilon \leq 1/n$ by $\underbrace{\varepsilon + \cdots + \varepsilon}_{n \text{ times}} \leq 1$.

Any finite subset of T' is contained in $T_N = T \cup \{0 < \varepsilon\} \cup \{\varepsilon \leq 1/n \mid n \leq N\}$ for some $N \in \mathbb{N}$, and T_N has a model: Expand $(\mathbb{R}; 0, 1, +, -, \times, \leq)$ by interpreting ε as $1/N$. So by compactness, T' has a model $\mathcal{R}'$.

The structure $\mathcal{R}'$ is elementarily equivalent to $\mathbb{R}$ in $\mathcal{L}_{or}$, but it has a positive element ε which is infinitesimal in the sense that it is less than every rational number. Since

$$T \models \forall x\, (0 < x \wedge x \leq 1/n \to \exists y\, (x \cdot y = 1 \wedge y \geq n)),$$

it follows that ε has an inverse ε^{-1} which is infinite, in the sense that it is greater than every natural number.

Example 1.27. Suppose T is a theory with an infinite model M, and let κ be any infinite cardinal. Then T has a model of cardinality $\geq \kappa$. This is part of the Löwenheim–Skolem theorem.

Let $\mathcal{L}' = \mathcal{L} \cup \{c_\alpha \mid \alpha \in \kappa\}$ where the c_α are κ-many new constant symbols. Let $T' = T \cup \{c_\alpha \neq c_\beta \mid \alpha \neq \beta\}$. A finite subset $\Delta \subseteq T'$ only asserts that finitely many of the new constants are distinct. So Δ has a model: expand M by interpreting the finitely many constants named in Δ as distinct elements, and interpret the rest of the constants arbitrarily. By compactness, T' has a model M'. Then $M' \models T$, and $|M'| \geq \kappa$, since M' contains κ-many distinct elements interpreting neq the constant symbols.

The last example shows that no first-order theory can pin down an infinite model uniquely up to isomorphism: there will always be a bigger model. This is a weakness of first-order logic, but the wealth of models is part of what makes model theory interesting.

2 Preservation and quantifier elimination

2.1 Maps between structures

Definition 2.1. If A and B are structures, a **homomorphism** $h\colon A \to B$ is a function such that:

- For every function symbol $f \in \mathcal{L}$ with $\mathrm{ar}(f) = n$, and for every tuple $\overline{a} \in A^n$, $h(f^A(a_1, \ldots, a_n)) = f^B(h(a_1), \ldots, h(a_n))$.
- For every relation symbol $R \in \mathcal{L}$ with $\mathrm{ar}(R) = n$, and for every tuple $\overline{a} \in A^n$, if $(a_1, \ldots, a_n) \in R^A$, then $(h(a_1), \ldots, h(a_n)) \in R^B$. We say h **preserves** R.

Definition 2.2. An **embedding** is a homomorphism $h\colon A \to B$ such that additionally:

- h is injective.
- For every relation symbol $R \in \mathcal{L}$ with $\mathrm{ar}(R) = n$, and for every tuple $\overline{a} \in A^n$, if $(h(a_1), \ldots, h(a_n)) \in R^B$, then $(a_1, \ldots, a_n) \in R^A$. We say h **reflects** R.

Definition 2.3. An **isomorphism** is a homomorphism $h\colon A \to B$ such that there exists an inverse homomorphism $h^{-1}\colon B \to A$. Equivalently, h is a surjective embedding. We write $A \cong B$ when A and B are isomorphic.

Suppose A and B are structures and $A \subseteq B$. Then A is a **substructure** of B if the inclusion map $A \to B$ is an embedding. That is, $f^A(\overline{a}) = f^B(\overline{a})$ for all function symbols $f \in \mathcal{L}$ of arity n and all $\overline{a} \in A^n$, and $\overline{a} \in R^A$ if and only if $\overline{a} \in R^B$ for all relation symbols $R \in \mathcal{L}$ of arity n and all $\overline{a} \in A^n$.

Suppose X is a subset of a structure AA. We say that X is **closed** if it is closed under the functions f^A for all function symbols $f \in \mathcal{L}$: if $\mathrm{ar}(f) = n$, then for all $\overline{a} \in X^n$, we have $f^A(\overline{a}) \in X$.

If X is closed, then there is a unique substructure of A with domain X, called the **induced substructure** on X, defined by:

$$f^X(\overline{a}) = f^A(\overline{a}), \text{ for all function symbols } f \in \mathcal{L} \text{ of arity } n \text{ and all } \overline{a} \in X^n.$$
$$R^X(\overline{a}) \iff R^A(\overline{a}), \text{ for all relation symbols } R \in \mathcal{L} \text{ of arity } n \text{ and all } \overline{a} \in X^n.$$

If X is an arbitrary subset of a structure A, then there is a smallest substructure of A containing X, namely the induced substructure on the closure of X. This is called the **substructure generated by** X and denoted $\langle X \rangle$.

Exercise 2. Show that the underlying set of $\langle X \rangle$ is

$$\{t^A(\overline{a}) \mid t \text{ is a term, and } \overline{a} \in X^n\}.$$

More precisely, show that this is the smallest closed subset of A containing X.

Exercise 3. Show that the intersection of a family of substructures of A is also a substructure of A, so that we can also $\langle X \rangle$ as the intersection of the family of all substructures of A containing X.

Homomorphisms and embeddings are particular kinds of structure-preserving maps between structures. A natural question is: which properties are preserved by which kinds of maps?

Definition 2.4. Let A and B be structures and $X \subseteq A$. A function $h\colon X \to B$ is called a **partial function** $A \dashrightarrow B$. For any formula φ, we say that a partial function $h\colon X \to B$ **preserves** φ if for any $\overline{a} \in X^n$,

$$\text{if } A \models \varphi(a_1, \ldots, a_n), \text{then } B \models \varphi(h(a_1), \ldots, h(a_n)).$$

And we say that h **reflects** φ if for any $\overline{a} \in X^n$,

$$\text{if } B \models \varphi(h(a_1), \ldots, h(a_n)), \text{then } A \models \varphi(a_1, \ldots, a_n).$$

Remark 2.5. A partial function h preserves φ if and only if it reflects $\neg\varphi$.

To simplify notation, when $\overline{a} = (a_1, \ldots, a_n)$ is a tuple, we write $h(\overline{a})$ for the tuple $(h(a_1), \ldots, h(a_n))$.

Lemma 2.6. *Let $h\colon A \to B$ be a homomorphism. Then for any term $t(\overline{x})$ and any interpretation $\overline{a}$ of $\overline{x}$, we have*

$$h(t^A(\overline{a})) = t^B(h(\overline{a})).$$

Proof. By induction on the complexity of terms. If t is a variable x_i, then $h(t^A(\overline{a})) = h(a_i) = t^B(h(\overline{a}))$.

If t is a composite term $f(t_1, \ldots, t_k)$, then

$$\begin{aligned} h(t^A(\overline{a})) &= h(f^A(t_1^A(\overline{a}), \ldots, t_k^A(\overline{a}))) \\ &= f^B(h(t_1^A(\overline{a})), \ldots, h(t_k^A(\overline{a}))) \\ &= f^B(t_1^B(h(\overline{a})), \ldots, t_k^B(h(\overline{a}))) \\ &= t^B(h(\overline{a})). \end{aligned}$$

□

Think of the next proposition as a generalization of the conditions for defining a homomorphism of groups $G \to H$ by defining a function on the generators of G. To be well-defined, the images of the generators in H have to satisfy all the same relations as the generators do in G. And the resulting homomorphism will be an embedding just in case the images of the generators do not satisfy any further relations.

Proposition 2.7. *Let A and B be structures and $X \subseteq A$ a subset such that $A = \langle X \rangle$. Let $h\colon X \to B$ be a partial function $A \dashrightarrow B$.*

(1) h extends to a homomorphism $A \to B$ if and only if it preserves all atomic formulas.

(2) h extends to an embedding $A \to B$ if and only if it preserves and reflects all atomic formulas (equivalently, it preserves all atomic and negated atomic formulas).

Proof. Suppose h extends to a homomorphism $A \to B$. Let $\varphi(\overline{x})$ be an atomic formula, and let $\overline{a}$ be an interpretation of $\overline{x}$ in X.

Case 1: φ is $t_1 = t_2$. If $A \models \varphi(\overline{a})$, this means $t_1^A(\overline{a}) = t_2^A(\overline{a})$, which implies $h(t_1^A(\overline{a})) = h(t_2^A(\overline{a}))$. By Lemma 2.6, $t_1^B(h(\overline{a})) = t_2^B(h(\overline{a}))$, so $B \models \varphi(h(\overline{a}))$. We have shown h preserves φ. If h extends to an embedding, then the implication above is an equivalence, using the fact that h is injective, so h reflects φ.

Case 2: φ is $R(t_1, \ldots, t_n)$. If $A \models \varphi(\overline{a})$, this means $(t_i^A(\overline{a}))_{i=1}^n \in R^A$, which implies $(h(t_i^A(\overline{a})))_{i=1}^n \in R^B$. By Lemma 2.6, $(t_i^B(h(\overline{a})))_{i=1}^n \in R^B$, so $B \models \varphi(h(\overline{a}))$. We have shown that h preserves φ. If h extends to an embedding, then the implication above is an equivalence, since h reflects R.

Conversely, suppose h preserves all atomic formulas. We will show that h extends to a homomorphism $h' \colon A \to B$. Since $A = \langle X \rangle$, every element of A can be written as $t^A(\overline{a})$ for some term t and some $\overline{a}$ from X. We define $h'(t^A(\overline{a})) = t^B(h(\overline{a}))$. This is well-defined, because if $t_1^A(\overline{a}) = t_2^A(\overline{a}')$, then $A \models t_1(\overline{a}) = t_2'(\overline{a}')$, which implies $B \models t_1(h(\overline{a})) = t_2(h(\overline{a}'))$, since h preserves atomic formulas. So $t_1^B(h(\overline{a})) = t_2^B(h(\overline{a}'))$. If h additionally reflects atomic formulas, then the implication above is an equivalence, so h' is injective.

Further, the above definition implies that h' preserves the function symbols in the language. Let $f \in \mathcal{L}$ be a function symbol, and let $\overline{t} = (t_1^A(\overline{a}), \ldots, t_n^A(\overline{a}))$ be a tuple of elements of A, which we may assume are all terms interpreted on the same tuple $\overline{a}$ from X, by expanding the contexts of the terms. Let s be the composite term $f(t_1, \ldots, t_n)$. Then

$$\begin{aligned} h'(f^A(\overline{t})) &= h'(s^A(\overline{a})) \\ &= s^B(h(\overline{a})) \\ &= f^B(t_1^B(h(\overline{a})), \ldots, t_n^B(h(\overline{a}))) \\ &= f^B(h'(t_1^A(\overline{a})), \ldots, h'(t_n^A(\overline{a}))) \\ &= f^B(h'(\overline{t})). \end{aligned}$$

Finally, let $R \in \mathcal{L}$ be a relation symbol, and let $\overline{t} = (t_1^A(\overline{a}), \ldots, t_n^A(\overline{a}))$ be a tuple of elements of A, as above. Let $\varphi(\overline{x})$ be the atomic formula

$$R(t_1(\overline{x}), \ldots, t_n(\overline{x})).$$

Then $\overline{t} \in R^A$ if and only if $A \models \varphi(\overline{a})$, which implies $B \models \varphi(h(\overline{a}))$, which means $(t_1^B(h(\overline{a})), \ldots, t_n^B(h(\overline{a}))) = h'(\overline{t}) \in R^B$. Thus h' is a homomorphism. If h also reflects all atomic formulas, then the implication above is an equivalence, so h' also reflects all relation symbols. □

Given an $\mathcal{L}$-structure A and a subset $C \subseteq A$, let $\mathcal{L}(C)$ be the language obtained from $\mathcal{L}$ by adding a new constant symbol for every element $c \in C$. When there is no chance for confusion, we will also denote the constant symbol

by c. We view A as an $\mathcal{L}(C)$ structure in the obvious way. In particular, $\mathcal{L}(A)$ is obtained by naming every element of A by a constant symbol.

Suppose C is a subset of a structure A. The **positive diagram** of C, denoted $\mathrm{Diag}^+(C)$, is the set of all atomic $\mathcal{L}(C)$-sentences true in A. That is, for every atomic $\mathcal{L}$-formula $\varphi(\overline{x})$ and every tuple $\overline{c}$ from C such that $A \models \varphi(\overline{c})$, the $\mathcal{L}(C)$-sentence $\varphi(\overline{c})$ is in $\mathrm{Diag}^+(C)$.

Similarly, the **diagram** of C, denoted $\mathrm{Diag}(C)$, is the set of all atomic and negated atomic $\mathcal{L}(C)$-sentences true in A.

When $\mathcal{L} \subseteq \mathcal{L}'$ are languages and A is an $\mathcal{L}'$-structure, we write $A|_{\mathcal{L}}$ for the **reduct** of A to the language $\mathcal{L}$, obtained by forgetting about the interpretations of the symbols in $\mathcal{L} \setminus \mathcal{L}'$.

Example 2.8. Let $\mathcal{L} = \{0, +, -\}$ be the language of abelian groups, and let $\mathcal{L}' = \{0, 1, +, -, \cdot\}$ be the language of rings. If R is a ring, the reduct $R|_{\mathcal{L}}$ is the underlying abelian group of R.

The following proposition now follows immediately from Proposition 2.7.

Proposition 2.9. *Let A be an $\mathcal{L}$-structure, let $C \subseteq A$ such that $A = \langle C \rangle$ (so C is a set of generators for A), and let B be a $\mathcal{L}(C)$-structure.*

(1) $B \models \mathrm{Diag}^+(C)$ if and only if the map $C \to B$ given by $c \mapsto c^B$ extends to a homomorphism $A \to B|_{\mathcal{L}}$.

(2) $B \models \mathrm{Diag}(C)$ if and only if the map $C \to B$ given by $c \mapsto c^B$ extends to an embedding $A \to B|_{\mathcal{L}}$.

The significance of Proposition 2.9 is that we can turn the problem of finding some homomorphism or embedding out of a structure A into the problem of finding a model for some theory, namely the (positive) diagram of A.

We have seen that embeddings preserve all atomic and negated atomic formulas. In fact, we get much more than this for free.

Definition 2.10. A formula φ is **quantifier-free** if it contains no quantifiers. That is, it is built up from atomic formulas and $\top$ and $\bot$ using only the connectives $\neg$, $\wedge$, and $\vee$.

A formula is **existential** if it has the form $\exists \overline{y}\, \varphi(\overline{x}, \overline{y})$, where φ is quantifier-free.

A formula is **universal** if it has the form $\forall \overline{y}\, \varphi(\overline{x}, \overline{y})$, where φ is quantifier-free.

In the above definitions, if $\overline{y} = (y_1, \ldots, y_m)$, then $\exists \overline{y}$ is shorthand for $\exists y_1\, \exists y_2\, \ldots\, \exists y_n$, and similarly for $\forall \overline{y}$. Note that $\overline{y}$ could be the empty tuple of variables, so every quantifier-free formula is both existential and universal.

Definition 2.11. Two formulas $\varphi(\overline{x})$ and $\psi(\overline{x})$ are **equivalent** if $A \models \varphi(\overline{a})$ if and only if $A \models \psi(\overline{a})$ for all structures A and all interpretations $\overline{a}$ of $\overline{x}$. That is, φ and ψ are equivalent if they define the same definable set in all structures.

Every universal formula $\forall \overline{y}\, \varphi(\overline{x}, \overline{y})$ is equivalent to the negation of an existential formula, $\neg \exists \overline{y}\, \neg \varphi(\overline{x}, \overline{y})$. Similarly, every existential formula is equivalent to the negation of a universal formula.

Exercise 4. Here is a classical result from propositional logic. Every quantifier-free formula is equivalent to one in disjunctive normal form: $\bigvee_{i=1}^{n} \bigwedge_{j=1}^{m} \varphi_{i,j}$, where $\varphi_{i,j}$ is atomic or negated atomic for all i and j. This can be proven by using de Morgan's laws and the distributivity of $\vee$ over $\wedge$: $\neg(P \wedge Q)$ is equivalent to $(\neg P \vee \neg Q)$, $\neg(P \vee Q)$ is equivalent to $(\neg P \wedge \neg Q)$, and $(P \vee Q) \wedge R$ is equivalent to $(P \wedge R) \vee (Q \wedge R)$.

Example 2.12. Consider the structure $(\mathbb{Z}; \leq)$ and the map $h \colon \mathbb{Z} \to \mathbb{Z}$ given by $n \mapsto 2n$. This is an embedding, since it is injective, and $a \leq b$ if and only if $2a \leq 2b$.

Let $\varphi(x, y)$ be the formula $\exists z\, (x < z < y)$. ($x < y < z$ is shorthand for $x < z \wedge z < y$.) This is an existential formula. It is preserved by h, since if there is some c such that $a < c < b$, then there is some c' such that $2a < c' < 2b$. For example, we can take $c' = h(c)$. On the other hand, φ is not reflected by h. For example, $\mathbb{Z} \models \varphi(h(1), h(2))$, since $2 < 3 < 4$, but $\mathbb{Z} \not\models \varphi(1, 2)$.

Example 2.13. Let G be a group, and let H be a substructure of G, so the inclusion map $i \colon H \to G$ is an embedding. Let φ be the universal sentence $\forall x \forall y\, (x \cdot y = y \cdot x)$. Then φ is reflected by i: this just says that a subgroup of an abelian group is abelian.

Proposition 2.14. *Let $h \colon A \to B$ be an embedding. Then:*

(1) h preserves and reflects all quantifier-free formulas.

(2) h preserves all existential formulas.

(3) h reflects all universal formulas.

Proof. Let φ be a quantifier-free formula. We prove (1) by induction on the complexity of φ. In the base case, if φ is atomic, then we already know h preserves and reflects φ, by Proposition 2.7. If φ is $\top$ of $\bot$, it is clearly preserved and reflected. If φ is $\neg\psi$, then since h preserves and reflects ψ by induction, it reflects and preserves φ. If φ is $\psi \vee \chi$, then $A \models \varphi(\overline{a})$ if and only if $A \models \psi(\overline{a})$ or $A \models \chi(\overline{a})$. By induction, this is equivalent to $B \models \psi(h(\overline{a}))$ or $B \models \chi(h(\overline{a}))$, which is equivalent to $B \models \varphi(h(\overline{a}))$. The argument for $\wedge$ is similar.

For (2), let $\exists \overline{y}\, \varphi(\overline{x}, \overline{y})$ be an existential formula, and suppose $A \models \exists \overline{y}\, \varphi(\overline{a}, \overline{y})$. Then there is some tuple $\overline{b}$ such that $A \models \varphi(\overline{a}, \overline{b})$. Since h preserves quantifier-free formulas like φ, we have $B \models \varphi(h(\overline{a}), h(\overline{b}))$. So $B \models \exists \overline{y}\, \varphi(\overline{a}, \overline{y})$.

For (3), every universal formula is equivalent to the negation of an existential formula. By (2), existential formulas are preserved by h, so universal formulas are reflected by h. □

It's now natural to ask whether there are functions which preserve (and reflect) *all* formulas.

Definition 2.15. A function $h\colon A \to B$ is an **elementary embedding** if it preserves all formulas. That is, for all formulas $\varphi(\overline{a})$ and all interpretations $\overline{a}$ of $\overline{x}$ in A,

$$\text{if } A \models \varphi(\overline{a}), \text{then } B \models \varphi(h(\overline{a})).$$

Note that if h preserves all formulas, then in particular it preserves the negation of every formula, so it also reflects every formula. Thus we can strengthen the above to $A \models \varphi(\overline{a})$ if and only if $B \models \varphi(h(\overline{a}))$.

In particular, elementary embeddings preserve and reflect all sentences, so if $h\colon A \to B$ is an elementary embedding, then $A \equiv B$.

If $M \subseteq N$ and the inclusion is an elementary embedding, we write $M \preccurlyeq N$, and say that M is an **elementary substructure** of N, or N is an **elementary extension** of M.

Example 2.16. The inclusion map $\mathbb{Q} \to \mathbb{R}$ is an embedding of fields but not an elementary embedding, so $\mathbb{Q} \not\preccurlyeq \mathbb{R}$. For example, let $\varphi(x)$ be the formula $\exists y\,(y \cdot y = x)$. Then we have $\mathbb{Q} \models \neg\varphi(2)$, but $\mathbb{R} \models \varphi(2)$.

Example 2.17. Consider the structure $(\mathbb{N}; \leq)$. The map $h\colon \mathbb{N} \to \mathbb{N}$ given by $n \mapsto n+1$ is an embedding of $\mathbb{N}$ in itself (and $\mathbb{N}$ is elementarily equivalent to itself), but h is not an elementary embedding. For example, letting $\varphi(x)$ be the formula $\forall y\,(x \leq y)$, we have $\mathbb{N} \models \varphi(0)$, but $\mathbb{N} \not\models \varphi(h(0))$. Note that two structures are elementarily equivalent if they agree on the truth of sentences, but an elementary embedding requires that the structures also agree on the satisfaction of formulas on tuples from the structure.

Exercise 5. Consider the structure $(\mathbb{N}; \leq)$. Show that if $h\colon \mathbb{N} \to \mathbb{N}$ is an elementary embedding, then h is the identity map.

It is harder to give examples of elementary embeddings, since the definition quantifies over all formulas. In the next section, we will see that for models of certain theories, elementary embeddings are easy to come by. For now, one obvious source of elementary embeddings is isomorphisms.

Exercise 6. Show that every isomorphism is an elementary embedding. In particular, isomorphic structures are elementarily equivalent.

Exercise 7. Show that the composition of two homomorphisms is a homomorphism. Do the same for embeddings and elementary embeddings.

We can also use the following proposition to construct elementary embeddings by compactness.

The $\mathcal{L}(A)$-theory $\mathrm{Th}_{\mathcal{L}(A)}(A)$ is called the **elementary diagram of** A, and denoted $\mathrm{EDiag}(A)$: it contains all of the information about all first-order formulas satisfied by all tuples from A. Similarly to Proposition 2.9, we have:

Proposition 2.18. *Let A be an $\mathcal{L}$-structure, and let B be an $\mathcal{L}(A)$-structure. Then $B \models \mathrm{EDiag}(A)$ if and only if the map $a \mapsto a^B$ is an elementary embedding $A \to B|_{\mathcal{L}}$.*

2.2 A test for quantifier elimination

Many of the concepts in model theory ($\models$, $\preccurlyeq$, $\mathrm{Th}(A)$, etc.) are difficult to understand concretely because they quantify over all formulas. Typically, the quantifier-free formulas are much easier to understand than general formulas, so it is very desirable to be able to reduce all formulas to quantifier-free ones. In the context of some theories, we can do this.

Definition 2.19. Let T be a theory. Two formulas $\varphi(\overline{x})$ and $\psi(\overline{x})$ (in the same variable context $\overline{x}$) are **T-equivalent** if $M \models \varphi(\overline{a})$ if and only if $M \models \psi(\overline{a})$ for all models $M \models T$ and all interpretations $\overline{a}$ of $\overline{x}$ in M. That is, φ and ψ are equivalent if they define the same definable set in all models of T.

Definition 2.20. A theory T has **quantifier elimination** (or **eliminates quantifiers**) if every formula is T-equivalent to a quantifier-free formula.

Example 2.21. Let T be the theory of fields. The formula $\exists y\,(xy = 1)$, expressing that x is invertible, is T-equivalent to the quantifier-free formula $x \neq 0$.

Similarly, the formula

$$\exists a\,\exists b\,\exists c\,\exists d\,(ax + bz = 1 \wedge ay + bw = 0 \wedge cx + dz = 0 \wedge cy + dw = 0),$$

expressing that the matrix $\begin{pmatrix} x & y \\ z & w \end{pmatrix}$ is invertible, is T-equivalent to the quantifier-free formula $xw - yz \neq 0$.

The formula $\varphi(a, b, c)$ defined by $\exists y\,(ay^2 + by + c = 0)$ is not T-equivalent to a quantifier-free formula. To see this, note for example that if $h\colon \mathbb{R} \to \mathbb{C}$ is the inclusion embedding, h preserves and reflects all quantifier-free formulas. But it does not reflect φ, since $\mathbb{R} \not\models \varphi(1, 0, 1)$, while $\mathbb{C} \models \varphi(1, 0, 1)$.

However, if we include $\leq$ in the language and move to the complete theory of $\mathbb{R}$, we have that φ is $\mathrm{Th}(\mathbb{R}; 0, 1, +, -, \cdot, \leq)$-equivalent to the quantifier-free formula $b^2 - 4ac \geq 0$.

We begin with a useful structural criterion for a single formula to be T-equivalent to a quantifier-free formula.

Theorem 2.22. *Let T be a theory and $\varphi(\overline{x})$ a formula. The following are equivalent.*

(1) φ is T-equivalent to a quantifier-free formula.

(2) Suppose M and N are models of T, A is any structure, and $g\colon A \to M$ and $h\colon A \to N$ are embeddings. For all $\overline{a}$ from A, if $N \models \varphi(h(\overline{a}))$, then $M \models \varphi(g(\overline{a}))$.

Proof. For (1) implies (2), suppose φ is T-equivalent to a quantifier-free formula ψ. Then in the context of (2), since g and h preserve and reflect ψ, we have

$$\begin{aligned} N \models \varphi(h(\overline{a})) &\Rightarrow N \models \psi(h(\overline{a})) \\ &\Rightarrow A \models \psi(\overline{a}) \\ &\Rightarrow M \models \psi(g(\overline{a})) \\ &\Rightarrow M \models \varphi(g(\overline{a})) \end{aligned}$$

For (2) implies (1), let Ψ be the set of all quantifier-free formulas $\psi(\overline{x})$ such that $T \models \forall \overline{x}\,(\varphi \to \psi)$. This is the set of all quantifier-free consequences of φ. If φ is to be T-equivalent to a quantifier-free formula, it must be in the set Ψ: our goal is to find a formula $\psi \in \Psi$ such that also $T \models \forall \overline{x}\,(\psi \to \varphi)$.

Note that Ψ is closed under finite conjunctions, since if

$$T \models \forall \overline{x}\,(\varphi \to \psi_i) \quad \text{for all } 1 \leq i \leq n,$$

then

$$T \models \forall \overline{x} \left(\varphi \to \bigwedge_{i=1}^{n} \psi_i \right).$$

If $\overline{x} = (x_1, \ldots, x_n)$, let $\mathcal{L}(\overline{c})$ be the expansion of $\mathcal{L}$ obtained by introducing new constant symbols $\overline{c} = (c_1, \ldots, c_n)$, and let $\Psi(\overline{c}) = \{\psi(\overline{c}) \mid \psi(\overline{x}) \in \Psi\}$.

Claim: It suffices to show that $T \cup \Psi(\overline{c}) \models \varphi(\overline{c})$.

Indeed, if $T \cup \Psi(\overline{c}) \models \varphi(\overline{c})$, then by the compactness theorem, there are finitely many formulas $\psi_1, \ldots, \psi_n \in \Psi$ such that

$$T \cup \{\psi_1(\overline{c}), \ldots, \psi_n(\overline{c})\} \models \varphi(\overline{c}).$$

Let $\psi = \bigwedge_{i=1}^{n} \psi_i$ (if $n = 0$, then $\psi = \top$). By the observation above, $\psi \in \Psi$, and we have that $T \models \psi(\overline{c}) \to \varphi(\overline{c})$. Since the constant symbols $\overline{c}$ are not mentioned in T, this implication is true for any tuple $\overline{c}$ from a model of T, and it follows that $T \models \forall \overline{x}\,(\psi(\overline{x}) \to \varphi(\overline{x}))$. But the reverse is true since $\psi \in \Psi$, so φ is T-equivalent to ψ.

Having established the claim, we show that $T \cup \Psi(\overline{c}) \models \varphi(\overline{c})$. So we pick a model $M \models T \cup \Psi(\overline{c})$ and show that $M \models \varphi(\overline{c})$. Let $C = \langle \overline{c}^M \rangle$, and let $g\colon C \to M$ be the inclusion. We would like to find a model for the theory $T \cup \mathrm{Diag}_M(\overline{c}) \cup \{\varphi(\overline{c})\}$.

Suppose for contradiction that this theory is not satisfiable. Then by compactness, there is a finite subset which is unsatisfiable, so there are finitely many atomic and negated atomic sentences $\theta_i(\overline{c}) \in \mathrm{Diag}_M(\overline{c})$ such that $T \cup \{\theta_i(\overline{c}) \mid 1 \leq i \leq n\} \cup \{\varphi(\overline{c})\}$ is not satisfiable. But then $T \cup \{\varphi(\overline{c})\} \models \neg \bigwedge_{i=1}^{n} \theta_i(\overline{c})$, so $T \models \varphi(\overline{c}) \to \bigvee_{i=1}^{n} \neg\theta_i(\overline{c})$, and

$$T \models \forall \overline{x} \left(\varphi(\overline{x}) \to \bigvee_{i=1}^{n} \neg\theta_i(\overline{x}) \right)$$

.

It follows that $\bigvee_{i=1}^{n} \neg\theta_i(\overline{x})$ is in Ψ, so $M \models \bigvee_{i=1}^{n} \neg\theta_i(\overline{c})$. But also $M \models \theta_i(\overline{c})$ for all $1 \leq i \leq n$, since $\theta_i(\overline{c}) \in \mathrm{Diag}_M(\overline{c})$. This is a contradiction.

Thus there is a model $N \models T \cup \mathrm{Diag}_M(\overline{c}) \cup \{\varphi(\overline{c})\}$. By the diagram lemma, the map $c_i \mapsto c_i^N$ extends to an embedding $h\colon C \to N$. Now by our assumption (2), we have $N \models \varphi(h(\overline{c}))$, so $M \models \varphi(g(\overline{c}))$, as desired. □

Note that the proof of Theorem 2.22 used the compactness theorem twice. Compactness is a fundamentally non-constructive proof technique, so the proof

doesn't give us any information on how to explicitly find the quantifier-free formula equivalent to φ. In certain circumstances, we can do better by giving an explicit algorithm for eliminating quantifiers.

Next we note that to prove that T has quantifier elimination, we only have to consider formulas of a particularly simple form. Semantically, it says that T has quantifier elimination just in case the coordinate projection of an intersection of sets defined by atomic and negated atomic formulas is a Boolean combination of sets defined by such formulas.

Definition 2.23. A **primitive formula** in context $\overline{x}$ has the form $\exists y \bigwedge_{i=1}^{n} \varphi_i$, where each formula φ_i is atomic or negated atomic in context $\overline{x}y$.

Theorem 2.24. *If every primitive formula is T-equivalent to a quantifier-free formula, then T has quantifier elimination.*

Proof. We prove by induction on the complexity of formulas that every formula is T-equivalent to a quantifier-free formula. The base cases and inductive steps are clear, and a formula of the form $\forall y\, \varphi$ can be rewritten as $\neg\exists y \neg\varphi$, so it suffices to handle the induction step for the existential quantifier.

So consider a formula φ in the form $\exists y\, \psi$. By induction, ψ is T-equivalent to a quantifier-free formula θ, so φ is T-equivalent to $\exists y\, \theta$. Writing θ in disjunctive normal form, φ is T-equivalent to

$$\exists y \left(\bigvee_{i=1}^{n} \bigwedge_{j=1}^{m} \varphi_{ij} \right),$$

where each φ_{ij} is atomic or negated atomic. It follows that φ is T-equivalent to

$$\bigvee_{i=1}^{n} \exists y \left(\bigwedge_{j=1}^{m} \varphi_{ij} \right),$$

since existential quantifiers distribute over disjunctions. For fixed i, the formula $\exists y \left(\bigwedge_{j=1}^{m} \varphi_{ij} \right)$ is primitive, so by hypothesis it is T-equivalent to a quantifier-free formula, and φ is T-equivalent to the disjunction of these n formulas. $\square$

Putting Theorem 2.22 and Theorem 2.24 together, we obtain the following test for quantifier elimination.

Corollary 2.25. *Let T be a theory. The following are equivalent:*

(1) T has quantifier elimination.

(2) Suppose M and N are models of T, A is any structure, $g\colon A \to M$ and $h\colon A \to N$ are embeddings, and φ is a primitive formula. For all $\overline{a}$ from A, if $N \models \varphi(h(\overline{a}))$, then $M \models \varphi(g(\overline{a}))$.

2.3 Algebraically closed fields

Before beginning to study the model theory of the real numbers, we consider the model theory of fields in general, and in particular the case of algebraically closed fields, which are a bit simpler than real closed fields. We work in the language of rings, $\mathcal{L}_r = \{0, 1, +, -, \cdot\}$.

The theory of fields, T_f, is axiomatized by the following sentences:

- $\forall x \, \forall y \, \forall z \, ((x + y) + z = x + (y + z))$
- $\forall x \, \forall y \, (x + y = y + x)$
- $\forall x \, (0 + x = x)$
- $\forall x \, (x + -x = 0)$
- $\forall x \, \forall y \, \forall z \, ((x \cdot y) \cdot z = x \cdot (y \cdot z))$
- $\forall x \, \forall y \, (x \cdot y = y \cdot x)$
- $\forall x \, (1 \cdot x = x)$
- $\forall x \, \forall y \, \forall z \, (x \cdot (y + z) = (x \cdot y) + (x \cdot z))$
- $\neg(0 = 1)$
- $\forall x \, (x = 0 \vee \exists y \, (x \cdot y = 1))$

Remark 2.26. Every atomic formula $\varphi(\overline{x})$ is T_f-equivalent to one of the form $p(\overline{x}) = 0$, where $p \in \mathbb{Z}[x_1, \ldots, x_n]$. Indeed, there are no relation symbols in the language of rings, so every atomic formula has the form $t_1 = t_2$, and by subtracting t_2 to the other side and liberally applying the distributive law, we find that $(t_1 - t_2) = 0$ is T_f-equivalent to a polynomial equation. In particular, if $K \models T_f$ is a field, and $\overline{a}$ is a tuple from K, then an atomic formula $\varphi(\overline{x}, \overline{a})$ with parameters $\overline{a}$ is equivalent in K to $p(\overline{x}) = 0$, where $p \in A[x_1, \ldots, x_n]$ and A is the subring of K generated by $\overline{a}$.

An important example of an atomic sentence is χ_p for prime p, which asserts that the field has characteristic 0:

$$\underbrace{1 + \cdots + 1}_{p \text{ times}} = 0.$$

We can also axiomatize the class of fields of characteristic 0 by

$$T_f \cup \{\neg \chi_p \mid p \text{ prime}\}.$$

Exercise 8. Show that there is no single sentence θ such that the models of $T_f \cup \{\theta\}$ are exactly the fields of characteristic 0. *Hint:* Use the compactness theorem.

The theory ACF of algebraically closed fields consists of the theory of fields T_f, together with a sentence φ_d for every degree $d \geq 1$, expressing that every monic polynomial of degree d with coefficients in the field has a root in the field:

$$\forall a_0 \ldots \forall a_{d-1} \exists y \, (y^d + a_{d-1} \cdot y^{d-1} + \cdots + a_1 \cdot y + a_0 = 0).$$

We write ACF_p for $\mathrm{ACF} \cup \{\chi_p\}$ when p is prime, and we write ACF_0 for $\mathrm{ACF} \cup \{\neg\chi_p \mid p \text{ prime}\}$.

Theorem 2.27. ACF *has quantifier elimination.*

Proof. We use the test in Corollary 2.25. So suppose K_1 and K_2 are algebraically closed fields, A is an $\mathcal{L}_r$, structure, and $g\colon A \to K_1$ and $h\colon A \to K_2$ are embeddings. Let φ be a primitive formula, $\exists y \bigwedge_{i=1}^n \varphi_i(\overline{x}, y)$, and let $\overline{a}$ be a tuple from A such that $K_2 \models \varphi(h(\overline{a}))$. We would like to show that $K_1 \models \varphi(g(\overline{a}))$.

Note that A is isomorphic to a subring $g(A)$ of K_1, so it is an integral domain. Let $A' = \mathrm{Frac}(A)$, the field of fractions of A. The embedding g extends to an embedding $g'\colon A' \to K_1$, with image the subfield of K_1 generated by $g(A)$, and similarly h extends to an embedding $h'\colon A' \to K_2$. Now let A'' be an algebraic closure of the field A'. Since K_1 and K_2 are algebraically closed, the embeddings g' and h' extend to embeddings $g''\colon A'' \to K_1$ and $h''\colon A'' \to K_2$, with images $F_1 \subseteq K_1$ and $F_2 \subseteq K_2$, the algebraic closures of $g'(A')$ in K_1 and $h'(A')$ in K_2, respectively.

Now let's analyze the primitive formula $\exists y \bigwedge_{i=1}^n \varphi_i(h(\overline{a}), y)$. Since each φ_i is atomic or negated atomic, we may assume by Remark 2.26 that $\varphi_i(h(\overline{a}), y)$ is a polynomial equality $p_i(y) = 0$ or an inequality $p_i(y) \neq 0$, where $p_i \in F_2[y]$. If any of the φ_i are equalities, then letting b be a witness to the existential quantifier in K_2, b is algebraic over F_2, and F_2 is algebraically closed, so $b \in F_2$. On the other hand, if each of the φ_i are inequalities, then since each polynomial p_i has only finitely many roots in K_2 and F_2 is infinite, we can find a witness $b \in F_2$ for the existential quantifier.

In either case, we have $b \in F_2 = h''(A'')$, so there is some $b' \in A''$ with $h(b') = b$, so $K_2 \models \bigwedge_{i=1}^n \varphi_i(h(\overline{a}), h(b'))$. Since embeddings preserve and reflect quantifier-free formulas, this implies $K_1 \models \bigwedge_{i=1}^n \varphi_i(g(\overline{a}), g(b'))$. In particular, $K_1 \models \varphi(g(\overline{a}))$. □

That seemed easy, but note that there is some serious algebra about extending embeddings hiding in the second paragraph of the proof.

Here are two exercises in which you can see our test for quantifier elimination in action.

Exercise 9. Let K be a field, let $\mathcal{L} = \{0, +, -\} \cup \{c \mid c \in K\}$ be the language of abelian groups with an additional unary function symbol c for each element $c \in K$. Write down the theory T of infinite vector spaces over K in this language. (To express that the vector space is infinite, it suffices to include the axiom $\exists x \, (x \neq 0)$ when K is infinite. When K is finite, include an axiom φ_n:

$$\exists x_1 \ldots \exists x_n \left(\bigwedge_{i \neq j} x_i \neq x_j \right)$$

for each $n \in \mathbb{N}$.)

Show that T has quantifier elimination. Where was the property that models are infinite used in the proof?

The next exercise shows that quantifier elimination can be quite sensitive to the choice of language.

Exercise 10. Let $s\colon \mathbb{N} \to \mathbb{N}$ be the successor function $n \mapsto n+1$. Show that $\mathrm{Th}(\mathbb{N}; s)$ does not have quantifier elimination, but $\mathrm{Th}(\mathbb{N}; s, 0)$ does (where 0 is a constant symbol naming the element 0). Note that $\mathrm{Th}(\mathbb{N}; s)$ refers to the complete theory of the structure $(\mathbb{N}; s)$. You do not need to write down axioms for this theory to solve the exercise, but you do need to think about what properties an arbitrary model for this theory must have.

The theory ACF is not complete, because it does not determine the characteristic: for any prime p, $\mathrm{ACF} \not\models \chi_p$ and $\mathrm{ACF} \not\models \neg\chi_p$. But as a first application of quantifier elimination, we will show that the theories ACF_p and ACF_0 are complete.

Proposition 2.28. *Suppose T is a satisfiable theory with quantifier elimination, and suppose further that there is a structure A such that for any model $M \models T$, there is an embedding $g\colon A \to M$. Then T is complete.*

Proof. Since T is satisfiable, it has a model $M \models T$, and there is an embedding $g\colon A \to M$. Let φ be a sentence. Then either $M \models \varphi$ or $M \models \neg\varphi$; without loss of generality, we assume $M \models \varphi$. We claim that $T \models \varphi$.

So let $N \models T$. By our assumption, there is an embedding $h\colon A \to N$. By quantifier elimination, φ is equivalent to a quantifier-free sentence, and by Theorem 2.22, $M \models \varphi$ implies $N \models \varphi$. □

The following result does not apply to theories in the language of rings, because $\mathcal{L}_r$ contains constant symbols. But it may help to clarify the relationship between quantifier elimination and completeness.

Corollary 2.29. *If $\mathcal{L}$ has no 0-ary function symbols (constant symbols) or 0-ary relation symbols (proposition symbols), then every satisfiable theory with quantifier elimination is complete.*

Proof. Suppose $\mathcal{L}$ has no 0-ary symbols, and T is a theory with quantifier elimination. Then there is a unique empty $\mathcal{L}$-structure: each n-ary function symbol is interpreted as the unique empty function $\emptyset^n \to \emptyset$, and each n-ary relation symbol is interpreted as the unique empty relation on $\emptyset^n$. And the empty structure embeds uniquely into every model of T, by the empty embedding. So T is complete, by Proposition 2.28.

Another way of proving this is to note that if T has quantifier elimination, then every sentence is T-equivalent to a quantifier-free sentence. But if there are no 0-ary symbols, the only quantifier-free sentences are $\top$ and $\bot$. So for every sentence φ, either $T \models \varphi \leftrightarrow \top$ or $T \models \varphi \leftrightarrow \bot$, which are equivalent to $T \models \varphi$ and $T \models \neg\varphi$, respectively. □

Corollary 2.30. *Let p be prime or 0. Then ACF_p is complete.*

Proof. Since $\mathrm{ACF} \subseteq \mathrm{ACF}_p$, ACF_p has quantifier elimination. If $p = 0$, we take $A = \mathbb{Q}$, and if p is prime, we take $A = \mathbb{F}_p$, the finite field of order p. Then A embeds in any field of characteristic p, and in particular in every model of ACF_p, so ACF_p is complete, by Proposition 2.28. □

Completeness in turn gives us the following interesting result, which makes precise an intuition that the behavior of algebraically closed fields of characteristic 0 is the "limit as p goes to ∞" of the behavior of algebraically closed fields of characteristic p.

Corollary 2.31 (Transfer principle for algebraically closed fields)**.** *Let φ be a sentence in the language of rings. The following are equivalent:*

(1) Some algebraically closed field of characteristic 0 satisfies φ.

(2) Every algebraically closed field of characteristic 0 satisfies φ.

(3) For all but finitely many primes p, some algebraically closed field of characteristic p satisfies φ.

(4) For all but finitely many primes p, every algebraically closed field of characteristic p satisfies φ.

Proof. (1)⇒(2): If $M \models \mathrm{ACF}_0$ and $M \models \varphi$, then $\mathrm{ACF}_0 \not\models \neg\varphi$, so $\mathrm{ACF}_0 \models \varphi$ by completeness.

(2)⇒(4): We have $\mathrm{ACF}_0 \models \varphi$, so by compactness there are finitely many primes $p_1, \ldots, p_n$ such that $\mathrm{ACF} \cup \{p_i \neq 0 \mid 1 \leq i \leq n\} \models \varphi$. For any prime q not among these finitely many exceptions, any algebraically closed field of characteristic q satisfies $\mathrm{ACF} \cup \{p_i \neq 0 \mid 1 \leq i \leq n\}$, hence satisfies φ.

(4)⇒(3): Trivial.

(3)⇒(1): Assume for contradiction that (1) fails. Then every algebraically closed field of characteristic 0 satisfies $\neg\varphi$. By (2)⇒(4) for $\neg\varphi$, we have that for all but finitely many primes p, every algebraically closed field of characteristic p satisfies $\neg\varphi$. This contradicts (3) (since there are infinitely many primes!). □

I'll end this section with some discussion on the significance of completeness and quantifier elimination.

If a theory T has a reasonable (meaning computable or computably enumerable) axiomatization, then as soon as we know T is complete, we have an algorithm for deciding which sentences are entailed by T (we say that T is a **decidable** theory). For any sentence φ, start searching for a proof $T \vdash \varphi$, and simultaneously start searching for a proof $T \vdash \neg\varphi$. Since T is complete, one of these searches eventually terminates.

But the algorithm described above is hopelessly inefficient. In many situations, a more efficient algorithm can be found via **effective quantifier elimination**, i.e. an algorithm for finding a quantifier-free formula which is T-equivalent to a given formula. Applying this algorithm to an arbitrary sentence φ produces

a quantifier-free sentence ψ which is T-equivalent to it. But a quantifier-free sentence ψ is just a Boolean combination of atomic sentences, the truth value of which can usually be easily checked.

The proof we gave above that ACF has quantifier elimination was entirely ineffective. An effective quantifier elimination algorithm exists for ACF; unfortunately, it is also very inefficient (doubly exponential running time in the size of the input).

Now let's think about what quantifier elimination for ACF means in the language of definable sets. By Remark 2.26, an atomic formula with parameters from K describes the zero-set in K^n of a polynomial in $K[x_1, \ldots, x_n]$. An **algebraic set** (a Zariski closed set) is one defined by a system of polynomial equations, i.e., a finite intersection of zero-sets of polynomials. A **constructible set** is a finite Boolean combination of algebraic sets; these are exactly the sets in K^n defined by quantifier-free formulas.

Now existential quantification corresponds to a coordinate projection of definable sets. It is not true in general that the coordinate projection of an algebraic set is algebraic. For example, the projection of the set $\{(x, y) \in K^2 \mid xy = 1\}$ onto the first coordinate is the set $\{x \in K \mid x \neq 0\}$, which is the complement of an algebraic set. But we do have the following theorem of algebraic geometry, which is equivalent to quantifier elimination for ACF.

Corollary 2.32 (Chevalley's theorem)**.** *In affine space over an algebraically closed field, a coordinate projection of a constructible set is constructible.*

The same is not true for the field $\mathbb{R}$ of real numbers. For example, consider the formula $\exists x\,(y = x^2)$. This formula defines the set $\{y \in \mathbb{R} \mid y \geq 0\}$. Geometrically, we can view it as the coordinate projection of the parabola in $\mathbb{R}^2$ onto the y-axis. This set is not defined by any quantifier-free formula, i.e. it is not constructible: An atomic formula in a single free variable is a polynomial equation $p(x) = 0$ or $p(x) \neq 0$, so it defines a finite or cofinite (finite complement) set, and a finite Boolean combination of finite or cofinite sets is finite or cofinite. But $\{x \in \mathbb{R} \mid x \geq 0\}$ is infinite and has infinite complement.

For a more complicated example, consider the formula $\exists x\,(ax^2 + bx + c = 0)$. This formula is ACF-equivalent to the formula $(a \neq 0 \vee b \neq 0 \vee c = 0)$. But in $\mathbb{R}$, it defines the set $\{(a, b, c) \in \mathbb{R}^3 \mid b^2 - 4ac \geq 0\}$, and it is possible to show that this set is not constructible.

Both of these examples suggest that if we want to understand formulas relative to the theory of the real field, we need pay attention to the linear ordering $\leq$ on $\mathbb{R}$ – even though this symbol is not included in the language of fields! If we were to include it in the language, both of the formulas above would become equivalent to quantifier-free formulas: $y \geq 0$ and $b^2 - 4ac \geq 0$, respectively. So to understand the model theory of the real field, we need to take a detour into the algebra of ordered fields.

3 Real algebra

3.1 Ordered (and orderable) fields

Definition 3.1. An **ordered ring** is a ring R equipped with a linear order $\leq$ such that for all a, b, and $c \in R$:

(a) If $a \leq b$, then $a + c \leq b + c$.

(b) If $a \leq b$ and $0 \leq c$, then $ac \leq bc$.

Example 3.2. We are primarily interested in ordered fields (and sometimes ordered integral domains, which are subrings of ordered fields). Our most basic example of an ordered field is $\mathbb{R}$ with its usual ordering. Any subfield of an ordered field is an ordered field (with the induced suborder), for example: $\mathbb{Q}$, $\mathbb{Q}[\sqrt{2}]$, and $\mathbb{Q}^r = \overline{\mathbb{Q}} \cap \mathbb{R}$, the field of real algebraic numbers.

Exercise 11. Consider $\mathbb{R}(t)$, the field of rational functions over $\mathbb{R}$. Define a linear order $\leq$ on $\mathbb{R}(t)$, making $\mathbb{R}(t)$ into an ordered field, such that $\leq$ extends the usual ordering on $\mathbb{R}$ and such that $r \leq t$ for all $r \in \mathbb{R}$. (You may find it easier to solve this exercise by defining a positive cone on $\mathbb{R}(t)$, rather than defining $\leq$ explicitly, see Proposition 3.10 below, or by defining an ordering on $\mathbb{R}[t]$ and using Theorem 3.15 below to extend this order to $\mathbb{R}(t)$.)

Proposition 3.3. *Let R be a non-zero ordered ring.*

(1) For all $a \in R$, $0 \leq a$ if and only if $-a \leq 0$.

(2) For all $a, b, c \in R$, if $a \leq b$ and $c \leq 0$, then $bc \leq ac$.

(3) For all $a \in R$, $0 \leq a^2$.

(4) For all $a_1, \ldots, a_k \in R$, $0 \leq a_1^2 + \cdots + a_k^2$.

(5) $-1 < 0 < 1$.

(6) R has characteristic 0.

Proof. (1) Suppose $0 \leq a$. Then $0 + -a \leq a + -a$, so $-a \leq 0$. Conversely, if $-a \leq 0$, then $-a + a \leq 0 + a$, so $0 \leq a$.

(2) By (1), $0 \leq -c$. Then $a \leq b$ implies $-ac \leq -bc$. Adding $(ac + bc)$ to both sides, $bc \leq ac$.

(3) We have $0 \leq a$ or $a \leq 0$. If $0 \leq a$, then multiplication by a preserves the inequality: $0 = 0a \leq aa = a^2$. If $a \leq 0$, then multiplication by a reverses the inequality by (2): $0 = 0a \leq aa = a^2$.

(4) By induction on k. The base case is (3). For the inductive step, suppose $0 \leq a_1^2 + \cdots + a_k^2$. Then $a_{k+1}^2 \leq a_1^2 + \cdots + a_k^2 + a_{k+1}^2$, and by (3), $0 \leq a_{k+1}^2$, so by transitivity $0 \leq a_1^2 + \cdots + a_{k+1}^2$.

(5) $1 = 1^2$, so by (3), $0 \leq 1$, and by (1), $-1 \leq 0$. In a non-zero ring, $0 \neq 1$ and $0 \neq -1$, so $-1 < 0 < 1$.

(6) Suppose for contradiction that R has characteristic $n > 0$. Then $-1 = n - 1 = \underbrace{1^2 + \cdots + 1^2}_{n-1 \text{ times}}$, so $0 \leq -1$, contradicting (5). □

Exercise 12. Show that there is a unique linear order $\leq$ on $\mathbb{R}$ such that $(\mathbb{R}; \leq)$ is an ordered field. Now do the same for $\mathbb{Q}$. (The two arguments will look rather different! But in both cases it could be useful to note that $a \leq b$ if and only if $0 \leq b - a$.)

Example 3.4. Some fields have more than one compatible order. The field $K = \mathbb{Q}[\sqrt{2}]$ has a canonical ordering $\leq$ inherited from $\mathbb{R}$, in which $-\sqrt{2} \leq 0 \leq \sqrt{2}$. But K admits an automorphism σ determined by $\sqrt{2} \mapsto -\sqrt{2}$, and this allows us to define a new linear order $\leq^\sigma$ on K by $a \leq^\sigma b \iff \sigma(a) \leq \sigma(b)$. Since σ is an automorphism, it is easy to see that $\leq^\sigma$ makes K into an ordered field. But in the new order, we have $\sqrt{2} \leq^\sigma 0 \leq^\sigma -\sqrt{2}$.

For any field K, we define

$$\Sigma K^2 = \{b_1^2 + \cdots + b_k^2 \mid b_1, \ldots, b_k \in K\}.$$

Definition 3.5. A field K is **formally real** if $-1 \notin \Sigma K^2$.

By Proposition 3.3(4) and (5), any ordered field is formally real: $-1 < 0$, so -1 cannot be a sum of squares.

Example 3.6. $\mathbb{C}$ is not formally real, since -1 is a square. It follows that there is no ordering $\leq$ on $\mathbb{C}$ such that $(\mathbb{C}; \leq)$ is an ordered field.

If the presence of -1 in the definition of formally real seems ad hoc, then solve the following exercises, which characterize formally real fields in ways that may seem more natural.

Exercise 13. Show that a field K is formally real if and only if zero can only be expressed as a sum of squares in the trivial way. That is, if $a_1^2 + \cdots + a_k^2 = 0$, then $a_1 = \cdots = a_k = 0$.

Exercise 14. Show that if $\mathrm{char}(K) \neq 2$, then K is formally real if and only if $\Sigma K^2 \neq K$. *Hint:* $x = \left(\frac{x+1}{2}\right)^2 - \left(\frac{x-1}{2}\right)^2$.

Our next goal is to show that any formally real field is orderable. For this purpose, it is convenient to reformulate the notion of an ordered field by focusing on the set of positive elements.

Definition 3.7. Let K be a field. A **prepositive cone** is a subset $P \subseteq K$ such that:

(a) P is closed under addition and multiplication: If $a, b \in P$, then $a + b \in P$ and $ab \in P$.

(b) $\Sigma K^2 \subseteq P$.

(c) $-1 \notin P$.

A **positive cone** is a prepositive cone such that additionally:

(d) For all $a \in K$, either $a \in P$ or $-a \in P$.

The motivating example for this definition is of course $P = \{a \in K \mid 0 \leq a\}$ in an ordered field $(K; \leq)$. Perhaps it would be better to say "non-negative" rather than "positive" here, but we will follow precedent from the literature.

Proposition 3.8. *Let $(K; \leq)$ be an ordered field, and let $P = \{a \in K \mid 0 \leq a\}$. Then P is a positive cone, and $a \leq b$ if and only if $b - a \in P$.*

Proof. (a) Let $a, b \in P$. Then $0 \leq a$ and $0 \leq b$. From $0 \leq b$ and $b = 0 + b \leq a + b$, we obtain $0 \leq a + b$. And also $0 = 0b \leq ab$.

(b) Proposition 3.3(4).

(c) Proposition 3.3(5).

(d) Proposition 3.3(1).

Finally, note that $a \leq b$ if and only if $0 = a - a \leq b - a$ if and only if $b - a \in P$. □

The reason for isolating the notion of prepositive cone is that condition (d) has a somewhat different character than conditions (a)-(c): rather than saying certain elements must be in or out of P, it requires us to make a decision, for every $a \in K$, whether a or $-a$ should be in P. To build a positive cone on a field, it will be useful to start with a prepositive cone, and then make these decisions one by one. The next lemma will be useful when working with prepositive cones. Note that part (2) says that these decisions are non-trivial: if $a \neq 0$, then a and $-a$ cannot both be in a prepositive cone P.

Lemma 3.9. *Let $P \subseteq K$ be a prepositive cone.*

(1) If $a \in P$ and $a \neq 0$, then $a^{-1} \in P$.

(2) If $a \in P$ and $-a \in P$, then $a = 0$.

Proof. (1) Since $a \neq 0$, a is invertible and $(a^{-1})^2 \in \Sigma K^2 \subseteq P$. So $a^{-1} = a(a^{-1})^2 \in P$.

(2) Suppose that $a \in P$ and $-a \in P$, but $a \neq 0$. By (1), $a^{-1} \in P$, so $-1 = (-a)a^{-1} \in P$, which is a contradiction. □

In the other direction, a positive cone on K determines a linear order on K.

Proposition 3.10. *Let $P \subseteq K$ be a positive cone. If we define*

$$a \leq b \iff b - a \in P,$$

then $(K; \leq)$ is an ordered field, and $P = \{a \in K \mid 0 \leq a\}$.

Proof. First, we need to check that $\leq$ is a linear order.

Reflexivity: For all $a \in K$, $a - a = 0 \in \Sigma K^2 \subseteq P$, so $a \leq a$.

Transitivity: For all $a, b, c \in K$, if $a \leq b$ and $b \leq c$, then $b - a \in P$ and $c - b \in P$, and $c - a = (c - b) + (b - a) \in P$, so $a \leq c$.

Antisymmetry: For all $a, b \in K$, if $a \leq b$ and $b \leq a$, then $b - a \in P$ and $a - b = -(b - a) \in P$. By Lemma 3.9(2), $b - a = 0$, and hence $a = b$.

Linearity: For all $a, b \in K$, either $b - a \in P$ or $a - b = -(b - a) \in P$. So either $a \leq b$ or $b \leq a$.

Having established that $\leq$ is a linear order, we need to check that it respects addition and multiplication on K. Let $a, b, c \in K$.

Addition: If $a \leq b$, then $b - a \in P$. Then $(b + c) - (a + c) = b - a \in P$, so $a + c \leq b + c$.

Multiplication: If $a \leq b$ and $0 \leq c$, then $b - a \in P$ and $c = c - 0 \in P$. So $bc - ac = (b - a)c \in P$, and hence $ac \leq bc$.

Finally, note that $0 \leq a$ if and only if $a = a - 0 \in P$. □

We can now show that formally real fields are orderable by the following strategy: Show that in every formally real field K, ΣK^2 is a prepositive cone, use Zorn's Lemma to extend this prepositive cone to a positive cone, and then use this positive cone to define an ordering on K. We will break this work up into a sequence of lemmas.

Lemma 3.11. *Suppose K is formally real. Then ΣK^2 is a prepositive cone.*

Proof. (b) and (c) are clear, so it suffices to prove (a). Let $a = a_1^2 + \cdots + a_k^2$ and $b = b_1^2 + \cdots + b_\ell^2$ be elements of ΣK^2. Then

$$a + b = a_1^2 + \cdots + a_k^2 + b_1^2 + \cdots + b_\ell^2 \in \Sigma K^2,$$

and

$$ab = (a_1^2 + \cdots + a_k^2)(b_1^2 + \cdots + b_\ell^2) = \sum_{i=1}^{k} \sum_{j=1}^{\ell} (a_i b_j)^2 \in \Sigma K^2.$$ □

Lemma 3.12. *Let P be a prepositive cone on K, and suppose $a \in K$ is such that $-a \notin P$. Then*

$$P[a] = \{x + ya \mid x, y \in P\}$$

is a prepositive cone such that $P \subseteq P[a]$ and $a \in P[a]$.

Proof. We have $P \subseteq P[a]$, since if $x \in P$, then $x = x + 0a \in P[a]$, and $a \in P[a]$, since $a = 0 + 1a \in P[a]$. Now we need to check that $P[a]$ is a prepositive cone.

(a) If $x + ya$ and $x' + y'a$ are in $P[a]$, then

$$(x + ya) + (x' + y'a) = (x + x') + (y + y')a \in P[a]$$
$$(x + ya)(x' + y'a) = (xx' + yy'a^2) + (xy' + x'y)a \in P[a].$$

Note that we used here the fact that $a^2 \in \Sigma K^2 \subseteq P'$.

(b) We have $\Sigma K^2 \subseteq P \subseteq P[a]$.

(c) Suppose for contradiction that $-1 \in P[a]$. Then we have $-1 = x + ya$ for some $x, y \in P$. Since $-1 \notin P$, $y \neq 0$. By Lemma 3.9(1), $y^{-1} \in P$, and we can write $-ya = x + 1$ and $-a = (x+1)y^{-1} \in P$, contradicting our assumption. □

Lemma 3.13. *Let K be a field. Then every prepositive cone $P \subseteq K$ extends to a positive cone $P \subseteq P' \subseteq K$.*

Proof. Let $(\mathcal{P}; \subseteq)$ be the set of all prepositive cones extending P, partially ordered by containment. It is easy to check that the union of a chain of prepositive cones is a prepositive cone, so by Zorn's Lemma there is a maximal prepositive cone P' containing P. It remains to show that P' is a positive cone.

Let $a \in K$ and suppose for contradiction that $a \notin P'$ and $-a \notin P'$. By Lemma 3.12, $P'[a]$ is a prepositive cone such that $P' \subsetneq P'[a]$, since $a \in P'[a]$. This contradicts maximality of P'. So $a \in P'$ or $-a \notin P'$. Since a was arbitrary, P' is a positive cone. □

Theorem 3.14. *Let K be a formally real field. Then there exists an order $\leq$ such that $(K; \leq)$ is an ordered field. Moreover, if $a \in K$ and $-a \notin \Sigma K^2$, then we can choose $\leq$ so that $0 \leq a$.*

Proof. Suppose K is formally real. Let $P = \Sigma K^2$. By Lemma 3.11, P is a prepositive cone on K. Let a be such that $-a \notin P$. If a is not specified, we can take $a = 1$. By Lemma 3.12, $P[a]$ is a prepositive cone, and by Lemma 3.13, $P[a]$ extends to a positive cone $P[a] \subseteq P'$. By Proposition 3.10, $(K; \leq)$ is an ordered field, where $a \leq b$ if and only if $b - a \in P'$. In particular, $0 \leq a$, since $a = a - 0 \in P[a] \subseteq P'$. □

We can also use the theory of positive cones to show that the ordering on an ordered integral domain R extends uniquely to its field of fractions $\mathrm{Frac}(R)$.

Theorem 3.15. *Let $(R; \leq)$ be an ordered integral domain. Let $K = \mathrm{Frac}(R)$. Then there is a unique ordering on K extending the ordering on R, such that K is an ordered field.*

Proof. We define the ordering on K by defining a positive cone on K. Let $P_R = \{a \in R \mid 0 \leq a\}$, and let

$$P = \{a/b \in K \mid a \in P_R \text{ and } b \in P_R\}.$$

Note that elements of K are really equivalence classes of fractions. So, for example, if $c \leq 0$ and $d < 0$ in R, then $c/d \in P_K$ since $c/d = (-c)/(-d)$ and $-c \geq 0$ and $-d > 0$ in R.

We check that P is a positive cone.

(a) Let $a/b, c/d \in P$, with $a, b, c, d \in P_R$. Then $\frac{a}{b} + \frac{c}{d} = \frac{ad+bc}{cd} \in P$, since $(ad + bc) \in P_R$ and $cd \in P_R$. And $\left(\frac{a}{b}\right)\left(\frac{c}{d}\right) = \frac{ab}{cd} \in P$, since $ab \in P_R$ and $cd \in P_R$.

(b) It suffices to check that every square is in P, since closure under addition then implies every sum of squares is in P. So let $a/b \in K$. Then $\left(\frac{a}{b}\right)^2 = \frac{a^2}{b^2} \in P$ since $a^2 \in P_R$ and $b^2 \in P_R$.

(c) To show $-1 \notin P$, we need to consider every fraction equivalent to -1. If $\frac{a}{b} = -1$, then $a = -b$ and $b \neq 0$. So if $-1 \in P$, then there is some $b \in R$ with $b \neq 0$ such that $b \in P_R$ and $-b \in P_R$. But then $b \geq 0$ and $-b \geq 0$ in R, which implies $b \leq 0$, and hence $b = 0$, contradiction.

(d) Suppose $\frac{a}{b} \in K$. If $0 \leq a$ and $0 < b$ in R, then $\frac{a}{b} \in P$. And as noted above, if $a \leq 0$ and $b < 0$ in R, then $\frac{a}{b} \in P$. So we may assume $0 \leq a$ and $b < 0$ or $a \leq 0$ and $0 < b$. In the first case, $-\frac{a}{b} = \frac{a}{-b}$, and $a, -b \in P_R$, and in the second case, $-\frac{a}{b} = \frac{-a}{b}$, and $-a, b \in P_R$. In either case, $-\frac{a}{b} \in P$.

So P induces an ordering on K by Proposition 3.10. To see that this ordering extends the ordering on R, note that if $a \leq b$ in R, then $(b - a) \geq 0$, so $\frac{b}{1} - \frac{a}{1} = \frac{b-a}{1} \in P$, and hence $\frac{a}{1} \leq \frac{b}{1}$ in K. Conversely, suppose $\frac{a}{1} \leq \frac{b}{1}$ in K, with $a, b \in R$. Then $\frac{b-a}{1} \in P$, so there is some $p, q \in P_R$ with $\frac{b-a}{1} = \frac{p}{q}$. It follows that $(b-a)q = p$. Suppose for contradiction that $(b-a) < 0$ in R. Then since $q > 0$, we have $p = (b-a)q < 0q = 0$, contradicting $p \geq 0$. Thus $b - a \geq 0$ in R, and hence $a \leq b$ in R.

For uniqueness, let $\leq'$ be another ordering on K extending the ordering on R, and let P' be the positive cone associated to $\leq'$. Let $a, b \in P_R$ with $b \neq 0$, so $0 \leq a$ and $0 \leq b$ in R. Then also $0 \leq' a$ and $0 \leq' b$ in K, so $a, b \in P'$. By Lemma 3.9(1), $b^{-1} \in P'$, so $ab^{-1} = a/b \in P'$. This shows $P \subseteq P'$. Now we want to prove that every positive cone is maximal. Concretely, let $c \in P'$, and assume for contradiction that $c \notin P$. Since P is a positive cone, $-c \in P \subseteq P'$, so c and $-c$ are both in P'. By Lemma 3.9, $c = 0 \in P$, contradiction. So $P = P'$. □

Note that to define the order on K directly in above theorem would have been a bit of a mess. We have $\frac{a}{c} \leq \frac{b}{d}$ if and only if $0 \leq \frac{b}{d} - \frac{a}{c} = \frac{bc-ad}{cd}$, and then one has to consider the relative signs of $bc - ad$ and cd.

Exercise 15. Show that if $(R; \leq)$ is an ordered ring which is not an integral domain, then there is some $a \in R$ such that $a^2 = 0$.

This motivates the following construction of an ordered ring which is not an integral domain. The ring $\mathbb{R}[\varepsilon]/(\varepsilon^2)$ is sometimes called the ring of "dual numbers", and it is relevant for the algebraic study of formal differentiation. We think of ε as an "infinitesimal" element in this ring, from which point of view the ordering suggests itself naturally,

Exercise 16. Show that the order on $R = \mathbb{R}[\varepsilon]/(\varepsilon^2)$ defined by $a + b\varepsilon \leq c + d\varepsilon$ if and only if $a < c$ or ($a = c$ and $b \leq d$) makes R into an ordered ring.

3.2 Real closed fields

The algebraically closed fields are the "richest" fields, in the sense that they contain roots to all non-constant polynomials. We'd like to isolate a similar class of the "richest" formally real (orderable) fields, and intuitively $\mathbb{R}$ should be an example. But formally real fields do not have roots to all polynomials (e.g., $x^2 = 1$), and it is a bit tricky to say exactly which polynomials should have roots.

So instead, we take inspiration from an equivalent characterization of algebraically closed fields: a field is algebraically closed if it has no proper algebraic extensions. This "maximality" condition can be easily generalized by restricting to the class of formally real fields, leading to the following definition.

Definition 3.16. A field is **real closed** if it is formally real and has no proper formally real algebraic extensions.

Example 3.17. $\mathbb{R}$ is real closed. Indeed, if $\mathbb{R} \subseteq K$ is an algebraic extension, then we may assume $\mathbb{R} \subseteq K \subseteq \mathbb{C}$, since $\mathbb{C}$ is algebraically closed. But then $[\mathbb{C} : K][K : \mathbb{R}] = [\mathbb{C} : \mathbb{R}] = 2$, so either $[K : \mathbb{R}] = 1$ and $K = \mathbb{R}$ (in which case the extension is not proper), or $[\mathbb{C} : K] = 1$ and $K = \mathbb{C}$ (in which case K is not formally real).

Soon, we will see that there are many equivalent and natural ways to define real closed fields. But this definition has the nice consequence that makes it easy to construct real closures.

Definition 3.18. Let K be a formally real field. Then a field R is a **real closure** of K if $K \subseteq R$ is an algebraic extension and R is real closed.

Theorem 3.19. *Every formally real field has a real closure.*

Proof. Let K be formally real, and let C be an algebraic closure of K. Let

$$\mathcal{R} = \{R \subseteq C \mid K \subseteq R \text{ and } R \text{ is formally real}\}.$$

We view $\mathcal{R}$ as a poset, partially ordered by inclusion. For Zorn's Lemma, suppose $(R_i)_{i \in I}$ is a chain in $\mathcal{R}$, and let $R_\infty = \bigcup_{i \in I} R_i$. Then $K \subseteq R_\infty \subseteq C$, and R_∞ is formally real: If $-1 = \sum_{j=1}^{n} a_j^2$ with $a_j \in R_\infty$ for all j, then there is some $i \in I$ such that $a_j \in R_i$ for all $1 \leq j \leq n$, contradicting the fact that R_i is formally real.

By Zorn's Lemma $\mathcal{R}$ has a maximal element R. Then R is a formally real algebraic extension of K, and it remains to show that R is real closed. Suppose $R \subseteq R'$ is an algebraic extension and R' is formally real. Since C is algebraically closed, we can embed R' in C over R, so that $K \subseteq R \subseteq R' \subseteq C$. Since R is maximal among real closed extensions of K contained in C, we have $R = R'$. This completes the proof. □

To understand real closed fields more concretely, we need to understand which algebraic extensions of formally real fields are formally real.

Lemma 3.20. *Suppose K is formally real and $c \in K$ is such that $-c$ is not a sum of squares. Then $K(\sqrt{c})$ is formally real. In particular, for all $a \in K$, either $K(\sqrt{a})$ or $K(\sqrt{-a})$ is formally real.*

Proof. If c is a square in K, then $K(\sqrt{c}) = K$ is formally real. If not, the the elements of $K(\sqrt{c})$ can be written uniquely in the form $a + b\sqrt{c}$ with $a, b \in K$. Suppose for contradiction that $K(\sqrt{c})$ is not formally real. Then

$$\begin{aligned} -1 &= \sum_{i=1}^{n} (a_i + b_i\sqrt{c})^2 \\ &= \sum_{i=1}^{n} (a_i^2 + 2a_i b_i \sqrt{c} + b_i^2 c) \\ &= \sum_{i=1}^{n} (a_i^2 + b_i^2 c) + \left(\sum_{i=1}^{n} 2a_i b_i \right) \sqrt{c} \end{aligned}$$

It follows that $-1 = \left(\sum_{i=1}^{n} a_i^2\right) + \left(\sum_{i=1}^{n} b_i^2\right) c$. Note that $\left(\sum_{i=1}^{n} b_i^2\right)^{-1} \neq 0$, since otherwise -1 is a sum of squares in K. Rearranging gives

$$-c = \left(1 + \sum_{i=1}^{n} a_i^2\right) \left(\sum_{i=1}^{n} b_i^2\right)^{-1}.$$

That is, $-c$ is a quotient of sums of squares. But recall that since K is formally real, ΣK^2 is a prepositive cone (Lemma 3.11), and prepositive cones are closed under addition and multiplication and inverses of nonzero elements (the last by Lemma 3.9(1)). So $-c$ is a sum of squares, contradicting our assumption.

For the in particular clause, let $a \in K$. Since K is formally real, ΣK^2 is a prepositive cone in K and by Lemma 3.9(2), if $a \in \Sigma K^2$ and $-a \in \Sigma K^2$, then $a = 0$, and $K(\sqrt{a}) = K(\sqrt{-a}) = K$ is formally real. If not, then either $-a$ or a is not a sum of squares, so either $K(\sqrt{a})$ or $K(\sqrt{-a})$ is formally real. □

Lemma 3.21. *Suppose K is formally real and $f(x) \in K[x]$ is an irreducible polynomial of odd degree. Then letting α be a root of f, $K(\alpha)$ is formally real.*

Proof. Suppose not. Then there is an irreducible polynomial f of *minimal* odd degree n such that $K(\alpha) = K[x]/(f)$ is not formally real. Note that $n > 1$, since if $\deg(f) = 1$, then $K(\alpha) = K$.

The elements of $K(\alpha)$ can be uniquely represented as $g(\alpha)$ where $g \in K[x]$ is a polynomial of degree at most $(n-1)$. So we can write $-1 = \sum_{i=1}^{m} (g_i(\alpha))^2$.

Lifting to $K[x]$, we have $-1 = \sum_{i=1}^{m} g_i^2 + hf$ for some polynomial h. First note that $h \neq 0$. Indeed, if $h = 0$, then we have $-1 = \sum_{i=1}^{m} g_i^2$ in $K[x]$, and evaluating at 0, $-1 = \sum_{i=1}^{m} (g_i(0))^2$ in K, contradicting the fact that K is formally real.

Since $h \neq 0$, $\deg(hf) = \deg(h) + \deg(f)$, and the leading term of $\sum_{i=1}^{m} g_i^2$ must cancel the leading term of hf, so $\deg(h) + n = \deg\left(\sum_{i=1}^{m} g_i^2\right)$. Since each

g_i has $\deg(g_i) \leq n-1$, $\deg\left(\sum_{i=1}^m g_i^2\right) \leq 2n-2$, and hence $\deg(h) \leq n-2$. Further, $\deg\left(\sum_{i=1}^m g_i^2\right)$ is even and n is odd, so $\deg(h)$ is odd.

Let h' be an irreducible factor of h of odd degree (if a polynomial has only even degree irreducible factors, it must have even degree), and let β be a root of h' (so β is also a root of h). Consider the equation $-1 = \sum_{i=1}^m g_i^2 + hf$, and evaluate at β in $K(\beta)$. We have $-1 = \sum_{i=1}^m (g_i(\beta))^2 + h(\beta)f(\beta) = \sum_{i=1}^m (g_i(\beta))^2$, demonstrating that $K(\beta)$ is not formally real. But then h' contradicts the minimality of the degree of f among irreducible polynomials of odd degree such that the field obtained by adjoining a root to K fails to be formally real. □

Theorem 3.22 (Artin–Schreier)**.** *Let R be a field. The following are equivalent:*

(1) R is real closed.

(2) R is formally real, every odd degree polynomial in $R[x]$ has a root in R, and for all $a \in R$, either a or $-a$ is a square in R.

(3) R is not algebraically closed, but $R(\sqrt{-1})$ is algebraically closed.

Proof. Let's agree to write i instead of $\sqrt{-1}$ in this proof.

(1)⇒(2): Suppose R is real closed. Then R is formally real by definition.

Let $f \in R[x]$ be a polynomial of odd degree, let g be an irreducible factor of f of odd degree, and let α be a root of g. By Lemma 3.21, $R(\alpha)$ is a formally real algebraic extension of R. Since R is real closed, $R = R(\alpha)$, and hence f has a root in R.

Now let $a \in R$. By Lemma 3.20, either $R(\sqrt{a})$ or $R(\sqrt{-a})$ is a formally real algebraic extension of R. Since R is real closed, $R = R(\sqrt{a})$ or $R = R(\sqrt{-a})$, so either a or $-a$ is a square in R.

(2)⇒(3): Since R is formally real, there is an ordering $\leq$ making R into an ordered field. For all $a \in R$, if $0 < a$, then $-a < 0$ and is not a square. Hence a is a square by our assumption. So a is a square in R if and only if $0 \leq a$. In particular, -1 is not a square, so R is not algebraically closed.

We now prove that $R(i)$ is algebraically closed. To start, we show that every element of $R(i)$ is a square. Indeed, if $a + bi \in R(i)$, then $(a+bi) = (c+di)^2$, where

$$c = \sqrt{\frac{a + \sqrt{a^2+b^2}}{2}} \quad \text{and} \quad d = \pm\sqrt{\frac{-a + \sqrt{a^2+b^2}}{2}},$$

where d is positive if $b \geq 0$ and negative if $b \leq 0$. When we write $\sqrt{x}$, we need to verify that $x \geq 0$, and we mean to take the positive square root (if $y^2 = x$, then also $(-y)^2 = x$, and either $y \geq 0$ or $-y \geq 0$). We will check in a moment that these formulas make sense. But first, let's compute:

$$\begin{aligned}
(c+di)^2 &= c^2 - d^2 + 2cdi \\
&= \frac{a + \sqrt{a^2+b^2}}{2} - \frac{-a + \sqrt{a^2+b^2}}{2} \pm 2i\sqrt{\frac{(a + \sqrt{a^2+b^2})(-a + \sqrt{a^2+b^2})}{4}} \\
&= a \pm i\sqrt{-a^2 + (a^2+b^2)} \\
&= a + bi,
\end{aligned}$$

since if $b \geq 0$, then $\sqrt{b^2} = b$ and the sign on d is $+$, while if $b \leq 0$, then $\sqrt{b^2} = -b$ and the sign on d is $-$.

Now to check that the the formulas for c and d actually describe elements of R. The square root of $a^2 + b^2$ exists, since $a^2 + b^2 \geq 0$. It remains to show that $\sqrt{a^2 + b^2} \pm a \geq 0$. Indeed, being formally real, R has characteristic 0, so we can multiply by $2^{-1} \geq 0$, which preserves positivity.

If $a = 0$, we're done, since $\sqrt{b^2} \geq 0$ by definition of the square root. If $a \neq 0$, we may assume $a > 0$ without loss of generality. We have

$$(\sqrt{a^2 + b^2} - a)(\sqrt{a^2 + b^2} + a) = (a^2 + b^2) - a^2 = b^2 \geq 0.$$

Now $0 < \sqrt{a^2 + b^2} + a$, since the right hand side is a sum of two positive elements of R. This was one of the inequalities we had to show. And it means we can multiply both sides of the inequality by $(\sqrt{a^2 + b^2} + a)^{-1}$, obtaining $\sqrt{a^2 + b^2} - a \geq 0$, which was the other inequality we had to show.

It follows that every quadratic polynomial $ax^2 + bx + c$ over $R(i)$ has a root in $R(i)$, by the quadratic formula, since the discriminant $(b^2 - 4ac)$ is a square in $R(i)$. So $R(i)$ has no algebraic extensions of degree 2. Our strategy for the rest of the proof is to show that if $R(i)$ has any proper algebraic extension, then it has one of degree 2, deriving a contradiction. To do this, we need some Galois theory.

Suppose $R(i) \subsetneq K$ is a proper finite extension. Then also $R \subseteq K$ is a finite extension of degree $2[K : R(i)]$. Extending further if necessary, we may assume K is Galois over R. Let $G = \mathrm{Gal}(K/R)$. Let H be a 2-Sylow subgroup of G, so $|G|/|H|$ is odd. Let F be the fixed field of H, so $R \subseteq F \subseteq K$. Then $[F : R] = |G|/|H|$ is odd. By the primitive element theorem, there is some $\alpha \in F$ such that $F = R(\alpha)$, and the minimal polynomial $f(x)$ of α over R is of odd degree. But then f has a root in R, and since f is irreducible over R, it follows that $\deg(f) = 1$, and hence $F = R$. But then $G = H$ is a 2-group (has order a power of 2).

Now the Galois group $G' = \mathrm{Gal}(K/R(i))$ is a subgroup of G and hence also a 2-group. It is a fact that every 2-group contains a subgroup of index 2, so we can find $H' \subseteq G'$ of index 2. Letting F' be the fixed field of H', we have $R(i) \subsetneq F' \subseteq K$, and $[F' : R(i)] = 2$. But this contradicts the fact, established above, that $R(i)$ has no algebraic extension of degree 2.

(3)⇒(1): First, we show that R is formally real. To do this, we show that every sum of squares in R is a square. So let $a, b \in R$. Since $R(i)$ is algebraically closed, $a + bi$ is a square in $R(i)$, so there exist $c, d \in R$ such that $(a + bi) = (c + di)^2 = (c^2 - d^2) + 2cdi$. It follows that $a = c^2 - d^2$ and $b = 2cd$. But then $a^2 + b^2 = (c^2 - d^2)^2 + (2cd)^2 = c^4 - 2c^2d^2 + d^4 + 4c^2d^2 = c^4 + 2c^2d^2 + d^4 = (c^2 + d^2)^2$. It follows that an arbitrary sum of squares in R is a square. Now since $R \neq R(i)$, -1 is not a square in R, and hence not a sum of squares in R.

Having shown that R is formally real, the fact that it is has no proper formally real algebraic extension follows by the same argument as in Example 3.17, with R in place of $\mathbb{R}$ and $R(i)$ in place of $\mathbb{C}$. □

The statement of the Artin–Schreier theorem often includes a fourth equivalent condition: the absolute Galois group $\mathrm{Gal}(\overline{R}/R)$ is finite and non-trivial (so as a consequence of the theorem, it turns out that if the absolute Galois group of a field is finite and nontrivial, it must be cyclic of order 2, since already $R(\sqrt{-1})$ is algebraically closed). The proof of this equivalence requires some more heavy lifting, so we'll omit it.

It is also interesting to note that the proof of the implication (2)⇒(3) is probably as close as it's possible to get to a purely algebraic proof of the fundamental theorem of algebra. Many proofs have a crucial topological or complex analytic component. And if we define the complex numbers as $\mathbb{C} = \mathbb{R}(i)$ and we want to show that $\mathbb{C}$ is algebraically closed, we need to use *some* basic facts about the real field. Condition (2) pares down the topological / analytic assumptions about $\mathbb{R}$ to some very innocuous properties that can be verified using the intermediate value theorem.

Here are some further examples of real closed fields.

Example 3.23. The field of **real algebraic numbers**

$$\mathbb{Q}^r = \overline{\mathbb{Q}} \cap \mathbb{R} = \{r \in \mathbb{R} \mid r \text{ is algebraic over } \mathbb{Q}\}$$

is a real closure of $\mathbb{Q}$. It is easy to verify that condition (2) of the Artin–Schreier theorem holds for $\mathbb{Q}^r$.

Example 3.24. For a field K, we define by $K((t))$ the field of **formal Laurant series** over K:

$$K((t)) = \left\{ \sum_{i=-k}^{\infty} a_i t^i \mid k \in \mathbb{N}, a_i \in K \text{ for all } i \geq -k \right\}.$$

That is, a formal Laurant series is a power series with arbitrary integer powers, in which the powers appearing are bounded below.

The field of **Puiseux series** over K is

$$K\langle\langle t \rangle\rangle = \bigcup_{n \geq 1} K((t^{1/n})) = \left\{ \sum_{i=-k}^{\infty} a_{i/n} t^{i/n} \mid k \in \mathbb{N}, a_{i/n} \in K \text{ for all } i \geq -k \right\}.$$

That is, a Puiseux series over $\mathbb{R}$ is a power series with arbitrary rational powers, in which the powers appearing around bounded below and have bounded denominator.

The reason for these restriction on the powers of t is so that multiplication is well-defined. Recall that in the product

$$\left(\sum_{i=-k}^{\infty} a_{i/n} t^{i/n} \right) \left(\sum_{j=-\ell}^{\infty} a_{j/m} t^{j/m} \right),$$

the coefficient of t^q is

$$\sum_{i/n + j/m = q} a_{i/n} b_{j/m}.$$

The fact that there are no infinite descending sequences in the possible values of i/n or j/n implies that this sum is finite.

When $K = \mathbb{R}$ (or any real closed field), the Puiseux series field $K\langle\langle t\rangle\rangle$ is real closed, and it is a real closure of the field of formal Laurant series $K((t))$. This can also be verified by checking that condition (2) of the Artin–Schreier theorem holds, but the verification is significantly harder in this case!

Example 3.25. The field of **Hahn series** over K is a generalization of the field of Puiseux series over K, which weakens the requirements on the powers of t as far as possible while ensuring that multiplication is well-defined. For an arbitrary ordered abelian group Γ, we define

$$K[[t^\Gamma]] = \left\{ \sum_{e\in\Gamma} a_e t^e \mid a_e \in K \text{ for all } e \in \Gamma, \text{and } \{e \in \Gamma \mid a_e \neq 0\} \text{ is well-ordered} \right\}.$$

The set $\{e \in \Gamma \mid a_e \neq 0\}$ is called the **support** of the Hahn series, and to say it is well-ordered means that it contains no infinite descending sequence.

When $K = \mathbb{R}$ (or any real closed field) and Γ is a divisible group (for example, when $\Gamma = \mathbb{Q}$ or $\Gamma = \mathbb{R}$), the Hahn series field $K[[t^\Gamma]]$ is real closed.

We have $\mathbb{R}\langle\langle t\rangle\rangle \subsetneq \mathbb{R}[[t^{\mathbb{Q}}]]$, for example

$$t^{-1} + t^{-1/2} + t^{-1/3} + t^{-1/4} + \dots$$

is a Hahn series but not a Puiseux series, since the support $\{-1/n \mid n \in \mathbb{N}\}$ is well-ordered, but the there is no bound on the denominators.

In all the examples of real closed fields above, we obtain an algebraically closed field by adjoining i, which has the effect of changing the field of coefficients in a power series field. So for example, the algebraic closures of $\mathbb{R}\langle\langle t\rangle\rangle$ and $\mathbb{R}[[t^{\mathbb{Q}}]]$ are $\mathbb{C}\langle\langle t\rangle\rangle$ and $\mathbb{C}[[t^{\mathbb{Q}}]]$, respectively.

3.3 Real closed ordered fields

We will now consider properties of real closed fields as ordered fields.

Proposition 3.26. *If R is a real closed field, then R admits a unique ordering, defined by $a \leq b$ if and only if $(b - a)$ is a square.*

Proof. Since R is formally real, it admits some ordering. Let $a, b \in R$. Then either $(b - a)$ or $-(b - a) = (a - b)$ is a square. If $(b - a)$ is a square, then for any ordering $\leq$ on R, we have $0 \leq b - a$, so $a \leq b$. The other case is similar. So there is a unique ordering on R. □

Let $(K, \leq)$ be an ordered field. We know that K is formally real, and thus K has a real closure $K \subseteq R$, which admits a unique ordering. Does the ordering on R extend the ordering on K?

The answer is no, not necessarily. But we can choose R to ensure that it does. For example, we have seen in Example 3.4 that $\mathbb{Q}(\sqrt{2})$ admits two

orderings, the "standard" one, in which $0 < \sqrt{2}$, and a "nonstandard" one in which $0 < -\sqrt{2}$. This is also a consequence of Theorem 3.14: neither $\sqrt{2}$ nor $-\sqrt{2}$ is a sum of squares in this field. The field $\mathbb{Q}^r$ of real algebraic numbers is a real closure of $\mathbb{Q}(\sqrt{2})$, and the unique ordering on $\mathbb{Q}^r$ extends the standard ordering on $\mathbb{Q}(\sqrt{2})$, but not the non-standard one.

On the other hand, letting $(K, \leq)$ be $\mathbb{Q}(\sqrt{2})$ with the nonstandard ordering, if we can turn $-\sqrt{2}$ into a square by moving to $K(\sqrt{-\sqrt{2}})$, then any real closure of the latter field must be ordered so that $0 < -\sqrt{2}$. This is the strategy we will use, but adjoining square roots to *all* positive elements.

Lemma 3.27. *Let $(K, \leq)$ be an ordered field, and let L be the field obtained by adjoining square roots to all positive elements of K. Then L is formally real.*

Proof. We can view L as the union of all fields of the form $K(\sqrt{a_1}, \ldots, \sqrt{a_k})$, with $k \geq 0$ and $0 < a_i$ in K for all $1 \leq i \leq k$. If $-1 = \sum_{j=1}^{n} b_j^2$ is a sum of squares in L, then there is some field $K(\sqrt{a_1}, \ldots, \sqrt{a_k})$ such that all b_j are in this field. Thus it suffices to prove that any field of the form $K(\sqrt{a_1}, \ldots, \sqrt{a_k})$ is formally real.

We will prove a stronger claim by induction on k: For any such field $K' = K(\sqrt{a_1}, \ldots, \sqrt{a_k})$, -1 is not a positive K-linear combination of squares in K'. That is, there are no $b_1, \ldots, b_k \in K'$ and $c_1, \ldots, c_k \in K$ with $0 \leq c_i$ for all $1 \leq i \leq k$, such that $-1 = \sum_{i=1}^{k} c_i b_i^2$. Note that a sum of squares is a positive K-linear combination of squares, since we can take all the c_i to be equal to 1.

The base case is handled by the fact that K is an ordered field: Any positive K-linear combination of squares in K is non-negative, thus is not equal to -1.

Now suppose the claim is true for $K' = K(\sqrt{a_1}, \ldots, \sqrt{a_k})$, and consider $K'(\sqrt{a_{k+1}})$. If a_{k+1} is a square in K', then $K'(\sqrt{a_{k+1}}) = K'$, and we are done. Otherwise, the elements of $K'(\sqrt{a_{k+1}})$ can be written uniquely in the form $d + e\sqrt{a_{k+1}}$, with $d, e \in K'$. Suppose for contradiction that

$$\begin{aligned} -1 &= \sum_{i=1}^{n} c_i (d_i + e_i \sqrt{a_{k+1}})^2 \\ &= \left(\sum_{i=1}^{n} c_i d_i^2 + c_i e_i^2 a_{k+1} \right) + \left(\sum_{i=1}^{n} 2 c_i d_i e_i \right) \sqrt{a_{k+1}} \end{aligned}$$

where $0 \leq c_i \in K$ and $d_i, e_i \in K'$ for all $1 \leq i \leq n$. Then

$$-1 = \sum_{i=1}^{n} (c_i d_i^2 + (c_i a_{k+1}) e_i^2),$$

and since $0 \leq c_i a_{k+1} \in K$ for all i, this is a positive K-linear combination of squares, contradicting the inductive hypothesis. □

Theorem 3.28. *Let $(K, \leq)$ be an ordered field. Then there is a real closure $K \subseteq R$ such that the unique ordering on R extends the ordering on K.*

Proof. Let L be the field obtained by adjoining square roots to all positive elements of K. By Lemma 3.27, L is formally real. By Theorem 3.19, L has a real closure R, which is also a real closure of K, since $K \subseteq L \subseteq R$ is an algebraic extension. If $a \leq b$ in K, then $0 \leq (b-a)$ in K, so $(b-a)$ is a square in L and hence in R, and thus $a \leq b$ in R. So the ordering on R extends the ordering on K, as desired. □

It follows that formally real fields which admit multiple orderings also admit multiple non-isomorphic real closures. More precisely, if $\leq_1$ and $\leq_2$ are distinct orderings of K, then Theorem 3.28 provides real closures $K \subseteq R_1$ and $K \subseteq R_2$ such that the unique orderings on R_1 and R_2 extend the orderings $\leq_1$ and $\leq_2$, respectively. Then there is no isomorphism $\sigma\colon R_1 \cong R_2$ such that σ fixes K. Indeed, letting $a \neq b$ be elements such that $a \leq_1 b$ but $b \leq_2 a$, we have that $(b-a) \in K$ is a square in R_1 but not in R_2.

On the other hand, once we fix an ordering on K, the real closure provided by Theorem 3.28 is unique up to isomorphism over K. Our next goal is to prove this. First, we give an additional characterization of real closed fields.

Definition 3.29. Let R be an ordered field. We say R satisfies the **intermediate value theorem for polynomials** (IVTP) if for all $f \in R[x]$ and $a, b \in R$ such that $a < b$ and such that $f(a)$ and $f(b)$ have opposite signs (meaning $f(a) < 0 < f(b)$ or $f(a) > 0 > f(b)$), there exists $c \in R$ such that $a < c < b$ and $f(c) = 0$.

Lemma 3.30. *Let K be an ordered field, and let $f = c_0 + c_1 x + \cdots + x^n \in K[x]$ be a monic polynomial. If $b > \max(1, |c_0| + \cdots + |c_{n-1}|)$, then $f(b) > 0$. If n is even, $f(-b) > 0$, and if n is odd, then $f(-b) < 0$. In particular if $f(a) = 0$, then $-b < a < b$.*

Proof. We have

$$\begin{aligned} b^n &> (|c_0| + \cdots + |c_{n-1}|)b^{n-1} \\ &\geq |c_0| + |c_1|b + \cdots + |c_{n-1}|b^{n-1} \\ &\geq -c_0 + -c_1 b + \cdots + -c_{n-1}b^{n-1}, \end{aligned}$$

where the second inequality follows from the fact that $b^i < b^j$ for $i < j$, since $b > 1$. It follows that $f(b) > 0$.

Now we can apply the same argument to the polynomial $f(-x)$ (which is monic if n is even) or $-f(-x)$ (which is monic if n is odd). Since the coefficients of $f(-x)$ and $-f(-x)$ agree with those of f up to sign, we also have $f(-b) > 0$ if n is even, and $-f(-b) > 0$, so $f(-b) < 0$, if n is odd.

Now increasing b or decreasing $-b$ retains the inequality, so any root of f is bounded between $-b$ and b. □

Theorem 3.31. *Let R be an ordered field. Then R is real closed if and only if R satisfies IVTP.*

Proof. Suppose R is real closed. Let $f \in R[x]$ and $a, b \in R$ with $a < b$ and $f(a) < 0 < f(b)$. We can factor f as a constant times a product of monic irreducible polynomials. If for every irreducible factor g of f, $g(a)$ and $g(b)$ have the same sign, then $f(a)$ and $f(b)$ have the same sign. Since this is not true, we may pick a monic irreducible factor g of f, such that $g(a)$ and $g(b)$ have opposite signs. It suffices to find a root of g between a and b.

Now since $R(i)$ is algebraically closed, g is linear or quadratic. If g is linear, then $g = x - c$, and c is the unique root of g. Since g is increasing, we have $g(a) < 0 < g(b)$, so $a - c < 0 < b - c$, and $a < c < b$.

If g is quadratic, then $g = x^2 + dx + e$, and since g has no root in R, the discriminant $d^2 - 4e < 0$, so $e - \frac{d^2}{4} > 0$. But then, completing the square, we can write $g(z) = (z + \frac{d}{2})^2 + (e - \frac{d^2}{4}) > 0$ for all $z \in R$, contradicting the fact that $g(a)$ and $g(b)$ have opposite signs.

Conversely, suppose R satisfies IVTP. Since R is an ordered field, it is formally real. We show that every odd degree polynomial in $R[x]$ has a root in R, and for all $a \in R$, either a or $-a$ is a square.

So suppose f is an odd degree polynomial in $R[x]$. Dividing by the leading coefficient of f, we may assume f is monic. By Lemma 3.30, there exists $b > 0$ such that $f(b) > 0$ and $f(-b) < 0$. Since $-b < 0 < b$, by IVTP, f has a root between $-b$ and b.

Now suppose $a \in R$. Since a or $-a$ is non-negative, it suffices to assume $0 < a$ and show a is a square (we already know 0 is a square). Consider the polynomial $f = x^2 - a$. We have $f(0) = -a < 0$ and $f(1+a) = (1+a)^2 - a = 1 + a + a^2 > 0$. Since $0 < 1 + a$, by IVTP, f has a root between 0 and $1 + a$. □

Our goal is to show that if R_1 and R_2 are two real closures of an ordered field $(K; \leq)$, whose unique orders extend the ordering on K, then R_1 and R_2 are isomorphic over K. The key ingredient we need to prove this is to show that any irreducible polynomial in $K[x]$ which has a root in R_1 also has a root in R_2. That is, we need to be able to determine, for any irreducible polynomial $f \in K[x]$, whether it has a root in a real closure of K, just by looking at f and $K[x]$. One method for determining this is called Sturm's algorithm.

Definition 3.32. Let $(K; \leq)$ be an ordered field, and let $f \in K[x]$ be a polynomial. A **Sturm sequence** for f is a sequence $f_0, f_1, \ldots, f_n \in K[x]$, such that:

(a) $f_0 = f$ and $f_1 = f'$, the formal derivative of f.

(b) For all $\alpha \in K$ and all $0 \leq i \leq n - 1$, we do not have $f_i(\alpha) = f_{i+1}(\alpha) = 0$.

(c) For all $\alpha \in K$ and all $1 \leq i \leq n - 1$, if $f_i(\alpha) = 0$, then $f_{i-1}(\alpha)$ and $f_{i+1}(\alpha)$ have opposite signs.

(d) f_n is a non-zero constant polynomial.

Proposition 3.33. *Suppose f is a non-constant polynomial which does not have multiple roots (in an algebraic closure of K). Then there exists a Sturm sequence for f.*

Proof. Let $f_0 = f$ and $f_1 = f'$, the formal derivative of f. Recall that since f does not have multiple roots, f and f' do not share any roots (in an algebraic closure of K), and hence have no common non-constant polynomial factors.

If f' is a constant, then f, f' is already a Sturm sequence with $n = 1$: f and f' do not share any roots, (c) is satisfied vacuously, and $f' \neq 0$, since K has characteristic 0. So assume f' is non-constant.

For $i \geq 1$, we define f_{i+1} by induction. Divide f_{i-1} by f_i in order to write $f_{i-1} = q_i f_i - f_{i+1}$, with $\deg(f_{i+1}) < \deg(f_i)$. Since degrees decrease in the sequence, we eventually arrive at a constant polynomial f_n.

Let's check that this is a Sturm sequence.

(b) In the base case, we have that $f_0 = f$ and $f_1 = f'$ do not share any roots (even in an algebraic closure of K). Suppose for contradiction that f_i and f_{i+1} share a root $\alpha \in K$, with $i \geq 1$. Then $f_{i-1}(\alpha) = q_i(\alpha) f_i(\alpha) - f_{i+1}(\alpha) = 0$, so f_{i-1} and f_i also share the root α, contradicting the inductive hypothesis.

(c) If $f_i(\alpha) = 0$, then $f_{i-1}(\alpha) = q_i(\alpha) f_i(\alpha) - f_{i+1}(\alpha)$ implies $f_{i-1}(\alpha) = -f_{i+1}(\alpha)$, so $f_{i-1}(\alpha)$ and $f_{i+1}(\alpha)$ have opposite signs.

(d) By construction, f_n is a constant, and $\deg(f_{n-1}) > \deg(f_n) = 0$. Suppose for contradiction that $f_n = 0$. Then f_{n-1} divides f_{n-2}, and by induction f_{n-1} divides f_i for all i: we have $f_{i-1} = q_i f_i - f_{i+1}$, so if f_{n-1} divides both f_i and f_{i+1}, then f_{n-1} divides f_{i-1}. But then f_{n-1} divides both f and f', contradicting the fact that they share no common non-constant polynomial factors. □

If $f_0, f_1, \ldots, f_n$ is a Sturm sequence for f, then for $c \in K$ such that $f_i(c) \neq 0$ for all $0 \leq i \leq n$, we define $v(c)$ to be the number of sign changes in the sequence $f_0(c), f_1(c), \ldots, f_n(c)$.

Example 3.34. Let $f = x^3 - x$. Then the proof of Proposition 3.33 provides the following Sturm sequence for f:

$$
\begin{aligned}
f_0 &= f = x^3 - x \\
f_1 &= f' = 3x^2 - 1 \\
f_2 &= \frac{2}{3}x \quad \text{since } x^3 - x = \left(\frac{1}{3}x\right)\left(3x^2 - 1\right) - \left(\frac{2}{3}x\right) \\
f_3 &= 1 \quad \text{since } 3x^2 - 1 = \left(\frac{9}{2}x\right)\left(\frac{2}{3}x\right) - 1
\end{aligned}
$$

The zeros of the polynomials in the Sturm sequence occur at -1, $-\frac{1}{\sqrt{3}}$, 0, $\frac{1}{\sqrt{3}}$, and 1. The following table displays the signs of these polynomials at these zeros and at points between them. We can also verify from the table that this

is a Sturm sequence for f.

	c_0	-1	c_1	$-\frac{1}{\sqrt{3}}$	c_2	0	c_3	$\frac{1}{\sqrt{3}}$	c_4	1	c_5
f_0	$-$	0	$+$	$+$	$+$	0	$-$	$-$	$-$	0	$+$
f_1	$+$	$+$	$+$	0	$-$	$-$	$-$	0	$+$	$+$	$+$
f_2	$-$	$-$	$-$	$-$	$-$	0	$+$	$+$	$+$	$+$	$+$
f_3	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$	$+$
$v(c)$	3		2		2		1		1		0

Note that v "counts" zeros of f_0: v drops by one between c_i and c_{i+1} exactly when there is a zero of f_0 in the interval (c_i, c_{i+1}). And the total number of zeros of f_0 between c_0 and c_5 is $v(c_0) - v(c_5) = 3 - 0 = 3$.

We would like to prove that the phenomenon observed in Example 3.34 holds in general. We need the following bit of "algebraic calculus".

Lemma 3.35. *Suppose f is a polynomial with $f(z) = 0$ and $f'(z) \neq 0$. Then for all sufficiently small $\varepsilon > 0$, we have that if $f'(z) > 0$, then $f(z-\varepsilon) < 0 < f(z+\varepsilon)$, and if $f'(z) < 0$, then $f(z-\varepsilon) > 0 > f(z+\varepsilon)$.*

Proof. Since $f(z) = 0$, we can write f as

$$a_1(x-z) + a_2(x-z)^2 + \cdots + a_n(x-z)^n,$$

and $a_1 = f'(z)$, as can be seen by taking the formal derivative of f as written above, and substituting z for x, which kills all terms except the new constant term a_1. This is the Taylor series expansion of f around $x = z$.

We have $f = a_1(x-z)(1+g)$, where

$$g = b_1(x-z) + b_2(x-z)^2 + \cdots + b_{n-1}(x-z)^{n-1}$$

and $b_i = a_{i+1}/a_1$. It suffices to show that for sufficiently small ε, $|g(z \pm \varepsilon)| < 1$, since then the sign of f is the same as the sign of $a_1(x-z)$ at $x = z \pm \varepsilon$.

Agreeing to take $\varepsilon < 1$, we have

$$|g(z \pm \varepsilon)| \leq |b_1|\varepsilon + |b_2|\varepsilon^2 + \cdots + |b_{n-1}|\varepsilon^{n-1} \leq \varepsilon \sum_{i=1}^{n-1} |b_i|,$$

since $\varepsilon > \varepsilon^2 > \cdots > \varepsilon^{n-1}$. So any $\varepsilon < \min(1, (\sum_{i=1}^{n-1} |b_i|)^{-1})$ will do. □

Theorem 3.36 (Sturm's Theorem)**.** *Suppose $(R; \leq)$ is a real closed field, f is a polynomial without multiple roots, $f_0, \ldots, f_n$ is a Sturm sequence for f, and $a, b \in R$ with $a < b$ and such that neither a nor b is a root of any polynomial f_i in the sequence. Then the number of roots of f in the interval $(a, b) \subseteq R$ is $v(a) - v(b)$.*

Proof. List the roots of all the polynomials f_i between a and b in ascending order: $z_1 < z_2 < \cdots < z_m$. Let $c_0 = a$ and $c_m = b$, and for all $0 < j < m$, pick

$c_j \in R$ such that $z_j < c_j < z_{j+1}$ (for example, $c_j = \frac{z_j + z_{j+1}}{2}$ works). Then we have

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

so it suffices to show that

$$(v(c_i) - v(c_{i+1})) = \begin{cases} 1 & \text{if } z_{i+1} \text{ is a root of } f \\ 0 & \text{otherwise.} \end{cases}$$

That is, we have reduced to the case where $a < z < b$ and z is the only root of any of the polynomials f_i in the interval (a, b).

For all $0 \leq i < n$, if z is *not* a zero of f_i, then by IVTP, $f_i(a)$ and $f_i(b)$ have the same sign. So if z is not a zero of f_i or of f_{i+1}, then $f_i(a)$ and $f_{i+1}(a)$ have the same sign if and only if $f_i(b)$ and $f_{i+1}(b)$ have the same sign. So going from f_i to f_{i+1}, the sign changes once at a and at b, or zero times at a and at b.

Since f_n is a non-zero constant, z is not a zero of f_n. For all $1 \leq i < n$, if z is a zero of f_i, then z is not a zero of $f_{i-1}(z)$ and $f_{i+1}(z)$ are nonzero and have opposite signs. So $f_{i-1}(a)$ and $f_{i+1}(a)$ also have opposite signs, and regardless of the sign of $f_i(z)$, the sign changes exactly once at a as we go from f_{i-1} to f_i to f_{i+1}. The same is true at b. It follows from all of the above that if z is not a zero of f, then $v(a) - v(b) = 0$.

In the last case, suppose z is a root of $f_0 = f$. By Lemma 3.35, if $f'(z) > 0$, then there is some $\varepsilon > 0$ with $a < z - \varepsilon < z < z + \varepsilon < b$ such that $f(z - \varepsilon) < 0$ and $f(z + \varepsilon) > 0$. By IVTP, f does not change sign between a and $z - \varepsilon$ or between $z + \varepsilon$ and b. So $f(a) < 0$ and $f(b) > 0$. But since z is not a zero of f', $f'(a)$ and $f'(b)$ have the same sign as $f'(z) > 0$. So going from f_0 to f_1, the sign changes at a but not at b. On the other hand, if $f'(z) < 0$, similar reasoning shows that $f(a) > 0$ and $f(b) < 0$, while $f'(a) < 0$ and $f'(b) < 0$. Again, the sign changes at a but not at b. So $v(a) - v(b) = 1$. □

Corollary 3.37. *Suppose $(K; \leq)$ is an ordered field and $K \subseteq R_1$ and $K \subseteq R_2$ are two real closures of K such that the unique orderings on R_1 and R_2 each extend $\leq$. If $f \in K[x]$ is irreducible, then f has a root in R_1 if and only if it has a root in R_2.*

Proof. Let R be any real closure of K such that the unique ordering on R extends $\leq$. Since f is irreducible and K has characteristic 0, f has no repeated roots (characteristic 0 fields are perfect). Let $f = f_0, f_1, \ldots, f_n$ be a Sturm sequence for f, computed in $K[x]$ by Proposition 3.33. The proof of Proposition 3.33 applies just as well in $R[x]$. so $f_0, f_1, \ldots, f_n$ is a Sturm sequence for f over R. By Lemma 3.30, there is some $b \in K$, depending only on the coefficients of f, such that all the roots of f in R lie in the interval $(-b, b)$. By Theorem 3.36, the number of roots of f in R is $v(-b) - v(b)$. This computation is independent of R, since it only depends on the signs of polynomials in $K[x]$ evaluated on elements of K. So it is the same for any real closed field whose ordering is compatible with $\leq$. □

We are finally ready to prove uniqueness of ordered real closures.

Theorem 3.38. *Suppose $(K;\leq)$ is an ordered field and $K \subseteq R_1$ and $K \subseteq R_2$ are two real closures of K such that the unique orderings on R_1 and R_2 each extend $\leq$. Then there is an isomorphism $\sigma\colon R_1 \cong R_2$ such that $\sigma|_K = id_K$.*

Proof. Consider the set of all pairs (F,σ), where $K \subseteq F \subseteq R_1$ is a subfield and $\sigma\colon F \to R_2$ is an ordered field embedding, such that $\sigma|_K = \mathrm{id}_K$. We partially order this set in the natural way: $(F,\sigma) \leq (F',\sigma')$ if $F \subseteq F'$ and $\sigma'|_F = \sigma$. By Zorn's Lemma, this set has a maximal element. It remains to show that if (F,σ) is maximal, then $F = R_1$ and σ is surjective, and hence an isomorphism.

We first claim that for any finite extension $F \subseteq F' \subseteq R_1$, there is an embedding of fields (not necessarily order-preserving) $\sigma'\colon F' \to R_2$ such that $\sigma'|_F = \sigma$. By the primitive element theorem, we can write $F' = F(\alpha)$ for some $\alpha \in F'$. Let $f \in F[x]$ be the minimal polynomial of α. Let $F_\sigma \cong F$ (isomorphic as ordered fields) be the image of F under σ, and let f_σ be the image of f in $F_\sigma[x]$. Since α is a root of f in R_1, by Corollary 3.37, f_σ has a root β in R_2. Then σ extends to an embedding of fields $\sigma'\colon F' = F(\alpha) \to F_\sigma(\beta) \subseteq R_2$ by $\alpha \mapsto \beta$.

Now let $\alpha \in R_1$, and let $F' = F(\alpha)$. We would like to extend σ to an *ordered* field embedding $\sigma'\colon F' \to R_2$. We have just seen that there is some embedding of fields $F' \to R_2$. In fact, there are only finitely many, $\sigma_1,\ldots,\sigma_n$, one for each root of the minimal polynomial of α in R_2. Suppose for contradiction that none of the σ_i are order-preserving. Then for each $1 \leq i \leq n$, there is some $\gamma_i \in F'$ such that $\gamma_i > 0$ but $\sigma_i(\gamma_i) < 0$. Since $\gamma_i > 0$ in R_1, we can let $\delta_i = \sqrt{\gamma_i} \in R_1$. Consider the field $F(\alpha,\delta_1,\ldots,\delta_n)$. This is a finite extension of F, so by the claim there is an embedding of fields $\sigma^*\colon F(\alpha,\delta_1,\ldots,\delta_n) \to R_2$. Then $\sigma^*|_{F'}$ is an embedding $F' \to R_2$, so $\sigma^*|_{F'} = \sigma_j$ for some $1 \leq j \leq n$. But then $\sigma_j(\gamma_j) = \sigma^*(\gamma_j) = \sigma^*(\delta_j)^2$ is a square in R_2, contradicting $\sigma_j(\gamma_j) < 0$.

Thus there is an ordered field embedding $\sigma'\colon F' \to R_2$ extending σ, and by maximality of (F,σ), $F' = F$, so $\alpha \in F$. Since α was arbitrary, $F = R_1$. Now $R_1 \cong \sigma(R_1) \subseteq R_2$ is a formally real algebraic extension but R_1 is real closed, so $\sigma(R_1) = R_2$, and σ is an isomorphism. □

By Theorem 3.28 and Theorem 3.38, we have that for a formally real field K, the real closures of K up to isomorphism over K are in bijection with the orderings of K.

Exercise 17. Show that $\mathbb{Q}(t)$ has $2^{\aleph_0}$-many orderings, and hence $2^{\aleph_0}$-many real closures up to isomorphism.

3.4 Quantifier elimination and some consequences

We return now to model theory. Let RCF be the theory of real closed fields, in the language $\mathcal{L}_r$ of rings. This consists of:

- The field axioms.

- Sentences asserting that the field is formally real: For each $n \in \mathbb{N}$,

$$\forall x_1 \ldots \forall x_n \, (x_1^2 + \cdots + x_n^2 + 1 \neq 0).$$

- Sentences asserting that every odd degree polynomial has a root: For each $n \in \mathbb{N}$,

$$\forall a_0 \ldots \forall a_{2n} \, \exists y \, (y^{2n+1} + a_{2n} y^{2n} + \cdots + a_1 y + a_0 = 0).$$

- A sentence asserting that every element or its additive inverse is a square:

$$\forall x \, \exists y \, ((y^2 = x) \vee (y^2 = -x)).$$

If we add the relation symbol $\leq$ to the language, along with a sentence defining $\leq$ in any real closed field, we obtain the theory $\mathrm{RCF}_{\leq}$ of real closed ordered fields, in the language $\mathcal{L}_{or}$ of ordered rings:

$$\mathrm{RCF}_{\leq} = \mathrm{RCF} \cup \{\forall x \, \forall y \, (x \leq y \leftrightarrow \exists z \, (z^2 = y - x))\}.$$

This is called a **definitional expansion** of RCF: we add a new symbol to the language and an axiom which completely determines the interpretation of this symbol in any model.

An alternative approach to axiomatizing real closed ordered fields would be to add the axioms of ordered fields to RCF (the axioms asserting that $\leq$ is a linear order which is preserved by addition and multiplication by non-negative elements). The resulting theory would have exactly the same models, and hence the same logical consequences, as the theory $\mathrm{RCF}_{\leq}$ defined above, so the difference is not important to us.

Theorem 3.39. $\mathrm{RCF}_{\leq}$ *has quantifier elimination.*

Proof. We use the test in Corollary 2.25. So suppose R_1 and R_2 are algebraically closed fields, A is an $\mathcal{L}_{or}$, structure, and $g \colon A \to R_1$ and $h \colon A \to R_2$ are embeddings. Let φ be a primitive formula, $\exists y \bigwedge_{i=1}^{n} \varphi_i(\overline{x}, y)$, and let $\overline{a}$ be a tuple from A such that $R_2 \models \varphi(h(\overline{a}))$. We would like to show that $R_1 \models \varphi(g(\overline{a}))$.

Note that A is isomorphic to an ordered subring $g(A)$ of R_1, so it is an ordered integral domain. Let $A' = \mathrm{Frac}(A)$, the field of fractions of A. Let B_1 be the subfield of R_1 generated by $g(A)$, and let B_2 be the subfield of R_2 generated by $h(A)$. Then g extends to an isomorphism of fields $g' \colon A' \cong B_1$, and this is an isomorphism of ordered fields by Theorem 3.15, since there is a unique ordering on $\mathrm{Frac}(A)$ extending the ordering on A. Similarly, h extends to an isomorphism of ordered fields $h' \colon A' \cong B_2$.

Let C_1 be the relative algebraic closure of B_1 in R_1. Then C_1 is real closed: it is a subring of a formally real field, so it is formally real, any odd degree polynomial over C_1 has a root in R_1, and hence in C_1, since C_1 is relatively algebraically closed, and for every $a \in C_1$, either a or $-a$ is a square in R_1, and hence in C_1, since C_1 is relatively algebraically closed. Similarly, letting

C_2 be the relative algebraic closure of B_2 in R_2, C_2 is real closed. Further C_1 and C_2 are both ordered real closures of A', so by Theorem 3.38, there is an isomorphism $\sigma\colon R_1 \cong R_2$ such that $\sigma \circ g' = h'$.

Now let's analyze the primitive formula $\exists y \bigwedge_{i=1}^n \varphi_i(h(\overline{a}), y)$. Since each φ_i is atomic or negated atomic, we may assume by the reasoning in Remark 2.26 that $\varphi_i(h(\overline{a}), y)$ is a polynomial equality $p_i(y) = 0$, negated equality $p_i(y) \neq 0$, inequality $p_i(y) \geq 0$, or negated inequality $\neg(p_i(y) \geq 0)$, where $p_i \in C_2[y]$.

Let $b \in R_2$ be a witness to the existential quantifier. If b is algebraic over C_2, then $b \in C_2$, since C_2 is relatively algebraically closed in R_2. On the other hand, suppose that b is not algebraic over C_2. For each φ_i, let $p_i(y) \in C_2[y]$ be the polynomial in the formula φ_i. Then b is not a root of any of the p_i, and in particular, none of the φ_i are polynomial equalities.

List all of the roots of all of the p_i in C_2 in increasing order:

$$z_1 < z_2 < \cdots < z_m.$$

Since C_2 is relatively algebraically closed in R_2, the $p_i(y)$ do not have additional roots in R_2, and by IVTP, the $p_i(y)$ do not change sign in R_2 in the intervals $(-\infty, z_1)$, (z_i, z_{i+1}) for all $1 \leq i < m$, and (z_m, ∞). In particular, for any b' in the same interval as b, we also have $R_2 \models \bigwedge_{i=1}^n \varphi_i(h(\overline{a}), b')$.

In particular, we can find a witness $b' \in C_2$. If $b < z_1$, let $b' = z_1 - 1 \in C_2$. If $b > z_n$, let $b' = z_n + 1 \in C_2$. And if $z_i < b < z_{i+1}$ for some i, let $b' = \frac{z_i + z_{i+1}}{2}$.

In either case (if b is algebraic over C_2 or if not), we have a witness $b' \in C_2$ to the existential quantifier. Since embeddings preserve and reflect quantifier-free formulas, we have

$$C_2 \models \bigwedge_{i=1}^n \varphi_i(h(\overline{a}), b'),$$

so

$$C_1 \models \bigwedge_{i=1}^n \varphi_i(\sigma^{-1}(h(\overline{a})), \sigma^{-1}(b')),$$

hence

$$R_1 \models \bigwedge_{i=1}^n \varphi_i(g(\overline{a}), \sigma^{-1}(b')),$$

and

$$R_1 \models \varphi(g(\overline{a})).$$ □

Corollary 3.40. $\mathrm{RCF}_{\leq}$ *and* RCF *are complete theories.*

Proof. Let $R \models \mathrm{RCF}_{\leq}$. Since every real closed field has characteristic 0, there is an embedding of fields $\mathbb{Q} \to R$. Since $\mathbb{Q}$ admits a unique ordering, this is an embedding of ordered fields. By Proposition 2.28, $\mathrm{RCF}_{\leq}$ is complete.

For RCF, it suffices to show that if φ is a $\mathcal{L}_r$-sentence and $\mathrm{RCF}_{\leq} \models \varphi$, then also $\mathrm{RCF} \models \varphi$. Then since $\mathrm{RCF}_{\leq}$ entails ψ or $\neg\psi$ for every $\mathcal{L}_{or}$-sentence ψ, also RCF entails ψ or $\neg\psi$ for every $\mathcal{L}_r$-sentence ψ.

So suppose $\mathrm{RCF}_{\leq} \models \varphi$. Let $M \models \mathrm{RCF}$. Then M has a (unique) expansion to a model $M_{\leq} \models \mathrm{RCF}_{\leq}$, and $M_{\leq} \models \varphi$. But φ is a $\mathcal{L}_r$-sentence, so its truth does not depend on the interpretation of $\leq$, and hence also $M \models \varphi$. Thus $\mathrm{RCF} \models \varphi$. □

It follows that a field is real closed if and only if it is elementarily equivalent to the real field $\mathbb{R}$. Let's take a moment to summarize the various characterizations of real closed fields that we have found.

Theorem 3.41. *Let R be a field. The following are equivalent:*

(1) R is formally real and has no proper formally real algebraic extensions.

(2) R is formally real, every odd degree polynomial in $R[x]$ has a root in R, and for all $a \in R$, either a or $-a$ is a square in R.

(3) R is not algebraically closed, but $R(\sqrt{-1})$ is algebraically closed.

(4) There exists an ordering $\leq$ on R making R an ordered field which satisfies the intermediate value theorem for polynomials.

(5) R is elementarily equivalent to $\mathbb{R}$.

As discussed at the end of Section 2.3, completeness has the following consequence.

Corollary 3.42. $\mathrm{RCF}_{\leq}$ *and* RCF *are decidable theories.*

It follows that basic Euclidean geometry is also decidable, by coordinatizing the plane as $\mathbb{R}^2$ and space as $\mathbb{R}^3$ and translating geometric propositions into formulas in the language of ordered rings.

A quantifier-free definable subset of R^n, where R is a real closed field, is called a **semialgebraic** set. More concretely, a semialgebraic set is a finite Boolean combination of sets defined by polynomial equations $p(\overline{x}) = 0$ and inequalities $p(\overline{x}) \geq 0$. In geometric terms, quantifier elimination for $\mathrm{RCF}_{\leq}$ gives us the following:

Corollary 3.43 (Tarski–Seidenberg)**.** *In affine space over a real closed field, a coordinate projection of a semialgebraic set is semialgebraic.*

Another important consequences of quantifier elimination is a condition called "model-completeness". While RCF does not have quantifier elimination, it is a model-complete theory.

Definition 3.44. A theory T is **model-complete** if every embedding between models of T is an elementary embedding.

Proposition 3.45. *Suppose T has quantifier elimination. Then T is model-complete.*

Proof. Let M and N be models of T, and let $h\colon M \to N$ be an embedding. Let $\varphi(\overline{x})$ be a formula and $\overline{a}$ a tuple from M. Then φ is T-equivalent to a quantifier-free formula ψ, and h preserves and reflects ψ, so

$$\begin{aligned} M \models \varphi(\overline{a}) &\iff M \models \psi(\overline{a}) \\ &\iff N \models \psi(h(\overline{a})) \\ &\iff N \models \varphi(h(\overline{a})). \end{aligned}$$

So h is an elementary embedding. □

Theorem 3.46. RCF *and* $\mathrm{RCF}_{\leq}$ *are model-complete.*

Proof. Since $\mathrm{RCF}_{\leq}$ has quantifier elimination, it is model-complete by Proposition 3.45.

For RCF, let $h\colon M \to N$ be an embedding of real closed fields. Let $M_{\leq}$ and $N_{\leq}$ be the unique expansions of M and N to models of $\mathrm{RCF}_{\leq}$. Then h is an embedding of ordered fields: If $a \leq b$ in M, then $(b-a)$ is a square in M, so it is a square in N, and hence $a \leq b$ in N. And conversely, if $a \not\leq b$ in M, then $b < a$, so $(a-b)$ is a square in M, and $b < a$ in N, so $a \not\leq b$ in N.

Since $\mathrm{RCF}_{\leq}$ is model-complete, h preserves and reflects all $\mathcal{L}_{or}$-formulas, and in particular all $\mathcal{L}_r$-formulas. So h is an elementary embedding. □

As an application of model-completeness, we can give an easy solution to Hilbert's 17th problem.

Definition 3.47. Let F be an ordered field, and let $f \in F[t_1, \ldots, t_n]$ be polynomial in multiple variables. We write $\widehat{f}$ for the polynomial function $F^n \to F$ defined by f. We say f is **positive semidefinite** if $\widehat{f}(a_1, \ldots, a_n) \geq 0$ for all $a_1, \ldots, a_n \in F$.

Hilbert asked whether every positive semidefinite polynomial over $\mathbb{R}$ can be expressed as a sum of squares of *rational functions* over $\mathbb{R}$. Some motivation: Hilbert had previously proved that there are positive semidefinite polynomials which *cannot* be expressed as sums of squares of *polynomials* (though the first explicit example was not given until 1967!). On the other hand, he proved that every positive semidefinite polynomial *in at most two variables* can be expressed as a sum of squares of *rational functions*. The problem was to remove the bound on the number of variables.

Emil Artin answered the question affirmatively in 1927, over arbitrary real closed fields. The observation that model theory gives a very simple proof is due to Abraham Robinson.

Theorem 3.48. *Let R be a real closed field. Every positive semidefinite polynomial $f \in R[t_1, \ldots, t_n]$ is a sum of squares of rational functions in $R(t_1, \ldots, t_n)$.*

Proof. Suppose $f \in R[t_1, \ldots, t_n]$ is not a sum of squares in $R(t_1, \ldots, t_n)$. The field $R(t_1, \ldots, t_n)$ is formally real (because the field of rational functions over

a formally real field is always formally real, or by Exercise 14), and by Theorem 3.14, there is an ordering $\leq$ on $R(t_1, \ldots, t_n)$ such that $f < 0$. Let R^* be an ordered real closure of $R(t_1, \ldots, t_n)$.

Now in R^*, $\widehat{f}(t_1, \ldots, t_n) = f < 0$, so $R^* \models \exists x_1 \ldots \exists x_n \, (f(x_1, \ldots, x_n) < 0)$, and by model-completeness of $\mathrm{RCF}_{\leq}$, also $R \models \exists x_1 \ldots \exists x_n \, (f(x_1, \ldots, x_n) < 0)$. This means there exist $a_1, \ldots, a_n \in R$ such that $\widehat{f}(a_1, \ldots, a_n) < 0$, so f is not positive semidefinite. □

4 o-minimality

4.1 Definition and examples

Before introducing o-minimality, I want to briefly discuss its predecessor, strong minimality. Below, whenever I say "definable", I mean definable with parameters.

Definition 4.1. Let T be a complete theory. We say T is **strongly minimal** if for every model $M \models T$, every definable subset of M^1 is finite or cofinite.

In other words, for any formula $\varphi(x, \overline{y})$, where x is a single variable, and any tuple $\overline{b} \in M^n$, the set $\varphi(M, \overline{b}) = \{a \in M \mid M \models \varphi(a, \overline{b})\}$ is finite or cofinite.

Definition 4.2. We say a structure M is **strongly minimal** if its complete theory $\mathrm{Th}(M)$ is strongly minimal.

We could also reasonably consider structures M such that every definable subset of M^1 is finite or cofinite, without requiring this to be true for every model of $\mathrm{Th}(M)$. An example is the structure $(\mathbb{N}; \leq)$. Such structures are called **minimal**. But they admit a much weaker structure theory than the strongly minimal structures.

Proposition 4.3. *Every algebraically closed field is strongly minimal.*

Proof. The complete theory of an algebraically closed field is ACF_p, where p is prime or 0. Let $M \models \mathrm{ACF}_p$. By quantifier elimination, every formula in one variable is equivalent in M to a Boolean combination of atomic formulas. An atomic formula is a polynomial equation, which defines a finite subset of M, or all of M (in the case of $0 = 0$). Now it is easy to check that the set of finite and cofinite subsets of M is closed under Boolean combinations (intersections, unions, and complements). □

Note that the definition of strong minimality places no restriction on definable subsets of M^n for $n > 1$. For example, in models of ACF_p, we get all affine algebraic varieties as definable sets. Nevertheless, strong minimality has powerful consequences. For example, in any strongly minimal theory T, canonical dimensions can be assigned to all definable sets and to all models. In the case of ACF_p, these notions specialize to algebraic dimension (of algebraic varieties)

and transcendence degree (over the prime field). A model of a strongly minimal theory T is uniquely determined up to isomorphism by its dimension, and it follows that T has a unique model up to isomorphism in each uncountable cardinality.

In the other direction, Baldwin and Lachlan (following Morley) proved that if T is a theory with a unique model up to isomorphism in some uncountable cardinality, then every model of T contains a strongly minimal definable set, and the dimension of this set again determines the model up to isomorphism. These ideas led directly to Shelah's stability theory and classification theory, which gave rise to much of modern model theory. And analogies between definable sets in arbitrary theories and algebraic geometry, via ACF_p, continue to be fruitful.

A linearly ordered structure has no hope of being strongly minimal. Nevertheless, theories such as $\mathrm{RCF}_<$ admit a "tame" geometry of definable sets similar to that exhibited by strongly minimal theories. The following perspective on strong minimality suggests an appropriate variant for ordered structures: the finite and cofinite sets are exactly those subsets of M^1 which are definable with parameters in the empty language (i.e., using just $=$ and the logical connectives). If M is an ordered structure, we can consider instead those subsets of M^1 which are definable with parameters in the language $\{\leq\}$. When $\leq$ is a dense linear order without endpoints, these turn out to be exactly the finite unions of points and intervals.

Definition 4.4. A linear order $\leq$ is **dense** if for all $a < b$ there exists c such that $a < c < b$. We say $\leq$ **has no endpoints** if for all a there exists b such that $a < b$ and there exists c such that $c < a$.

Definition 4.5. Let T be a complete theory in a language which includes a binary relation symbol $\leq$, such that $\leq$ is a dense linear order without endpoints in every model of T. We say T is **o-minimal** if for every model $M \models T$, every definable subset of M^1 (with parameters) is a finite union of points and intervals.

It should be clarified that by an **interval**, I mean an open interval, i.e., a set of the form

$$\begin{aligned}(a,b) &= \{c \in M \mid a < c < b\},\\ (-\infty,b) &= \{c \in M \mid c < b\},\\ (a,\infty) &= \{c \in M \mid a < c\}, \text{ or}\\ (-\infty,\infty) &= M\end{aligned}$$

where $a, b \in M$. The restriction that the (non-infinite) endpoints are in M is crucial. For example, we do not consider $(0,\pi)$ to be an interval in $(\mathbb{Q},\leq)$. If we instead defined an interval to be a convex set with no greatest or least element, we would have defined **weakly o-minimal theories**. There is much less to say about weakly o-minimal theories, so we will not consider them further.

We do not lose anything by restricting to open intervals, since a closed interval is a finite union of points and open intervals: $[a, b] = \{a\} \cup (a, b) \cup \{b\}$. Also, every set can be written as a finite *disjoint* union of points and intervals. Indeed, if a point fails to be disjoint from another point or interval, it is contained in that point or interval and can be removed. And if two intervals (a, b) and (c, d) fail to be disjoint, then $(a, b) \cup (c, d) = (\min(a, c), \max(b, d))$.

In the definition of o-minimality above, I required that $\leq$ be a dense linear order without endpoints. The standard definition only requires that $\leq$ be a linear order. However, all of the important examples of o-minimal theories are densely ordered, and many of the key results about o-minimal theories require density as an additional hypothesis. For our purposes it will be simpler to assume it everywhere.

Definition 4.6. We say a structure $(M; \leq, \dots)$ is **o-minimal** if its complete theory $\mathrm{Th}(M)$ is o-minimal.

Remark 4.7. Similarly to the distinction between minimal and strongly minimal structures, we could have defined "M is o-minimal" to mean that every definable subset of M^1 is a finite union of points and intervals, and "M is strongly o-minimal" to mean the same is true for every model of $\mathrm{Th}(M)$. But Knight, Pillay, and Steinhorn proved that every o-minimal structure in this sense is strongly o-minimal, so we will not use the term "strongly o-minimal".

Proposition 4.8. *Every ordered real closed field is o-minimal.*

Proof. The complete theory of an ordered real closed field is $\mathrm{RCF}_{<}$. Let $M \models \mathrm{RCF}_{<}$. By quantifier elimination, every formula in one variable is equivalent in M to a Boolean combination of atomic formulas. An atomic formula is a polynomial equation $p(x) = 0$ or inequality $p(x) \geq 0$. The former defines a finite subset of M (a finite union of points) or all of M (the interval $(-\infty, \infty)$). The latter defines a finite union of points (the zeros of p in M) and intervals (the intervals between the zeros of p on which p is positive). Note that p does not change sign between its zeros, by IVTP.

It remains to show that the finite unions of points and intervals are closed under union, intersection, and complement. So let $\mathcal{U} = \bigcup_{i=1}^{m} U_i$ and $\mathcal{V} = \bigcup_{j=1}^{n} V_j$, where each U_i and V_j is a point or an interval.

- $\mathcal{U} \cup \mathcal{V} = (\bigcup_{i=1}^{m} U_i) \cup (\bigcup_{j=1}^{n} V_j)$ is a finite union of points and intervals.
- $\mathcal{U} \cap \mathcal{V} = \bigcup_{i=1}^{m} \bigcup_{j=1}^{n} (U_i \cap V_j)$. An intersection of two points or of a point and an interval is either empty or a point. And an intersection of two intervals is empty or an interval: $(a, b) \cap (c, d) = (\max(a, c), \min(b, d))$ if $\max(a, c) < \min(b, d)$ and empty otherwise. Here a and c may be $-\infty$ and b and d may be ∞.
- $\mathcal{U}^c = \bigcap_{i=1}^{m} U_i^c$. By closure under intersections, it suffices to show that U_i^c is a finite union of points and intervals. If $U_i = \{a\}$ is a point, then $U_i^c = (-\infty, a) \cup (a, \infty)$. If $U_i = (a, b)$, then $U_i^c = (-\infty, a) \cup \{a\} \cup \{b\} \cup (b, \infty)$.

Similar decompositions work for $(-\infty, b)$ and (a, ∞). And if $U_i = M$, then $U_i^c = \emptyset$, the empty union. $\square$

Example 4.9. Here are some examples (and non-examples) of o-minimal structures.

- $(\mathbb{R}; \leq)$ and $(\mathbb{Q}; \leq)$. These structures are elementarily equivalent, with complete theory DLO, the theory of dense linear orders without endpoints. DLO has quantifier elimination, and o-minimality follows easily.

- $(\mathbb{R}; \leq, +, 0, -)$ and $(\mathbb{Q}; \leq, +, 0, -)$. These structures are elementarily equivalent, with complete theory ODAG, the theory of ordered divisible abelian groups. ODAG has quantifier elimination, and o-minimality follows easily.

- $(\mathbb{R}; \leq, +, \times, 0, 1, -)$, the real ordered field. We showed above that this structure is o-minimal. The other structures we will consider are expansions of this one, so we introduce the following terminology: $\overline{R}$ is the real ordered field, so $(\overline{R}; f)$ is a structure in the language $\mathcal{L}_{or} \cup \{f\}$.

- The rational ordered field $(\mathbb{Q}; \leq, +, \times, 0, 1, -)$ is *not* o-minimal. For example, the formula $x^2 \leq 2$ defines $(-\sqrt{2}, \sqrt{2})$, which is not a finite union of intervals in $\mathbb{Q}$, since $\sqrt{2} \notin \mathbb{Q}$. Also, $\exists y\, (y^2 = x)$ defines an infinite set which does not contain any interval, since the non-squares are dense in $\mathbb{Q}$.

- $\mathbb{R}_{\exp} = (\overline{\mathbb{R}}; e^x)$ is o-minimal. This is a hard theorem, due to Alex Wilkie in 1991. The theory of this structure does not admit quantifier elimination, but it is model-complete. No explicit axiomatization is known – axiomatizing this theory would lead to deciding the real Schanuel conjecture, which would answer many open problems in transcendence theory.

- $(\overline{\mathbb{R}}; \sin)$ is *not* o-minimal. The set defined by $\sin(x) = 0$ is $\{n\pi \mid n \in \mathbb{Z}\}$, which is not a finite union of points and intervals.

- On the other hand, $(\overline{\mathbb{R}}; \sin|_{[0,2\pi]})$ is o-minimal. Here $\sin|_{[0,2\pi]}$ is a unary function defined by $x \mapsto \sin(x)$ if $x \in [0, 2\pi]$ and $x \mapsto 0$ if $x \notin [0, 2\pi]$.

- More generally, let f be a real-valued analytic function in n variables, defined an an open set containing the unit cube $[0, 1]^n$. Then $(\mathbb{R}, \widehat{f})$ is o-minimal, where $\widehat{f}$ is a function defined by $\widehat{f}(\overline{a}) = f(\overline{a})$ for $\overline{a} \in [0, 1]^n$ and $\widehat{f}(\overline{a}) = 0$ otherwise. This is due to van den Dries in 1986, building on work of Gabrielov in the 1960s, who showed that the complete theory of $(\mathbb{R}, \widehat{f})$ is model-complete.

- $\overline{\mathbb{R}}_{\text{an,exp}}$ is the structure defined by adding e^x and *all* restricted analytic functions $\widehat{f}$ as defined above. This structure is again o-minimal, as shown by van den Dries and Miller in 1992.

4.2 The order topology and definable functions

Definition 4.10. Let $(M; \leq, \dots)$ be an ordered structure and $A \subseteq M$. We say that b is an **upper bound** for A if $A \leq b$ for all $a \in A$. We say that M is **complete** if every non-empty subset of M which has an upper bound has a least upper bound.

Order-completeness is responsible for most of the nice topological and analytic properties of $\mathbb{R}$. But it is well-known that every complete ordered field is isomorphic to $\mathbb{R}$. As a result, all other real closed fields, and hence many o-minimal structures, fail to be complete.

For an explicit example, the real closure of the rationals $\mathbb{Q}^r$, is not complete, since $(0, \pi)$ has an upper bound but no least upper bound in $\mathbb{Q}^r$. As another typical example, if R is a real closed field with "infinite" elements, i.e. such that there exists $t \in R$ with $n \leq t$ for all $n \in \mathbb{N}$, then $\mathbb{N}$ has no least upper bound in R: If t is an upper bound for $\mathbb{N}$, then so is $t - 1$.

However, if we restrict our attention to definable sets in an o-minimal context, we regain completeness. We say that o-minimal structures are **definably complete**.

Proposition 4.11. *Suppose $(M; \leq, \dots)$ is o-minimal and $X \subseteq M$ is a definable set. If X has an upper bound in M, then X has a least upper bound $\sup(X) \in M$.*

Proof. Since X is definable, it can be decomposed into a finite disjoint union of points and intervals. Let U be the greatest point or interval in this decomposition. If $U = (a, b)$ is an interval, then since X has an upper bound in M, $b \neq \infty$. Then $\sup(X) = b$. If $U = \{c\}$ is a point, then $\sup(X) = c$. □

A similar argument shows that if a definable set in an o-minimal structure M has a lower bound in M, then it has a greatest lower bound $\inf(X) \in M$. If X is not bounded above or below, we often write $\sup(X) = \infty$ or $\inf(X) = -\infty$, respectively.

In any ordered structure $(M; \leq, \dots)$, we consider the **order topology**, in which the basic open sets are the intervals. We also consider the product topology on M^n for all n, in which the basic open sets are the boxes: products of intervals

$$(a_1, b_1) \times \cdots \times (a_n, b_n).$$

Note that every basic open set is definable. This allows us to work with the topology definably. For example, we have the following.

Proposition 4.12. *Let $(M; \leq, \dots)$ be an ordered structure, and let $A \subseteq M^n$ be a definable set for some n. Then the topological closure, interior, and boundary of A: $\mathrm{cl}(A)$, $\mathrm{int}(A)$, and $\mathrm{bd}(A)$, are definable sets.*

Proof. Let $\varphi_A(\overline{x}, \overline{b})$ define A. A point $\overline{a} \in M^n$ is in the interior of A if it has a basic open neighborhood which is contained in A. This property can be defined as follows:

$$\exists \overline{y}\, \exists \overline{z} \left(\left(\bigwedge_{i=1}^{n} y_i < x_i < z_i \right) \wedge \forall \overline{w} \left(\left(\bigwedge_{i=1}^{n} y_i < w_i < z_i \right) \rightarrow \varphi_A(\overline{w}, \overline{b}) \right) \right).$$

The closure of A can be defined similarly, or as the complement of the interior of the complement of A. The boundary of A is the closure minus the interior, so this is also definable. □

The order topology is Hausdorff in any ordered structure, but otherwise it may be rather badly behaved. For example, the order topology on $\mathbb{Q}^r$ agrees with the topology induced on $\mathbb{Q}^r$ as a subspace of $\mathbb{R}$, and $\mathbb{Q}^r$ is totally disconnected. There are also real closed fields in which the only compact sets are finite.

However, if we again restrict our attention to definable sets in o-minimal structures, order is restored, using definable completeness.

Definition 4.13. Let $(M; \leq, \dots)$ be an ordered structure. A set $A \subseteq M^n$ is **definably connected** if it is definable and it cannot be written as a union of two disjoint non-empty open definable subsets of A (here open means open in the subspace topology on A).

Proposition 4.14. *Let $(M; \leq, \dots)$ be o-minimal. Then every interval (a, b) and every generalized interval $[a, b]$, $[a, b)$, or $(a, b]$ is definably connected.*

Proof. Let I be an interval or a generalized interval. Suppose $I = U_1 \cup U_2$, where U_1 and U_2 are disjoint, non-empty, definable, and open in A. Pick $a \in U_1$ and $b \in U_2$. Without loss of generality, $a < b$.

Since U_1 is open, it can be written as a disjoint union of open intervals, one of contains a. Let s be the right endpoint of this interval, and note that $a < s \leq b$, so $s \in I$. Then $s \notin U_1$, so $s \in U_2$. Since U_2 is open in I, there is a basic open neighborhood $s \in (s', s'')$ such that $(s', s'') \cap I \subseteq U_2$. By denseness of the order, we can pick some t with $s' < t < s$. But then $t \in (s', s'') \subseteq U_2$ and $t \in (a, s) \subseteq U_1$, contradicting disjointness of U_1 and U_2. □

Note that denseness of the order was crucial to the proof (we used it in Claim 2). If $a < b$ in M and there is no c with $a < c < b$, then $M = (-\infty, b) \cup (a, \infty)$, and M fails to be definably connected.

Much of our work on o-minimal structures will be concerned with controlling the behavior of definable functions.

Definition 4.15. Let $X \subseteq M^n$ and $Y \subseteq M^n$ be definable sets in an arbitrary structure M. A function $f \colon X \to Y$ is a **definable function** if its graph $\Gamma(f)$ is definable:

$$\Gamma(f) = \{(\overline{a}, \overline{b}) \mid f(\overline{a}) = \overline{b}\} \subseteq X \times Y \subseteq M^{n+m}.$$

Typically, we will be interested in single-valued definable functions, i.e., those with codomain M. Examples in real closed fields include all polynomials p (defined by $p(x) = y$), rational functions p/q (defined by $(q(x) \neq 0) \wedge (p(x) = yq(x))$), and root functions (for example, the square root function is defined by $(y \geq 0) \wedge (y^2 = x)$).

Proposition 4.16. *The images and preimages of definable sets under definable functions are definable.*

Proof. Suppose $\varphi_f(\overline{x}, \overline{y})$ defines a function $f\colon X \to Y$. Let $A \subseteq X$ be defined by $\psi_A(\overline{x})$, and let $B \subseteq Y$ be defined by $\psi_B(\overline{y})$. Then the image $f(A)$ is defined by

$$\exists \overline{x}\, (\psi_A(\overline{x}) \wedge \varphi_f(\overline{x}, \overline{y}))$$

and the preimage $f^{-1}(B)$ is defined by

$$\exists \overline{y}\, (\psi_B(\overline{y}) \wedge \varphi_f(\overline{x}, \overline{y})).$$ □

In an ordered structure M, we say a definable function $f\colon X \to Y$ is **continuous** if it is continuous with respect to the order topology (more properly, the subspace topologies on X and Y induced by the order topology), i.e., the preimage of any open set in Y is open in X. Note that it suffices to check continuity on basic open sets, which are definable. So if M is o-minimal, a function $f\colon M \to M$ is continuous if and only if the pre-image of any interval is a finite union of intervals.

Proposition 4.17. *Suppose M is o-minimal, $X \subseteq M^n$ is definably connected, and $f\colon X \to M^m$ is a continuous definable function. Then $f(X)$ is definably connected.*

Proof. By Proposition 4.16, $f(X)$ is definable. Suppose $f(X) = U_1 \cup U_2$, where U_1 and U_2 are disjoint non-empty open definable subsets of $f(X)$. Let $V_1 = f^{-1}(U_1)$ and $V_2 = f^{-1}(U_2)$. Then $X = V_1 \cup V_2$, and V_1 and V_2 are disjoint non-empty open definable subsets of X, contradicting the fact that X is definably connected. □

Corollary 4.18 (Definable intermediate value theorem)**.** *Suppose M is o-minimal, and $f\colon [a, b] \to M$ is a continuous definable function. Then for all c between $f(a)$ and $f(b)$, c is in the image of f.*

Proof. By Proposition 4.14, $[a, b]$ is definably connected, and by Proposition 4.17, $X = f([a, b])$ is definably connected. Let $c \in M$ such that $f(a) < c < f(b)$ (the case that $f(b) < c < f(a)$ is similar). Suppose for contradiction that $c \notin X$.

Let $U_1 = (-\infty, c) \cap X$ and $U_2 = (c, \infty) \cap X$. Then U_1 and U_2 are disjoint definable open subsets of X. They are non-empty, since $f(a) \in U_1$ and $f(b) \in U_2$, and $U_1 \cup U_2 = X$, since $c \notin X$. This contradicts definable connectedness. □

The definable intermediate value theorem gives us yet one more characterization of real closed fields.

Exercise 18. Let R be an ordered field. Show that constant functions are definable and continuous, $f(x) = x$ is definable and continuous, and sums and products of definable continuous functions are definable and continuous. Conclude that every polynomial function $R \to R$ is a definable continuous function.

Corollary 4.19. *Let R be an ordered field. Then R is o-minimal if and only if it is real closed.*

Proof. We have seen (Proposition 4.8) that every real closed field is o-minimal. Conversely, suppose R is o-minimal. By Exercise 18, every polynomial function $R \to R$ is continuous and definable. Then Corollary 4.18 implies that R satisfies the intermediate value theorem for polynomials, and by Theorem 3.31, R is real closed. □

Thus far, we have only recovered some basic theorems of real analysis in the o-minimal setting, essentially by inserting the word "definable" everywhere. In the case that the o-minimal structure we are considering expands $(\mathbb{R}; \leq)$, we have not gained anything new.

Our next goal is to prove something which is definitely *not* a classical result of real analysis: every single-variable definable function in an o-minimal structure is piecewise continuous and constant or strictly monotone.

Definition 4.20. Let $f\colon (a,b) \to M$ be a definable function. We say that f is **constant** on (a,b) if $f(c) = f(d)$ for all $c < d$ in (a,b). We say that f is **strictly increasing** on (a,b) if $f(c) < f(d)$ for all $c < d$ in (a,b). We say that f is **strictly decreasing** on (a,b) if $f(c) > f(d)$ for all $c < d$ in (a,b). We say that f is **strictly monotone** on (a,b) if it is strictly increasing or strictly decreasing.

Theorem 4.21 (Monotonicity Theorem)**.** *Let M be an o-minimal structure, and let $f\colon (a,b) \to M$ be a definable function. Then there are finitely many points $a_0 < a_1 < \cdots < a_k$ with $a_0 = a$ and $a_k = b$ such that on each interval (a_i, a_{i+1}), f is continuous and either constant or strictly monotone.*

Toward this theorem, we prove three lemmas. In all three, M is an o-minimal structure, $I \subseteq M$ is an interval, and $f\colon I \to M$ is a definable function.

We use repeatedly the following consequence of o-minimality: A definable set is infinite if and only if it contains an interval. Indeed, if a definable set fails to contain an interval, then it is a finite union of points, and hence finite. For the converse, every interval is infinite, by denseness of the order.

Lemma 4.22. *There is a subinterval of I on which f is constant or injective.*

Proof. Case 1: There is some $b \in f(I)$ such that $f^{-1}(\{b\})$ is infinite. By o-minimality, the definable set $f^{-1}(\{b\})$ contains an interval $I' \subseteq I$, and f is constant on I' with value b.

Case 2: For all $b \in f(I)$, the preimage of b is finite. Then since I is infinite, $f(I)$ is an infinite definable set, so it contains an interval J. Define $g\colon J \to I$ by $g(b) = \min\{a \in I \mid f(a) = b\}$, and note that g is an injective definable function. Since J is infinite, $g(J)$ is an infinite definable set, so it contains an interval $I' \subseteq I$. And f is injective on I', since $g \circ f$ is the identity. □

Lemma 4.23. *If f is injective, then there is a subinterval of I on which f is strictly monotone.*

Proof. Write $I = (a, b)$. For each $x \in I$, we consider the local behavior near x. The interval (a, x) is decomposed into two disjoint definable sets:

$$L_+(x) = \{y \in (a, x) \mid f(y) > f(x)\} \text{ and } L_-(x) = \{y \in (a, x) \mid f(y) < f(x)\}.$$

One of these sets has supremum x, so exactly one of them contains an interval (c, x) for some $c \in (a, x)$. We say x has left-type $+$ if $(c, x) \subseteq L_+(x)$ and x has left-type $-$ if $(c, x) \subseteq L_-(x)$.

Similarly, the interval (x, b) is decomposed into two definable sets:

$$R_+(x) = \{y \in (x, b) \mid f(y) > f(x)\} \text{ and } R_-(x) = \{y \in (x, b) \mid f(y) < f(x)\}.$$

Exactly one of these sets contains an interval (x, c) for some $c \in (x, b)$, and we say x has right-type $+$ if $(x, c) \subseteq R_+(x)$ and x has right-type $-$ if $(x, c) \subseteq R_-(x)$.

Altogether, each point in (a, b) has a type $++$, $+-$, $-+$, or $--$, where the first symbol indicates the left-type and the second indicates the right-type. Further, the set of points of each type is definable. For example, the points of type $+-$ are definable by

$$\exists c \exists d\, ((c < x < d) \wedge \forall y\, ((c < y < x) \to f(y) > f(x)) \wedge \forall z\, ((x < z < d) \to (f(z) < f(x)))).$$

Thus the interval (a, b) is a union of four definable sets, so at least one of these is infinite and hence contains an interval. Replacing I with this subinterval, we may assume that all points of I have the same type.

Case 1: Every point in $I = (a, b)$ has type $-+$. We will show that f is strictly increasing on I. The case $+-$ is similar, and f is strictly decreasing.

Fix $x \in I$. Since x has right-type $+$, there is an interval $(x, c) \subseteq R_+(x)$. Let s be the maximal right endpoint of that interval (this means that when we decompose $R_+(x)$ into a disjoint union of points and intervals, (x, s) is one of the intervals in the decomposition). Suppose for contradiction that $s < b$.

First, we show $f(x) < f(s)$. Since s has left-type $-$, we can pick $z < s$ sufficiently close to s such that $f(z) < f(s)$. Then $f(x) < f(z) < f(s)$.

Now since s has right-type $+$, we can pick $s < s'$ sufficiently close to s so that $f(s) < f(y)$ for all $y \in (s, s')$. For all $z \in (x, s')$, either $z \leq s$, in which case $f(x) < f(z)$, or $s < z < s'$, in which case $f(x) < f(s) < f(z)$. So $(x, s') \subseteq R_+(x)$, contradicting maximality of s.

Thus $s = b$, and the entire interval $(x, b) \subseteq R_+(x)$, so for all $y \in I$ with $x < y$, $f(x) < f(y)$. But x was arbitrary, so f is strictly increasing on I.

Case 2: Every point in I has type $++$. We will show that this is impossible. The case $--$ can also be shown to be impossible, by a similar argument.

We say a point $x \in I$ is left-heavy if there exists an open interval $x \in (y_1, y_2)$ such that for all $z_1 \in (y_1, x)$ and $z_2 \in (x, y_2)$, $f(z_1) > f(z_2)$. Similarly, we say $x \in I$ is right-heavy if the same holds but $f(z_1) < f(z_2)$. No point can be both left-heavy and right-heavy. Note that the sets of left-heavy and right-heavy points are definable.

We claim that if I is an interval on which every point has type $++$, then I contains a subinterval I' on which every point is left-heavy. A symmetric argument shows that I' contains a subinterval I'' on which every point is right-heavy. But then every point of I'' is both left-heavy and right-heavy, which is a contradiction. Let us prove the claim.

Let $A = \{x \in I \mid \text{for all } y \in (x, b), f(x) < f(y)\}$. A is definable, so if it is infinite, it contains an interval J. Pick some $y \in J$. Since y has left-type $+$, we can find some $x < y$ sufficiently close so that $x \in J$ and $f(x) > f(y)$. But this contradicts $x \in J$.

Thus A is finite, and we can restrict our attention to the subinterval $J = (\max(A), b)$. We have $(\star)$: For all $x \in J$, there exists $y \in (x, b)$ such that $f(x) > f(y)$.

Now fix $c \in J$. Consider the sets $R_+(c)$ and $R_(c)$. Exactly one of them contains an interval of the form (y, b). Let s the minimal left endpoint of that interval.

Suppose for contradiction that $(s, b) \subseteq R_+(c)$. By $(\star)$ for s, there exists $t \in (s, b)$ such that $f(s) > f(t) > f(c)$. So $s \in R_+(c)$. But since s has left-type $+$, there exists $s' < s$ such that for all $y \in (s', s)$, $f(y) > f(s) > f(c)$. Then $(s', b) \subseteq R_+(c)$, contradicting minimality of s.

Thus $(s, b) \subseteq R_-(c)$. Since s has right-type $+$, there exists y sufficiently close to s so that $s < y < b$ and $f(s) < f(y) < f(c)$. In particular, $c < s$. Now $[s, b) \subseteq R_-(c)$. Exactly one of $R_+(c)$ and $R_-(c)$ contains an interval of the form (s', s), and it cannot be $R_-(c)$, by minimality of s. So there exists $s' < s$ with $(s', s) \subseteq R_+(c)$ and $(s, b) \subseteq R_-(c)$. Thus s is left-heavy.

We have shown that for any $c \in J$, we can find $c < c' \in J$ such that c' is left-heavy. Repeating, we see that there are infinitely many left-heavy points in J. Since the set of left-heavy points is definable, it contains an interval, and we have proved the claim and completed the proof of the lemma. □

Lemma 4.24. *If f is strictly monotone, then there is a subinterval of I on which f is continuous.*

Proof. Since f is strictly monotone, it is injective. Then $f(I)$ is an infinite definable set, so it contains an interval J. Let $a, b \in J$ with $a < b$, and let a' and b' be their preimages, so $f(a') = a$ and $f(b') = b$. If f is strictly increasing, $a' < b'$, and f is an order-preserving bijection between the intervals $I' = (a', b') \subseteq I$ and $(a, b) \subseteq J$. If f is strictly decreasing, $b' < a'$, and f is an order-reversing bijection between the intervals $I' = (b', a') \subseteq I$ and $(a, b) \subseteq J$. In either case, the preimage of a subinterval of (a, b) is an interval, so f is continuous on I'. □

Proof of the Monotonicity Theorem. We have $f\colon (a, b) \to M$ a definable function in an o-minimal structure. Let's say f is good at a point $x \in (a, b)$ if there is some interval containing x such that f is continuous and constant or strictly monotone on that interval. Let

$$X = \{x \in (a, b) \mid f \text{ is good at } x\}.$$

Then X is definable. Consider the definable set $(a,b) \setminus X$. If it is infinite, it contains an interval I such that for all $x \in I$, f is not good at x. But then by Lemmas 4.22, 4.23, and 4.24, I contains a subinterval J such that f is constant (and hence continuous) or strictly monotone and continuous on J. For any $x \in J$, f is good on x, contradiction.

So $(a,b) \setminus X$ is finite. Enumerate its elements as $a_1 < \cdots < a_{k-1}$ and set $a_0 = a$ and $a_k = b$. It remains to show that on each interval $I = (a_i, a_{i+1})$, f is continuous and either constant or strictly monotone on I.

Let $x \in I$. We know that f is good at x, so we have three cases.

Case 1: f is constant on an interval containing x. In particular, the definable set $\{y \in J \mid f(y) = f(x)\}$ contains an interval around x. Let $c \geq a_i$ be the minimal endpoint of such an interval, and let $d \leq a_{i+1}$ be the maximal endpoint of such an interval. If $a_i < c$, then f is good at c, so f is continuous and either constant or strictly monotone on an interval (c', c'') containing c, with $c'' \leq x$. But f is constant on (c, c'') with value $f(x)$, so it must be constant on (c', c'') with value $f(x)$, and hence also on (c', d), contradicting minimality of c. Thus $a_i = c$. Similarly, $d = a_{i+1}$, and f is constant on (a_i, a_{i+1}).

Case 2: f is continuous and strictly increasing on an interval containing x. Let $Y = \{y \in I \mid f \text{ is continuous and strictly increasing on } (y, x)\}$, and let $c = \inf(Y)$. Note $a_i \leq c < x$.

If $a_i < c$, then f is good at c, so f is continuous and either constant or strictly monotone on an interval (c', c'') containing c. But f is strictly increasing on (c, c''), so it must be strictly increasing on (c', c''), and hence on $(c', d) = (c', c'') \cup (c, d)$. Also, if a function is continuous on two open sets, it is continuous on their union, so f is continuous on (c', d). This contradicts minimality of c. Thus $a_i = c$.

Let $Z = \{z \in I \mid f \text{ is continuous and strictly increasing on } (x, z)\}$, and let $d = \sup(Z)$. By a similar argument, $d = a_{i+1}$.

Now fixing an interval $x \in (x', x'')$ such that f is continuous and strictly increasing on (x', x''), we have for all $y \in Y$ and $z \in Z$, $(y, z) = (y, x) \cup (x', x'') \cup (x, z)$, so f is continuous and strictly increasing on (y, z). Now we can write $(a_i, a_{i+1}) = \bigcup_{y \in Y, z \in Z} (y, z)$, and thus f is continuous and strictly increasing on (a_i, a_{i+1}).

Case 3: This argument is exactly the same as the one for Case 2. □

Let's call the points $a = a_0 < a_1 < \cdots < a_k = b$ in the statement of the monotonicity theorem the **breakpoints** of the definable function.

4.3 Limits and derivatives

We'll use the Monotonicity Theorem to study limits and derivatives and show that definable functions in o-minimal structures behave like the functions of "naive" single-variable calculus.

Definition 4.25. Let M be an ordered structure, and let f be a function $D \to M$ for some domain $D \subseteq M$. First, we define left-hand limits at $b \in M$. Suppose the domain of f contains an interval (a, b).

- We say $\lim_{x\to b^-} f(x) = L$ for $L \in M$ if for all $l < L < r$, there exists $b' < b$ such that if $x \in (b', b)$, then $f(x) \in (l, r)$.
- We say $\lim_{x\to b^-} f(x) = \infty$ if for all $m \in M$, there exists $b' < b$ such that if $x \in (b', b)$, then $f(x) \in (m, \infty)$.
- We say $\lim_{x\to b^-} f(x) = -\infty$ if for all $m \in M$, there exists $b' < b$ such that if $x \in (b', b)$, then $f(x) \in (-\infty, m)$.

Next, we define right-hand limits at $b \in M$. Suppose the domain of f contains an interval (b, c).

- We say $\lim_{x\to b^+} f(x) = L$ for $L \in M$ if for all $l < L < r$, there exists $b' > b$ such that if $x \in (b, b')$, then $f(x) \in (l, r)$.
- We say $\lim_{x\to b^+} f(x) = \infty$ if for all $m \in M$, there exists $b' > b$ such that if $x \in (b, b')$, then $f(x) \in (m, \infty)$.
- We say $\lim_{x\to b^+} f(x) = -\infty$ if for all $m \in M$, there exists $b' > b$ such that if $x \in (b, b')$, then $f(x) \in (-\infty, m)$.

Finally, we define two-sided limits at $b \in M$. Suppose the domain of f contains $(a, b) \cup (b, c)$. For $L \in M \cup \{-\infty, \infty\}$, we say $\lim_{x\to b} f(x) = L$ if $\lim_{x\to b^-} f(x) = L$ and $\lim_{x\to b^+} = L$.

Definition 4.26. We say a function $f\colon (a, b) \to M$ is **continuous at a point** $c \in (a, b)$ if $\lim_{x\to c} f(x) = f(c)$.

Exercise 19. Show that a function $f\colon (a, b) \to M$ is continuous in the order topology (meaning that the preimage of every open set is open) if and only if f is continuous at every point in (a, b). Show also that a function $f\colon [a, b] \to M$ is continuous in the order topology if and only if f is continuous at every point in (a, b) and $\lim_{x\to a^+} f(x) = f(a)$ and $\lim_{x\to b^-} f(x) = f(b)$.

Corollary 4.27. *Let M be o-minimal, and let $f\colon (a, b) \to M$ be a definable function. For all $c \in (a, b)$, the limits $\lim_{x\to c^-} f(x)$ and $\lim_{x\to c^+} f(x)$ exist in $M \cup \{\infty, -\infty\}$. Also $\lim_{x\to a^+} f(x)$ and $\lim_{x\to b^-} f(x)$ exist in $M \cup \{\infty, -\infty\}$.*

Proof. Let $a = a_0 < \cdots < a_k = b$ be the breakpoints of f. On each interval $I = (a_i, a_{i+1})$, f is continuous, so for all $c \in I$, $\lim_{x\to c} f(x) = f(c)$. For a breakpoint a_i with $0 < i \leq k$, if f is constant on (a_{i-1}, a_i) with value L, then $\lim_{x\to a_i^-} f(x) = L$.

If f is strictly increasing on (a_{i-1}, a_i), then

$$\lim_{x\to a_i^-} f(x) = \sup\{f(c) \mid c \in (a_{i-1}, a_i)\}.$$

Indeed, letting L be this supremum, either $L \in M$ or $L = \infty$. In either case, for all $l < L$, l is not an upper bound for the values of f, so there is some $c \in (a_{i-1}, a_i)$ such that $l < f(c) \leq L$. Then since f is strictly increasing, for all

$x \in (c, a_i)$, we have $f(x) \in (f(c), L] \subseteq (l, L]$. Similarly, if f is strictly decreasing on (a_{i-1}, a_i), then

$$\lim_{x \to a_i^-} f(x) = \inf\{f(c) \mid c \in (a_{i-1}, a_i)\}.$$

Similar arguments can be applied to the right-hand limits at breakpoints a_i with $0 \leq i < k$. □

Corollary 4.28 (Definable Extreme Value Theorem)**.** *Let $f\colon [a, b] \to M$ be a continuous definable function. Then f takes on a minimum and a maximum value on $[a, b]$.*

Proof. By the monotonicity theorem, we can divide $[a, b]$ into finitely many sub-intervals $[a_i, a_{i+1}]$ such f is constant or strictly monotone on (a_i, a_{i+1}) for all i. It suffices to show that f takes on a minimum and a maximum value on $[a_i, a_{i+1}]$, since then the minimum value of f will be the minimum of these finitely many minimum values, and similarly for the maximum. So we may assume f is constant and strictly monotone on (a, b).

Since f is continuous, $\lim_{x \to a^+} f(x) = f(a)$ and $\lim_{x \to b^-} f(x) = f(b)$. If f is constant on (a, b), then it is constant on $[a, b]$ with value c, and c is the minimum and maximum value of f. If f is strictly increasing on (a, b), then $f(a)$ is the infimum and $f(b)$ is the supremum of the values of f on (a, b). So $f(a)$ is the minimum value of f and $f(b)$ is the maximum value of f. A similar argument applies if f is strictly decreasing on (a, b). □

To define derivatives, we need to be able to add and multiply. By an **o-minimal field**, I mean a structure M in a language containing the language of ordered rings, such that M is an ordered field and M is o-minimal. Note that the language may contain other symbols (like the exponential function). By Corollary 4.19, any o-minimal field is real closed.

Definition 4.29. Let M be an o-minimal field, let $I \subseteq M$ be open, and let $f\colon I \to M$ be a definable function. We say that f is **differentiable** at a point $x \in I$, with derivative $a \in M$, if

$$\lim_{h \to 0} (h^{-1})(f(x + h) - f(x)) = a.$$

We write $f'(x) = a$.

We also define the functions

$$f'_-(x) = \lim_{h \to 0^-} (h^{-1})(f(x + h) - f(x))$$
$$f'_+(x) = \lim_{h \to 0^+} (h^{-1})(f(x + h) - f(x)).$$

By Corollary 4.27, f'_- and f'_+ take values in $M \cup \{-\infty, \infty\}$ for all $x \in I$. Note that if f fails to be differentiable at x, it is because $f'_-(x) \neq f'_+(x)$ or because one of these functions takes on the value $\pm\infty$.

Exercise 20. If $f\colon I \to M$ is a definable function, then $f'(x)$ is also a definable function with domain $\{x \in I \mid f \text{ is differentiable at } x\}$. And $f^-(x)$ and $f^+(x)$ are definable functions, with domain $\{x \in I \mid f'_-(x) \neq \pm\infty\}$ and $\{x \in I \mid f'_+(x) \neq \pm\infty\}$.

Exercise 21. Suppose $I \subseteq M$ is open, $f, g\colon I \to M$ are definable functions, and $c \in M$.

1. If $f = c$ is constant, then $f' = 0$.
2. $(f+g)' = f' + g'$ where f and g are differentiable.
3. $(cf)' = cf'$ where f is differentiable.

Theorem 4.30 (Definable Rolle's Theorem)**.** *Let M be an o-minimal field, and suppose $f\colon [a,b] \to M$ is definable and continuous on $[a,b]$, and differentiable on (a,b). If $f(a) = f(b)$, then for some $c \in (a,b)$, $f'(c) = 0$.*

Proof. By the definable extreme value theorem, f takes on a maximum value and a minimum value on $[a,b]$. If the maximum and minimum are equal, then f is constant, and $f'(c) = 0$ for all $c \in (a,b)$. If not, then since $f(a) = f(b)$, either the maximum or minimum occurs at a point $c \in (a,b)$. Suppose it is the maximum (the minimum case is similar).

Let $h > 0$ such that $c + h \in [a,b]$. Then $f(c+h) \leq f(c)$, so $(h^{-1})(f(c+h) - f(c)) \leq 0$, and $f'_-(c) \leq 0$. Now let $h < 0$ such that $c + h \in [a,b]$. Again $f(c+h) \leq f(c)$, but $h^{-1} < 0$, so $(h^{-1})(f(c+h) - f(c)) \geq 0$, and $f'_+(c) \geq 0$. Since f is differentiable at c, $f'(c) = f'_+(c) = f'_-(c)$, so it must be equal to 0. □

Theorem 4.31 (Definable Mean Value Theorem)**.** *Let M be an o-minimal field, and suppose $f\colon [a,b] \to M$ is definable and continuous on $[a,b]$, and differentiable on (a,b). Then for some $c \in (a,b)$, $f(b) - f(a) = (b-a)f'(c)$.*

Proof. Consider the function $g(x) = f(x) - r(x-a)$ where $r = \frac{f(b)-f(a)}{b-a}$. Then g is definable and continuous (Exercise 18) on $[a,b]$ and differentiable on (a,b) with derivative $g'(x) = f'(x) - r$ (Exercise 21). Since $g(a) = g(b)$, by the definable Rolle's theorem there exists $c \in (a,b)$ such that $g'(c) = 0$. Then $f'(c) - r = 0$, so $f'(c) = r$, and $(b-a)f'(c) = f(b) - f(a)$, as desired. □

Corollary 4.32. *Let M be an o-minimal field, let $I \subseteq M$ be an interval, and suppose $f\colon I \to M$ is definable and continuous on I. If $f'(c) = 0$ for all $c \in I$, then f is constant on I.*

Proof. Choose any $a < b$ in I. Then the hypotheses of the definable mean value theorem hold on $[a,b]$, so there exists $c \in (a,b)$ such that $f(b) - f(a) = (b-a)f'(c) = 0$. So $f(b) = f(a)$, and f is constant on I. □

Lemma 4.33. *Let M be an o-minimal field, $I \subseteq M$ an interval, and $f\colon I \to M$ a definable function. If f is strictly increasing on I, then $f'_+(x) \geq 0$ (possibly*

$f'_+(x) = \infty$) for all $x \in I$. If f is continuous and $f'_+(x) > 0$ (possibly $f'_+(x) = \infty$) for all $x \in I$, then f is strictly increasing on I.

The same statements are true if we replace f'_+ by f'_-.

The analogous statements are also true if we replace increasing by decreasing, and positive by negative.

Proof. Suppose f is strictly increasing on I. Let $x \in I$. For all $h > 0$ such that $x + h \in I$, we have $f(x + h) - f(x) > 0$, so $f'_+(x) \geq 0$. The arguments for f'_- and for the cases that f is strictly decreasing on I are similar.

Suppose now that f is continuous and $f'_+(x) > 0$ for all $x \in I$. Let $a_0 < \cdots < a_k$ be the breakpoints of f from the monotonicity theorem. If there is some interval $J = (a_i, a_{i+1})$ such that f is constant or strictly decreasing on I, then $f'_+(x) \leq 0$ for all $x \in J$ by the previous part, contradiction. Further, f is continuous, so $f(a_i) = \lim_{x \to a_i^+} f(x) = \inf(f(J))$ for $0 < i < k$ and $f(a_{i+1}) = \lim_{x \to a_{i+1}^-} f(x) = \sup(f(J))$ for $0 \leq i < k - 1$. It follows that f is strictly increasing on each generalized interval $(a_0, a_1]$, $[a_i, a_{i+1}]$, and $[a_{k-1}, a_k)$, and hence f is strictly increasing on all of I.

The arguments for f'_- and for the cases that $f'_+(x) < 0$ or $f'_-(0) < 0$ for all $x \in I$ are similar. □

Theorem 4.34. *Let M be an o-minimal field and $f\colon (a, b) \to M$ a definable function. Then f is differentiable at all but finitely many points.*

Proof. By the monotonicity theorem, after throwing away finitely many breakpoints and restricting out attention to the intervals between them, we may assume that f is continuous on (a, b).

Claim: There are only finitely many points $x \in (a, b)$ such that $f'_+(x) = \pm\infty$ or $f'_-(x) = \pm\infty$.

Suppose for contradiction the definable set $\{x \in (a, b) \mid f'_+(x) = \infty\}$ is infinite. The cases for $-\infty$ and f'_- are handled similarly. Then this set contains an interval I, and by Lemma 4.33, f is strictly increasing on I. Again by Lemma 4.33, $f'_-(x) \geq 0$ for all $x \in I$. In particular, $f'_-(x) \neq -\infty$ for all $x \in I$.

Now I is partitioned into the definable sets $\{x \in I \mid f'_-(x) = \infty\}$ and $\{x \in I \mid f'_-(x) \in M\}$, so one of them contains a subinterval $J \subseteq I$. We will show that we get a contradiction in either case.

Case 1: $f'_-(x) = \infty$ for all $x \in J$. Since f is strictly increasing, it is injective, so it has an inverse function $g\colon f(J) \to J$, which is definable. Since f is continuous, every subinterval of J is definably connected, so its image is definably connected, and hence an interval (since f is strictly increasing). Thus the preimage of an interval under g is an interval, so g is continuous.

We wish to show that $g'(y) = 0$ for all $y \in f(J)$. Then by Corollary 4.32, g is constant function, contradicting the fact that J is infinite.

Let $y \in f(J)$ and $x = g(y) \in J$, so $f(x) = y$. Let $0 < \varepsilon$ be small. Then since $\lim_{h \to 0} \frac{f(x+h)-f(x)}{h} = \infty$, there is some $0 < \delta$ such that for all $z \neq x$ in $(x - \delta, x + \delta)$, taking $h = z - x$, we have $\frac{f(z)-f(x)}{z-x} > 1/\varepsilon$. Let $0 < \delta'$ be small enough so that for all $z' \in (y - \delta', y + \delta')$, we have $g(z') \in (x - \delta, x + \delta)$

(using continuity of g). Then for all $z' \in (y-\delta', y+\delta')$ with $z' \neq y$, we have $g(z') \in (x-\delta, x+\delta)$ and $g(z') \neq x$, so

$$\frac{f(g(z')) - f(x)}{g(z') - x} > 1/\varepsilon$$

and taking reciprocals,

$$\frac{g(z') - g(y)}{z' - y} < \varepsilon.$$

Thus $g'(y) = 0$, as desired.

Case 2: $f'_-(x) \in M$ for all $x \in J$. By the monotonicity theorem, there is a subinterval $J' \subseteq J$ on which f'_- is continuous. Pick $c \in J'$ and some $m \in M$ such that $f'_-(c) < m$. Since f'_- is continuous, there is some interval $c \in J'' \subseteq J'$ such that $f'_-(x) < m$ for all $x \in J''$.

Now consider the definable function $g(x) = f(x) - mx$. We have $g'_-(x) = f'_-(x) - m < 0$ for all $x \in J''$. But for all $x \in J''$ we have $f'_+(x) = \infty$, so for all $l \in M$ and all sufficiently small $h > 0$, $\frac{f(x+h)-f(x)}{h} > (l+m)$, so $\frac{f(x+h)-m(x+h)-(f(x)-mx)}{h} > l$, and $g'_+(x) = \infty$. By Lemma 4.33, g is both strictly increasing and strictly decreasing on J'', which is a contradiction.

By the claim and the monotonicity theorem applied to f'_+ and f'_-, after throwing away finitely many points and restricting our attention to the intervals between them, we may assume that f'_+ and f'_- take their values in M and are both continuous on (a,b). We will show that $f'_+ = f'_-$, so f is differentiable on (a,b).

Suppose that for some $c \in (a,b)$, $f'_+(c) \neq f'_-(c)$. Assume $f'_-(c) < f'_+(c)$ (the other case is similar). Pick some $m \in M$ such that $f'_-(c) < m < f'_+(c)$. Since f'_- and f'_+ are continuous, there is some interval $c \in I \subseteq (a,b)$ such that $f'_-(x) < m$ and $f'_+(x) > m$ for all $x \in I$.

Now consider the definable function $g(x) = f(x) - mx$. We have $g'_-(x) = f'_-(x) - m < 0$ and $g'_+(x) = f'_+(x) - m > 0$ for all $x \in I$. By Lemma 4.33, g is both strictly increasing and strictly decreasing on I, which is a contradiction. □

Recall that a function is C^k at a point a if the functions $f, f', f'', \ldots, f^{(k)}$ are defined and continuous at a.

Corollary 4.35. *Let M be an o-minimal field and $f\colon (a,b) \to M$ a definable function. Let $k \in \mathbb{N}$. Then f is C^k at all but finitely many points.*

Proof. By induction on k. When $k = 0$, a C^0 function is just a continuous function. By the Monotonicity Theorem, f is continuous at all but finitely many points.

Now let $k > 0$. By the monotonicity theorem and Theorem 4.34, there are finitely many points $a = c_0 < c_1 < \cdots < c_n = b$ such that f is continuous and differentiable on each interval $I = (c_i, c_{i+1})$. Then f' is a definable function $I \to M$. By induction, f' is C^k at all but finitely many points in I, so f is C^{k+1} at all but finitely many points in I. Putting together the c_i and the finitely many exceptional points in each interval proves the corollary. □

4.4 Cell decomposition

We have seen that o-minimality, which is a condition on definable subsets of M, has strong consequences for the structure of definable functions $M \to M$. Our next goal is to prove that o-minimality also has strong consequences for the structure of definable subsets of M^n and definable functions $M^n \to M$.

For a definable set $X \subseteq M^n$, we define

$$C(X) = \{f \colon X \to M \mid f \text{ is continuous and definable}\}$$
$$C_\infty(X) = C(X) \cup \{-\infty, \infty\}$$

where we view $\pm\infty$ as "constant functions" on X with value $\pm\infty$. For $f, g \in C_\infty(X)$, we write $f < g$ if $f(x) < g(x)$ for all $x \in X$. Note that for any $f \neq -\infty$ and any $g \neq \infty$, we have $-\infty < f$ and $g < \infty$.

Definition 4.36. We define by induction on n what it means for a set $X \subseteq M^n$ to be a **cell**. Simultaneously, we define the **type** of the cell, which is a binary sequence of length n.

- Base case: $M^0 = \{()\}$, the singleton set containing the empty tuple. The only cell in M^0 is $\{()\}$, and its type is the empty sequence.
- Inductive case: There are two kinds of cell in M^{n+1}.
  - **Thin:** Let $X \subseteq M^n$ be a cell of type s, and let $f \in C(X)$. Then
  $$\Gamma(f) = \{(\overline{a}, b) \mid \overline{a} \in X \text{ and } f(\overline{a}) = b\} \subseteq M^{n+1}$$
  is a cell of type $s0$ (append a 0 to the end of s).
  - **Wide:** Let $X \subseteq M^n$ be a cell of type s, and let $f, g \in C_\infty(X)$ such that $f < g$. Then
  $$(f, g) = \{(\overline{a}, b) \mid \overline{a} \in X \text{ and } f(\overline{a}) < b < g(\overline{a})\} \subseteq M^{n+1}$$
  is a cell of type $s1$ (append a 1 to the end of s).

Let $X \subseteq M^n$ be a cell of type s. For $k < n$, we say X is **thin in component** k if $s(k) = 0$ and X is **wide in component** k if $s(k) = 1$.

Remark 4.37. Here are some observations about the definition.

(1) It is worth considering the special case of cells in M^1. The only cell in M^0 is $X = \{()\}$. Let $f \colon X \to M^1$ be a function. Then f is determined by its value on $()$, which is a point $b \in M$. Its graph is $\Gamma(f) = \{b\}$. Now any such function is continuous and definable (since $\{b\}$ is definable by $x = b$), so $C(X)$ is the set of all such functions.

Thus a thin cell in M^1 is exactly a point $\{a\}$, and a wide cell in M^1 is exactly an interval (a, b), $(-\infty, b)$, or (a, ∞).

(2) What about cells in M^2? A cell of type 00 is again a point. A cell of type 01 is an interval in a vertical fiber above a point: $\{(x, y) \mid x = a \text{ and } y \in (b, c)\}$. A cell of type 10 is the graph of a continuous definable function $f\colon (a, b) \to M$. And a cell of type 11 is (f, g) for some $f, g \in C_\infty((a, b))$ with $f < g$.

(3) Note that the definition of cell is not symmetric in the order of components. For example, writing $\sigma\colon M^2 \to M^2$ for the map $(x, y) \mapsto (y, x)$, the image of a cell under σ is not necessarily a cell. You can see this by observing how much more restrictive cells of type 01 are than cells of type 10. The privileging of the last component makes inductive arguments on dimension much easier.

(4) Write $\pi\colon M^{n+1} \to M^n$ for the projection map which drops the last coordinate. It follows immediately from the definition that if $X \subseteq M^{n+1}$ is a cell, then $\pi(X) \subseteq M^n$ is a cell. This is one of the benefits of privileging the last component.

(5) It is easy to show by induction on n that every cell in M^n is a definable set.

Our goal is to prove the following theorem, which generalizes the definition of o-minimality to definable sets in higher dimensions.

Theorem 4.38 (Cell Decomposition)**.** *Let M be an o-minimal structure and $X \subseteq M^n$ a definable set. Then X is a finite disjoint union of cells in M^n.*

Example 4.39. Let's consider a cell decomposition for the closed unit ball

$$B_2 = \{(x, y) \mid x^2 + y^2 \leq 1\} \subseteq \mathbb{R}^2.$$

We have $\pi(B_2) = [-1, 1]$ in $\mathbb{R}^1$, and each cell in $\mathbb{R}^2$ projects to a cell in $\mathbb{R}^1$, so we first take a cell decomposition $[-1, 1] = \{-1\} \cup (-1, 1) \cup \{1\}$ of the interval.

Now

$$B_2 = \{(-1, 0)\} \cup \Gamma(f) \cup \Gamma(g) \cup (f, g) \cup \{(1, 0)\},$$

where $f\colon (-1, 1) \to \mathbb{R}$ is the function defined by $(x^2 + y^2 = 1) \wedge (y < 0)$ and $g\colon (-1, 1) \to \mathbb{R}$ is the function defined by $(x^2 + y^2 = 1) \wedge (0 < y)$. These cells have type 00, 10, 10, 11, and 00, respectively.

Similarly, we can give a cell decomposition of the closed unit ball B_3 in $\mathbb{R}^3$. We have $\pi(B_3) = B_2$, and again, each cell in $\mathbb{R}^3$ projects to a cell in $\mathbb{R}^2$, so it makes sense to start with the cell decomposition of B_2.

The fibers over the points $\{(-1, 0)\}$ and $\{(1, 0)\}$ are just the antipodal points $(-1, 0, 0)$ and $(1, 0, 0)$, which are cells. The fibers over the points on $\Gamma(f)$ and $\Gamma(g)$ above are single points with z-value 0. We can form the upper and lower circumferences of the intersection of B_3 with the xy-plane as cells: the graphs of the constant function with value 0 over the cells $\Gamma(f)$ and $\Gamma(g)$. Finally, the fibers over the points on the open ball (f, g) above are closed intervals. We can finish with the following cells: the lower half-sphere (the graph of the function on (f, g) defined by $(x^2 + y^2 + z^2 = 1) \wedge (z < 0)$), the upper half-sphere (the

graph of the function on (f, g) defined by $(x^2 + y^2 + z^2 = 1) \wedge (0 < z)$), and the open unit ball (the wide cell between these two definable functions).

The above example suggests a strategy for proving the cell decomposition theorem by induction on n. Let's outline this strategy and point out what additional results we need to make it work.

Let $X \subseteq M^{n+1}$ be a definable set. By induction, $\pi(X) \subseteq M^n$ has a cell decomposition. Let Y be one of the cells in the decomposition. Then for all $\overline{a} \in Y$, the fiber $X_a = \{b \in M \mid (\overline{a}, b) \in X\}$ is definable, so it is a finite union of points and intervals. The points should be pieced together across various fibers to form thin cells over Y, and the intervals should be pieced together across various fibers to form wide cells over Y.

One issue is that "shape" of the fiber $X_{\overline{a}}$ may change as $\overline{a}$ varies across Y. For $\overline{a}, \overline{b} \in Y$, $X_{\overline{a}}$ may consist of a point followed by an interval, while $X_{\overline{b}}$ may consist of 7 points, followed by 4 intervals, followed by a point. After checking that the shape of the fiber $X_{\overline{a}}$ is a definable property of $\overline{a}$, we can further decompose Y into subsets corresponding to fiber shapes. But if we want a finite cell decomposition, we need to prove that only finitely many shapes are possible.

If we can accomplish this, then we can assume that for all $\overline{a} \in Y$, the fiber $X_{\overline{a}}$ has a fixed shape. Now we can define functions $f \colon Y \to M$ such that $f(\overline{a})$ is a point or an interval endpoint in the fiber $X_{\overline{a}}$. But we cannot directly use these functions to define cells, since they may be discontinuous.

To finish, we need a higher-dimensional version of the monotonicity theorem, which says that for any definable function $f \colon Y \to M$ with $Y \subseteq M^n$, we can partition Y into finitely many definable pieces, such that the restriction of f to each piece is continuous. It turns out that we can also prove this statement by induction on n, but we need to use cell decomposition for definable subsets of M^n as a hypothesis. So this leads us to prove both cell decomposition for subsets of M^n and piecewise continuity for definable functions $M^n \to M$ by an intwined induction on n.

Theorem 4.40 (Piecewise Continuity)**.** *Let M be an o-minimal structure. For every definable function $f \colon X \to M$, where $X \subseteq M^n$, there is a decomposition $X = X_1 \cup \cdots \cup X_k$ of X into a finite union of definable sets such that $f|_{X_i}$ is continuous for each $1 \leq i \leq k$.*

The proof of the cell decomposition and piecewise continuity theorems is lengthy, so we'll carry it out in pieces.

Proof of the main theorems, Part 1. Let M be an o-minimal structure. We prove the following statements by induction on n:

$(1)_n$ Every definable set $X \subseteq M^n$ is a finite disjoint union of cells.

$(2)_n$ For every definable function $f \colon X \to M$, where $X \subseteq M^n$, there is a decomposition $X = X_1 \cup \cdots \cup X_k$ of X into a finite union of definable sets such that $f|_{X_i}$ is continuous for each $1 \leq i \leq k$.

Our base cases are $n = 0$ and $n = 1$. The cases when $n = 1$ actually follow by our general inductive arguments, but we handle them separately to avoid confusion.

$(1)_0$: If $X \subseteq M^0$, then $X = \emptyset$ (the empty union) or $X = \{()\}$ (a cell).

$(2)_0$: If $f\colon X \to M$ is a definable function, where $X \subseteq M^0$, then f is trivially continuous.

$(1)_1$ follows from o-minimality.

$(2)_1$ follows from the monotonicity theorem: By o-minimality, we can decompose X into a finite union of points and intervals, and the restriction of f to any point is trivially continuous. For each interval (a, b) in the decomposition of X, by the monotonicity theorem, we can decompose (a, b) into a finite union of points (the breakpoints) and intervals (between them), such that f is continuous on each interval. Again, f is trivially continuous when restricted to each breakpoint. □

Toward the inductive step $(1)_{n+1}$, we need some preparatory lemmas. First, we show that $(1)_n$ implies that we can simultaneously decompose finitely many definable sets into cells.

Lemma 4.41. *Assume every definable subset of M^n is a finite disjoint union of cells. Let $X_1, \ldots, X_k \subseteq M^n$ be definable sets. Then there is a partition of M^n into a set $\mathcal{C}$ of finitely many cells such that for all i, X_i is a finite union of cells from $\mathcal{C}$.*

Proof. First, we make a "Venn diagram". For each i, let $X_i' = M^n \setminus X_i$. For each set $S \subseteq \{1, \ldots, k\}$, define

$$X_S = \bigcap_{i \in S} X_i \cap \bigcap_{i \notin S} X_i'.$$

Each set X_S is definable, the set X_S are pairwise disjoint, and we have

$$X_i = \bigcup_{S \text{ s.t. } i \in S} X_S \quad \text{for all } 1 \leq i \leq k$$
$$M^n = \bigcup_{S \subseteq \{1,\ldots,k\}} X_S.$$

By hypothesis, each X_S is a disjoint union of finitely many cells. Let $\mathcal{C}$ be the set of all these cells, for all $S \subseteq \{1, \ldots, k\}$. Then the cells in $\mathcal{C}$ partition M^n, since the X_S do. And for each $1 \leq i \leq k$, X_i is a disjoint union of cells from $\mathcal{C}$, since each X_S is. □

Next, we analyze the fibers of definable sets. To do this, the language of definable families will be useful.

Definition 4.42. Let $\varphi(\overline{x}, \overline{y})$ be a formula in context $\overline{xy}$, with its variables partitioned into two tuples $\overline{x}$ and $\overline{y}$, and let $\psi(\overline{x})$ be a formula in context $\overline{x}$. The **definable family of definable sets** defined by φ and ψ is

$$\{\varphi(\overline{a}, M) \subseteq M^n \mid \overline{a} \in \psi(M) \subseteq M^m\}$$

where m is the length of $\overline{x}$ and n is the length of $\overline{y}$.

If $X \subseteq M^{m+n}$ is the set defined by φ, we often write $X_{\overline{a}}$ for the definable set $\varphi(\overline{a}, M) \subseteq M^n$, and if $Y \subseteq M^m$ is the set defined by ψ, we often write $(X_{\overline{a}})_{\overline{a} \in Y}$ for the definable family of definable sets. Note that $X_{\overline{a}}$ is the fiber of of the projection map $\pi \colon X \to M^m$ over the point $\overline{a} \in M^m$.

Lemma 4.43 (Uniform finiteness)**.** *Let M be an o-minimal structure, and let $(X_{\overline{a}})_{\overline{a} \in Y}$ be a definable family of definable sets, where $X_{\overline{a}} \subseteq M^1$ for all $\overline{a} \in Y$. If each set $X_{\overline{a}}$ is finite, then there exists $N \in \mathbb{N}$ such that $|X_{\overline{a}}| \leq N$ for all $\overline{a} \in Y$.*

Proof. Suppose not. Let $\varphi(\overline{x}, y)$ be the formula defining the family. Let $\mathcal{L}'$ be the language obtained by adding a tuple $\overline{c}$ of new constant symbols to $\mathcal{L}$, where $\overline{c}$ has the same length as $\overline{x}$.

Consider the theory $T = \mathrm{Th}(M) \cup \{\varphi_n \mid n \in \mathbb{N}\} \cup \{\chi\}$, where φ_n is the sentence expressing that $|X_{\overline{c}}| \geq n$:

$$\exists y_1 \ldots \exists y_n \left(\bigwedge_{i \neq j} y_i \neq y_j \wedge \bigwedge_{i=1}^{n} \varphi(\overline{c}, y_i) \right)$$

and χ is the sentence expressing that $X_{\overline{c}}$ does not contain an interval:

$$\neg \exists y \, \exists y' \, \forall z \, ((y < z < y') \to \varphi(\overline{c}, z)).$$

By the compactness theorem, T is satisfiable. Indeed, any finite subset of T is contained in $T_N = \mathrm{Th}(M) \cup \{\varphi_n \mid n \leq N\} \cup \{\chi\}$, and interpreting $\overline{c}$ as a tuple $\overline{a} \in Y$ such that $|X_{\overline{a}}| \geq N$, we have $(M, \overline{a}) \models T_N$ (note $X_{\overline{a}}$ does not contain an interval, since it is finite).

But if $(N, \overline{c}) \models T$ is a model, $N \models \mathrm{Th}(M)$, so N is an o-minimal structure. But $X_{\overline{c}} \subseteq N^1$ is an infinite definable set in N which does not contain an interval, contradicting o-minimality. □

Remark 4.44. The use of compactness in the proof is "cheating", in a sense. It is possible only because we defined "o-minimal structure" by asserting the o-minimality condition in every model of the complete theory of the structure. As discussed in Remark 4.7, the original definition of "o-minimal structure" only asserted the o-minimality condition for definable subsets that structure.

It is possible to give a direct combinatorial proof of the uniform finiteness lemma, without using compactness or our stronger definition of o-minimality, and this lemma is crucial to the Knight–Pillay–Steinhorn theorem that if a single structure is o-minimal in the weaker sense, then every model of its complete theory is o-minimal. But this direct proof is much trickier (it is comparable in difficulty to the proof of the monotonicity theorem), and by the Knight–Pillay–Steinhorn theorem, we don't lose any examples by defining o-minimality in the way we did, which allows for this easier model-theoretic proof.

When working with a definable family of definable sets $(X_{\overline{a}})_{\overline{a} \in Y}$, we will often refer to a property of definable sets or an operation on definable sets being **uniformly definable** in the family.

To say that a property P of the $(X_{\overline{a}})_{\overline{a} \in Y}$ is uniformly definable means that there is a formula $\chi(\overline{x})$ such that for all $\overline{a} \in Y$, $M \models \chi(\overline{a})$ if and only if P is true of $X_{\overline{a}}$. The proof of Lemma 4.43 relied crucially on the observation that "$X_{\overline{a}}$ is finite" is uniformly definable by "$X_{\overline{a}}$ does not contain an interval".

If we have some operation F on definable sets, we say it is uniformly definable if there is a definable family of definable sets $(Z_{\overline{a}})_{\overline{a} \in Y}$ such that for all $\overline{a} \in Y$, $Z_{\overline{a}} = F(X_{\overline{a}})$. The next proposition gives three examples.

Proposition 4.45. *In an ordered structure, closure, interior, and boundary (interior minus closure) of definable sets are uniformly definable.*

Proof. We already showed in Proposition 4.12 that the closure, interior, and boundary of a definable set are definable. It remains to verify that these definitions were uniform.

For example, if $\varphi(\overline{x}, \overline{y})$ is a formula defining a family of definable sets $(X_{\overline{a}})_{\overline{a} \in Y}$, each defined by $\varphi(\overline{a}, \overline{y})$, then let $\chi(\overline{x}, \overline{y})$ be the following formula:

$$\exists \overline{z}\, \exists \overline{w}\, \left(\left(\bigwedge_{i=1}^{n} z_i < y_i < w_i \right) \wedge \forall \overline{y}' \left(\left(\bigwedge_{i=1}^{n} z_i < y_i' < w_i \right) \rightarrow \varphi(\overline{x}, \overline{y}') \right) \right).$$

We have that $\chi(\overline{a}, \overline{y})$ defines the interior of $X_{\overline{a}}$ for all $\overline{a} \in Y$, so $\chi(\overline{x}, \overline{y})$ defines the family $(\mathrm{int}(X_{\overline{a}}))_{\overline{a} \in Y}$.

The same considerations apply to closures and boundaries. □

In an o-minimal structure M, if a definable set $X \subseteq M^1$ is decomposed into a disjoint union of points and intervals, then $\mathrm{bd}(X)$ consists of the isolated points and the endpoints of the intervals. For example, if $X = \{0\} \cup (0,1) \cup (1,2) \cup \{3\}$, then $\mathrm{bd}(X) = \{0, 1, 2, 3\}$. It follows that $\mathrm{bd}(X)$ is finite. And if we list $\mathrm{bd}(X)$ as $b_1 < b_2 < \cdots < b_k$, then for each interval $I = (-\infty, b_1)$, (b_i, b_{i+1}), or (b_k, ∞), either $I \subseteq X$ or I is disjoint from X.

We define the **shape** of X, $\mathrm{sh}(X)$, to be a binary sequence $t_0 s_1 t_1 s_2 \ldots s_k t_k$ of length $2k+1$, where $s_i = 1$ if $b_i \in X$ and $t_i = 1$ if $(b_i, b_{i+1}) \subseteq X$, defining $b_0 = -\infty$ and $b_{k+1} = \infty$. Note that not every sequence can arise as a shape: we will not have $t_{i-1} = s_i = t_{i+1} = 1$ for any i, since then b_i would be an interior point.

Lemma 4.46. *Let M be an o-minimal structure. Let $(X_{\overline{a}})_{\overline{a} \in Y}$ be a definable family of definable sets, such that $X_{\overline{a}} \subseteq M^1$ for all $\overline{a} \in Y$.*

(1) Y can be partitioned into finitely many definable sets $Y = Y_1 \cup \cdots \cup Y_n$ such that for each i, and for all $\overline{a}, \overline{a}' \in Y_i$, $|\mathrm{bd}(X_{\overline{a}})| = |\mathrm{bd}(X_{\overline{a}'})|$ and $\mathrm{sh}(X_{\overline{a}}) = \mathrm{sh}(X_{\overline{a}'})$.

(2) Fix an i and let $k = |\mathrm{bd}(X_{\overline{a}})|$ for all $\overline{a} \in Y_i$. For all $1 \leq j \leq k$, write $b_j \colon Y_i \to M$ for the function which sends $\overline{a}$ to the j^{th} element of $\mathrm{bd}(X_{\overline{a}})$ when listed in increasing order. Then b_j is a definable function.

Proof. By the discussion above, for all $\overline{a} \in Y$, $\mathrm{bd}(X_{\overline{a}})$ is finite. By Lemma 4.43, there is an upper bound N such that $|\mathrm{bd}(X_{\overline{a}})| \leq N$ for all $\overline{a} \in Y$. Now for each $\overline{a} \in Y$, $\mathrm{sh}(X_{\overline{a}})$ is a binary sequence of length at most $2N+1$, and in particular only finitely many shapes are possible.

For each shape $s = t_0 s_1 t_1 s_2 \dots s_k t_k$ with $k \leq N$, we wish to show that $Y_s = \{\overline{a} \in Y \mid \mathrm{sh}(X_{\overline{a}}) = s\}$ is a definable set (possibly empty). To define Y_s, we write down the following conditions:

1. There exist $b_1 < \cdots < b_k$ in $\mathrm{bd}(X_{\overline{a}})$, and any point in $\mathrm{bd}(X_{\overline{a}})$ is equal to one of the b_i.
2. For each $1 \leq i \leq i$, if $s_i = 1$, then $b_i \in X_{\overline{a}}$, and if $s_i = 0$, then $b_i \notin X_{\overline{a}}$.
3. For each $0 \leq i \leq k$, if $t_i = 1$, then $(b_i, b_{i+1}) \subseteq X_{\overline{a}}$, and if $t_i = 0$, then $(b_i, b_{i+1}) \cap X_{\overline{a}} = \emptyset$, where we set $b_0 = -\infty$ and $b_k = \infty$.

Given such a set Y_s, the function $b_j \colon Y_s \to M$ is defined by: there exist $y_1 < \cdots < y_k$ in $\mathrm{bd}(X_{\overline{a}})$, and $b_j(\overline{a}) = y_j$. □

Proof of the main theorems, Part 2. We assume $(1)_n$ and $(2)_n$ and prove $(1)_{n+1}$.

Let $X \subseteq M^{n+1}$ be a definable set and $Y = \pi(X) \subseteq M^n$. By Lemma 4.46(1), Y can be partitioned into finitely many definable sets $Y = Y_1 \cup \cdots \cup Y_m$ such that for each i, $\mathrm{sh}(X_{\overline{a}})$ is the same for all $\overline{a} \in Y_i$. It suffices to show that for each i, $X \cap \pi^{-1}(Y_i)$ is a finite union of cells, since $X = \bigcup_{i=1}^m (X \cap \pi^{-1}(Y_i))$.

So replacing Y with Y_i and X with $X \cap \pi^{-1}(Y_i)$, we may assume that $\mathrm{sh}(X_{\overline{a}}) = t_0 s_1 t_1 s_2 \dots s_k t_k$ for all $\overline{a} \in Y$. By Lemma 4.46(2), writing $b_j \colon Y_i \to M$ for the function which sends $\overline{a}$ to the j^{th} element of $\mathrm{bd}(X_{\overline{a}})$ when listed in increasing order, b_j is a definable function for $1 \leq j \leq k$.

For each j, by $(2)_n$, Y can be written as a finite union $Y = Y_1^j \cup \cdots \cup Y_{m_j}^j$ of definable sets such that $b_j|_{Y_\ell^j}$ is continuous for each $1 \leq \ell \leq m_j$. By $(1)_n$, we can apply Lemma 4.41 to the finitely many definable sets Y_ℓ^j for $1 \leq j \leq k$ and $1 \leq \ell \leq m_j$, obtaining a finite collection of cells $\mathcal{C}$ which partitions M^n and such that each Y_ℓ^j is a union of cells from $\mathcal{C}$. Now $\mathcal{C}_Y = \{C \in \mathcal{C} \mid C \subseteq Y\}$ is a partition of Y. And for each cell $C \in \mathcal{C}_Y$, for each j, C is contained in some set Y_ℓ^j, so $b_j|_C$ is continuous.

Now $X = \bigcup_{C \in \mathcal{C}_Y} (X \cap \pi^{-1}(C))$, so it suffices to show that each set $X_C = (X \cap \pi^{-1}(C))$ is a finite disjoint union of cells. We have arranged that $\pi(X_C) = C$ is a cell, every fiber of X_C has the same shape $t_0 s_1 t_1 s_2 \dots s_k t_k$, and the functions $b_j|_C \colon C \to M$ which pick out the boundary points of the fibers are continuous. So we have

$$X_C = \bigcup_{s_j = 1} \Gamma(b_j|_C) \cup \bigcup_{t_j = 1} (b_j|_C, b_{j+1}|_C),$$

where we define $b_0 = -\infty$ and $b_{k+1} = \infty$. □

The hardest part is yet to come! Toward the inductive step $(2)_{n+1}$, we need more lemmas. To begin, we will establish some facts of independent interest about the topology and dimension of cells.

Definition 4.47. Let $X \subseteq M^n$ be a cell of type s. The **dimension** of X is $\dim(X) = \sum_{i=1}^{n} s(n)$. This is the number of components in which X is wide.

Proposition 4.48. *A cell $X \subseteq M^n$ with $\dim(X) = n$ is an open set in M^n.*

Proof. By induction on n. In the base case, M^0 is a discrete space with one element, so the unique cell in M^0 (which has dimension 0) is open.

Now suppose $X \subseteq M^{n+1}$ with $\dim(X) = n+1$. Let $Y = \pi(X) \subseteq M^n$. Then Y is a cell, and X is wide over Y, so $X = (f, g)$, with $f, g \in C_\infty(Y)$. Since X is wide in every component, so is Y, and $\dim(Y) = n$. By induction, Y is open.

For any $p \in X$, we can write $p = (y, c)$ with $y \in Y$, and $f(y) < c < g(y)$. Pick some $a, b \in M$ with $f(y) < a < c < b < g(y)$, and define

$$Z = Y \cap f^{-1}(-\infty, a) \cap g^{-1}(b, \infty),$$

which is an open neighborhood of y in M^n by continuity of f and g (if $f = -\infty$ or $g = \infty$, omit this term from the intersection). Let B be an open box such that $y \in B \subseteq Z$. Then $p \in B \times (a, b) \subseteq X$, since for all $z \in B$ and all $d \in (a, b)$, we have $z \in Y$ and $f(z) < a < d < b < g(z)$, so $(z, d) \in X$. This shows X is open. □

The next result is a strong converse: finitely many cells of dimension less than n in M^n cannot cover an open box. In particular, a single cell of dimension less than n is not open.

Proposition 4.49. *Suppose $X_1, \ldots, X_k \subseteq M^n$ are cells with $\dim(X_i) < n$ for all $1 \leq i \leq k$. Then $\bigcup_{i=1}^{k} X_i$ has empty interior.*

Proof. We prove the statement by induction on k. The base case $k = 0$ is trivial, since the empty union $\emptyset$ has empty interior.

Now suppose for contradiction that $\left(\bigcup_{i=1}^{k} X_i\right) \cup X$ has non-empty interior, and let B be an open box which is contained in the union. It suffices to find some smaller open box $B' \subseteq B$ which is disjoint from X, since then $B' \subseteq \bigcup_{i=1}^{k} X_i$, contradicting the inductive hypothesis that $\bigcup_{i=1}^{k} X_i$ has empty interior.

Since $\dim(X) < n$, there is some $1 \leq j \leq n$ such that X is thin in component j. Let j be the least with this property, and write π^* for the projection onto the first $j-1$ coordinates. If $Y = \pi^*(X)$, then Y is an open cell, and there is a function $f \in C(Y)$ such that every point in X has the form $(y, f(y), z)$, where $y \in Y \subseteq M^{j-1}$ and $z \in M^{n-j}$.

Let $\overline{b} \in B$ be a point such that $\pi^*(\overline{b}) \in Y$ (if there is no such point in B, then B is already disjoint from X, and we are done), and let b_j be its j-coordinate. Since B is an open box, we may assume $b_j \neq f(\pi^*(\overline{b_j}))$, by changing b_j if necessary without leaving B. Suppose $b_j < f(\pi^*(\overline{b}))$ (the other case is similar).

Pick some c with $b_j < c < f(\pi^*(\overline{b}))$. Since f is continuous, $U = f^{-1}(c, \infty) \subseteq M^{j-1}$ is open. Then $U \times (-\infty, c) \times M^{n-j}$ is an open neighborhood of $\overline{b}$ which is disjoint from X. So we can find a small open box $\overline{b} \in B' \subseteq B$ which is disjoint from X, as desired. □

In one dimension, we have the useful principle that if an infinite definable set is partitioned into finitely many pieces, one of them contains an interval. The following result generalizes this to higher dimensions.

Corollary 4.50. *Assume that every definable subset of M^n is a finite disjoint union of cells. Suppose $X \subseteq M^n$ is a definable set with non-empty interior. If $X = X_1 \cup \cdots \cup X_k$, where each X_i is definable, then for some $1 \leq i \leq k$, X_i has non-empty interior.*

Proof. Decomposing each X_i as a finite disjoint union of cells, we can write X as a finite union of cells, each of which is contained in some X_i. Since X has non-empty interior, by Proposition 4.49, at least one of these cells C has dimension n. By Proposition 4.48, C is open, and $C \subseteq X_i$ for some i, so X_i has non-empty interior. □

Proposition 4.51. *Let $X \subseteq M^n$ be a cell with $\dim(X) = d$. Then X is definably homeomorphic to an open cell $U \subseteq X^d$ of dimension d.*

Proof. We proceed by induction on n. In the base case, when $n = 0$, $X \subseteq M^0$ is the singleton cell of dimension 0, which is open in M^0.

Now let $X \subseteq M^{n+1}$ be a cell of dimension d.

Case 1: X is thin. Let $Y = \pi(X) \subseteq M^n$ and $X = \Gamma(f)$, with $f \in C(Y)$. Then $\pi \colon X \to \pi(X)$ is a definable continuous bijection, and its inverse $y \mapsto (y, f(y))$ is also continuous, so π is a definable homeomorphism. Since $\dim(Y) = \dim(X) = d$, we are done by the inductive hypothesis.

Case 2: X is wide. Let $Y = \pi(X) \subseteq M^n$ and $X = (f, g)$, with $f, g \in C_\infty(Y)$ and $f(y) < g(y)$ for all $y \in Y$. Since $\dim(Y) = \dim(X) - 1 = d - 1$, by induction there is a definable homeomorphism $h \colon Y \to U \subseteq M^{d-1}$, where U is an open cell of dimension $d - 1$. Now $f \circ h^{-1}$ and $g \circ h^{-1}$ are in $C_\infty(U)$, and $f \circ h^{-1}(u) < g \circ h^{-1}(u)$ for all $u \in U$, so $V = (f \circ h^{-1}, g \circ h^{-1}) \subseteq M^d$ is a cell of dimension d. And the map $X \to V$ which is h on the first $n - 1$ components and the identity on the last component is a definable homeomorphism. □

We use the following lemma to "bootstrap" continuity up a dimension. When we say a function f is **monotone**, we mean that $a \leq b$ implies $f(a) \leq f(b)$ or $a \leq b$ implies $f(a) \geq f(b)$. Note that this is weaker than the notion of strictly monotone used in the monotonicity theorem. In particular, constant functions are monotone but not strictly monotone.

Lemma 4.52. *Suppose X is a topological space and $(R_1; \leq)$ and $(R_2; \leq)$ are dense linear orders without endpoints. We equip R_1 and R_2 with the order topology and $X \times R_1$ with the product topology. Suppose $f \colon X \times R_1 \to R_2$ is a function such that:*

(i) For all $x \in X$, $f(x, \cdot) \colon R_1 \to R_2$ is continuous and monotone.

(ii) For all $r \in R_1$, $f(\cdot, r) \colon X \to R_2$ is continuous.

Then f is continuous.

Proof. Let $J \subseteq R_2$ be an interval, and let $(x, r) \in f^{-1}(J)$. In order to show that $f^{-1}(J)$ is open, we must find an open neighborhood $x \in U \subseteq X$ and an interval $r \in I \subseteq R_1$ such that $U \times I \subseteq f^{-1}(J)$.

Since $f(x, \cdot)$ is continuous, there is an interval $I = (a, b)$ with $a < r < b$ such that $f(x, a) \in J$ and $f(x, b) \in J$. Now since $f(\cdot, a)$ and $f(\cdot, b)$ are continuous, there is an open neighborhood $x \in U \subseteq X$ such that for all $x' \in U$, $f(x', a) \in J$ and $f(x', b) \in J$.

Now let $(x', r') \in U \times I$. We have $f(x', a) \in J$ and $f(x', b) \in J$, and since $a < r' < b$ and $f(x', \cdot)$ is monotone, either $f(x', a) \leq f(x', r') \leq f(x', b)$ or $f(x', a) \geq f(x', r') \geq f(x', b)$. In either case, since J is an interval, $f(x', r') \in J$. So $(x, r) \in U \times I \subseteq f^{-1}(J)$. □

Proof of the main theorems, Part 3. We assume $(1)_{n+1}$, as well as $(1)_m$ and $(2)_m$ for all $m \leq n$, and prove $(2)_{n+1}$.

Let $f \colon X \to M$ be a definable function, with $X \subseteq M^{n+1}$, and let $Y = \pi(X) \subseteq M^n$. Then each point in X has the form (y, c) with $y \in Y$ and $c \in X_y$, the fiber of X above y. We say f is **well-behaved** at (y, c) if there is an open box $B \subseteq M^n$ with $y \in B$ and an interval $(a, b) \subseteq M$ with $a < c < b$ such that:

(a) $B \times (a, b) \subseteq X$.

(b) For all $z \in B$, the function $f(z, \cdot) \colon (a, b) \to M$ is continuous and monotone.

(c) The function $f(\cdot, c) \colon B \to M$ is continuous.

Let $W \subseteq X$ be the set of points at which f is well-behaved. Note that W is a definable set. By $(1)_{n+1}$, we can write W and $X \setminus W$ each as a union of cells. Putting these decompositions together, we obtain a cell decomposition of X such that each cell is contained in W or disjoint from W.

If we can decompose each cell into a finite union of definable sets on which f continuous, we are done. So let C be a cell in the decomposition.

Case 1: $\dim(C) = d < n + 1$. Then by Proposition 4.51, there is a definable homeomorphism $g \colon C \to U$, where $U \subseteq M^d$ is an open cell. Applying $(2)_d$ to $h = f \circ g^{-1} \colon U \to M$, we can decompose $U = U_1 \cup \cdots \cup U_k$ with each U_i definable, such that $h|_{U_i}$ is continuous for each i. Let $C_i = g^{-1}(U_i)$ for each i. Then each C_i is definable, $C = C_1 \cup \ldots C_k$, and since g is a homeomorphism, $f|_{C_i} = h|_{U_i} \circ g|_{C_i}$ is continuous.

Case 2: $\dim(C) = n + 1$. By Proposition 4.48, C is an open set. We will show that C is not disjoint from W, from which is follows that $C \subseteq W$.

Claim: f is well-behaved at some point in C.

Since C is open, we can pick some open box $B_0 \times (a, b) \subseteq C$, where $B_0 \subseteq M^n$ is an open box. For each $x \in B_0$, consider the definable function $f(x, \cdot) \colon (a, b) \to M$. Let $\lambda(x)$ be the supremum of those values $c \in (a, b)$ such that $f(x, \cdot)$ is continuous and monotone on (a, c). Then λ is a definable function $B_0 \to M$, and by the monotonicity theorem, $\lambda(x) \in (a, b]$ for all $x \in B_0$.

By $(2)_n$, we can decompose B_0 into finitely many definable sets on which λ is continuous. By $(1)_n$ and Corollary 4.50, one of these definable sets contains an open box $B_1 \subseteq B_0$. Pick some $d \in (a, b)$ such that λ takes on a value greater than d at some point in B_1. Then since λ is continuous on B_1, $\lambda^{-1}(d, \infty)$ is non-empty and open in B_1, and we can find a smaller open box $B_2 \subseteq B_1$ such that $\lambda(x) > d$ for all $x \in B_2$. Note that this ensures that for all $z \in B_2$, the function $f(z, \cdot)\colon (a, d) \to M$ is continuous and monotone.

Fix some $c \in (a, d)$, and consider the function $f(\cdot, c)\colon B_2 \to M$. By $(2)_n$, we can decompose B_2 into finitely many definable sets on which $f(\cdot, c)$ is continuous. By $(1)_n$ and Corollary 4.50, one of these definable sets contains an open box $B_3 \subseteq B_2$ on which $f(\cdot, c)$ is continuous. Picking any $y \in B_3$, the box $B_3 \times (a, d)$ witnesses that f is well-behaved at the point (y, c).

Having established that C is contained in W, we have that f is well-behaved at every point of C. We show that for every point $p \in C$, f is continuous on an open box around p. Then C admits an open cover on which f is continuous, so f is continuous on all of C.

Let $p = (y, c)$, and let $p \in B \times (a, b)$ be a box witnessing that f is well-behaved at p. We apply Lemma 4.52 to the space B, the orders (a, b) and M, and the function $f|_{B\times(a,b)}$. By (b), for all $z \in B$, the function $f(z, \cdot)\colon (a, b) \to M$ is continuous and monotone, which establishes (i).

Point (c) is weaker: it only gives us that $f(\cdot, c)\colon B \to M$ is continuous, while for (ii), we need that for all $d \in (a, b)$, the function $f(\cdot, d)\colon B \to M$ is continuous. But for any $d \in (a, b)$ and any $z \in B$, since f is also well-behaved at (z, d), there is an open box $z \in B'$, which by shrinking we may assume is contained in B, such that $f(\cdot, d)\colon B' \to M$ is continuous. Then B admits an open cover by such boxes on which $f(\cdot, d)$ is continuous, so $f(\cdot, d)$ is continuous on all of B, which establishes (ii) and completes the proof. □

4.5 Dimension

As an application of cell decomposition, we can extend the notion of dimension from cells to arbitrary definable sets. In this section, M is always an o-minimal structure.

Definition 4.53. Let $X \subseteq M^n$ be a definable set. The **dimension** of X, $\dim(X)$, is the maximal d such that there is a cell $C \subseteq X$ with $\dim(C) = d$. We define $\dim(\emptyset) = -\infty$.

It may be somewhat troubling that this notion of dimension relies on the notion of cell. The definition of cell privileges the coordinate directions, and their ordering, while dimension should be an invariant geometric notion. For example, letting $\tau\colon M^2 \to M^2$ be defined by $\tau(x, y) = (y, x)$, the image of a cell under τ is not necessarily a cell, so one may worry that X and $\tau(X)$ may have different dimensions.

We will deal with these concerns by providing an alternative "invariant" characterization of the dimension, which has the following stronger consequence:

if $f\colon X \to Y$ is a definable bijection (not necessarily continuous!) then $\dim(X) = \dim(Y)$. We say the dimension is a **definable invariant** of definable sets.

Since we are overloading the notation dim, we should verify that if X is a cell, the $\dim(X)$ according to Definition 4.53 agrees with $\dim(X)$ according to Definition 4.47. In the following proposition, the dimension of a cell is the number of components on which is it wide.

Proposition 4.54. *Let X be a cell of dimension d. Then d is the maximal integer such that there exists a cell $Y \subseteq X$ with $\dim(Y) = d$.*

Proof. Certainly X contains a cell of dimension d (itself). And if $Y \subseteq X$ is a cell and X is thin in component j, then Y is also thin in component j. So $\dim(Y) \leq d$. □

Note that since M^n is a cell which is wide in each component, we have the desirable property that $\dim(M^n) = n$.

The next result is a companion to Corollary 4.50. Together, they are crucial to the main results on dimension.

Lemma 4.55. *Let $X \subseteq M^n$ be a definable set with non-empty interior and $f\colon X \to M^n$ an injective definable function. Then $f(X)$ has non-empty interior.*

Proof. By induction on n. If $n = 0$, then X is a singleton, and the only function $M^0 \to M^0$ is the identity, so $f(X)$ is a singleton, which is open in M^0.

The inductive argument handles the case $n = 1$, but it is easy to see directly: If $X \subseteq M$ has non-empty interior, then X is infinite. Since f is injective, $f(X) \subseteq M$ is infinite, so by o-minimality $f(X)$ contains an interval.

Now suppose $X \subseteq M^{n+1}$. By cell decomposition, $f(X) = C_1 \cup \cdots \cup C_m$ is a union of cells, and $X = f^{-1}(C_1) \cup \cdots \cup f^{-1}(C_m)$. By Corollary 4.50, there is some i such that $f^{-1}(C_i)$ has non-empty interior. By piecewise continuity, we can further decompose $f^{-1}(C_i)$ into finitely many definable sets $Y_1 \cup \cdots \cup Y_k$ such that $f|_{Y_j}\colon Y_j \to C_i$ is continuous. By Corollary 4.50, there is some j such that Y_j has non-empty interior.

Let $B \times (a, b) \subseteq Y_j$ be an open box, where $B \subseteq M^n$ is an open box. If $\dim(C_i) = n + 1$, then C_i is open by Proposition 4.48, so $f(X)$ has non-empty interior and we are done.

So assume for contradiction that $\dim(C_i) \leq n$. By projecting out one of the coordinates on which C_i is thin, we obtain a definable homeomorphism $h\colon C_i \to C$, where $C \subseteq M^n$ is a cell. Let $g\colon B \times (a, b) \to C$ be the continuous injective definable function obtained by composing f with h.

Now fix some $c \in (a, b)$. The function $g^* = g(\cdot, c)\colon B \to M^n$ is definable and injective, so by induction its image has non-empty interior. Let $B' \subseteq g^*(B)$ be an open box, and let $y \in B'$ and $x \in B$ such that $g^*(x) = g(x, c) = y$.

Now g is continuous, so $g^{-1}(B')$ is an open neighborhood of (x, c) in M^{n+1}. So taking $c' \neq c$ sufficiently close to c, $g(x, c') \in B'$. But then there exists $x' \in B$ such that $g(x', c) = g^*(x') = g(x, c')$. This contradicts injectivity of g, since $c \neq c'$. □

Theorem 4.56. *Let $X \subseteq M^n$ be a definable set with $\dim(X) = d$. Then d is the maximal integer such that there exists a non-empty open definable set $U \subseteq M^d$ and a definable injective function $f\colon U \to X$.*

Proof. Since $\dim(X) = d$, there is a cell $C \subseteq X$ with $\dim(C) = d$. By Proposition 4.51, there is a definable homeomorphism $h\colon C \to U \subseteq M^d$, where U is an open cell. Then $h^{-1}\colon U \to X$ is a definable injective function.

It remains to show that d is maximal with this property. That is, if $U \subseteq M^m$ is an open definable set and $f\colon U \to X$ is a definable injective function, then $m \leq d$.

Let $X = C_1 \cup \cdots \cup C_k$ be a cell decomposition of X. We have $\dim(C_i) \leq d$ for all i. Now $U = \bigcup_{i=1}^{k} f^{-1}(C_i)$, and by Corollary 4.50, there is some i such that $f^{-1}(C_i)$ has non-empty interior. Let $d' = \dim(C_i)$. Then $d' \leq d$, so it suffices to show $m \leq d'$.

Suppose for contradiction that $d' < m$. By Proposition 4.51, there is a definable homeomorphism $h\colon C_i \to U \subseteq M^{d'}$, where U is an open cell. Let $\overline{a} \in M^{m-d'}$ be arbitrary, and let $g\colon U \to M^m$ be given by $\overline{u} \mapsto \overline{ua}$. Then $V = g(U)$ is a cell of dimension d', formed from U by taking the graph of a constant function with value a_i at each stage. Then $\dim(V) = d' < m$, so V has empty interior, but $g \circ h \circ f$ is an injective definable function $f^{-1}(C_i) \to V$, contradicting Lemma 4.55. □

Corollary 4.57. *Let $X \subseteq M^n$ and $Y \subseteq M^m$ be definable sets and $f\colon X \to Y$ a definable bijection. Then $\dim(X) = \dim(Y)$.*

Proof. For any d, if there exists an open definable set $U \subseteq M^d$ and a definable injective function $g\colon U \to X$, then $f \circ g\colon U \to Y$ is a definable injective function. So by Theorem 4.56, $\dim(X) \leq \dim(Y)$. Replacing f by $f^{-1}\colon Y \to X$ shows $\dim(Y) \leq \dim(X)$, by the same argument. □

In addition to invariance under definable bijection, the o-minimal dimension satisfies the following reasonable properties.

Theorem 4.58. *Let $X, Y \subseteq M^n$ be definable sets.*

(*a*) $\dim(X) \leq n$.

(*b*) *If $Y \subseteq X$, then* $\dim(Y) \leq \dim(X)$.

(*c*) $\dim(X \cup Y) = \max(\dim(X), \dim(Y))$.

Proof. (*a*) If $C \subseteq X \subseteq M^n$ is a cell of type s, then by definition $\dim(C) \leq n$, since $\dim(C) = \sum_{i=1}^{n} s(i)$, and each $s(i)$ is 0 or 1. So $\dim(X) \leq n$.

(*b*) Any cell contained in Y is contained in X, so $\dim(Y) \leq \dim(X)$.

(*c*) By (*b*), $\dim(X) \leq \dim(X \cup Y)$ and $\dim(Y) \leq \dim(X \cup Y)$, so

$$\max(\dim(X), \dim(Y)) \leq \dim(X \cup Y).$$

For the converse, suppose $\dim(X \cup Y) = d$. By Theorem 4.56, there is an open definable set $U \subseteq M^d$ and an injective definable function $f\colon U \to X \cup Y$. Then $U = f^{-1}(X) \cup f^{-1}(Y)$, so by Corollary 4.50, either $f^{-1}(X)$ or $f^{-1}(Y)$ has non-empty interior. Assume without loss of generality that there is an open box $B \subseteq f^{-1}(X)$. Then $f|_B\colon B \to X$ is a definable injective function, so $d \leq \dim(X) \leq \max(\dim(X), \dim(Y))$ by Theorem 4.56. □

Corollary 4.59. *Let X be a definable set, and let $X = \bigcup_{i=1}^k C_i$ be any cell decomposition. Then* $\dim(X) = \max_{1 \leq i \leq k} \dim(C_i)$.

Proof. Immediate from Theorem 4.58(c). □

Next, we want to understand fiberwise properties of dimension. Recall that for a set $X \subseteq M^{m+n}$ and a point $\overline{a} \in M^m$, we define

$$X_{\overline{a}} = \{\overline{b} \in M^n \mid \overline{a}\overline{b} \in X\}.$$

Exercise 22. Let $C \subseteq M^{n+m}$ be a cell of type $s_1 \dots s_{m+n}$, and let $\pi\colon M^{m+n} \to M^m$ be the projection on the first m coordinates. Show that $\pi(C)$ is a cell of type $s_1 \dots s_m$ and for all $\overline{a} \in \pi(C)$, $C_{\overline{a}}$ is a cell of type $s_{m+1} \dots s_n$. Conclude that $\dim(C) = \dim(\pi(C)) + \dim(C_{\overline{a}})$ for all $\overline{a} \in \pi(C)$.

Proposition 4.60. *Let $X \subseteq M^{m+n}$ be a definable set, and let $\pi\colon M^{m+n} \to M^m$ be the projection on the first m coordinates. For $d \in \mathbb{N}$, let*

$$X(d) = \{\overline{a} \in \pi(X) \mid \dim(X_{\overline{a}}) = d\}.$$

Then $X(d)$ is definable and $\dim(X \cap \pi^{-1}(X(d))) = \dim(X(d)) + d$.

Proof. Let $\mathcal{C}$ be a cell decomposition of X, so $X = \bigcup_{Y \in \mathcal{C}} Y$. Define $\pi(\mathcal{C}) = \{\pi(Y) \mid Y \in \mathcal{C}\}$. By Lemma 4.41, we can find a cell decomposition $\mathcal{C}'$ of $\pi(X)$ such that each cell in $\pi(\mathcal{C})$ is a disjoint union of cells in $\mathcal{C}'$. Then for each $Z \in \mathcal{C}'$ and each cell $Y \in \mathcal{C}$ such that $Z \subseteq \pi(Y)$, the set $Y \cap \pi^{-1}(Z)$ is also a cell. So by further decomposing the cells in $\mathcal{C}$, we may assume that for each $Y \in \mathcal{C}$, $\pi(Y) \in \mathcal{C}'$.

For all $Z \in \mathcal{C}'$, define $\mathcal{C}_Z = \{Y \in \mathcal{C} \mid \pi(Y) = Z\}$. Then

$$X = \bigcup_{Z \in \mathcal{C}'} \bigcup_{Y \in \mathcal{C}_Z} Y.$$

Now for any $Z \in \mathcal{C}'$ and $\overline{a} \in Z$, we have $X_{\overline{a}} = \bigcup_{Y \in \mathcal{C}_Z} Y_{\overline{a}}$. And by Exercise 22, $\dim(Y) = \dim(Z) + \dim(Y_{\overline{a}})$. So by Theorem 4.58(c),

$$\begin{aligned} \dim(X_{\overline{a}}) &= \max_{Y \in \mathcal{C}_Z} \dim(Y_{\overline{a}}) \\ &= \max_{Y \in \mathcal{C}_Z} (\dim(Y) - \dim(Z)). \end{aligned}$$

This shows that the dimension of $X_{\overline{a}}$ is constant for all $\overline{a} \in Z$. It follows that for all d, $X(d)$ is a union of cells in $\mathcal{C}_Z$, and hence it is definable.

Also, for all $\overline{a} \in Z \subseteq X(d)$, we have

$$\max_{Y \in \mathcal{C}_Z} \dim(Y) = \dim(Z) + \dim(X_{\overline{a}}) = \dim(Z) + d.$$

Now writing $X(d) = Z_1 \cup \cdots \cup Z_k$, we have

$$\begin{aligned} X \cap \pi^{-1}(X(d)) &= \bigcup_{i=1}^{k} (X \cap \pi^{-1}(Z_i)) \\ &= \bigcup_{i=1}^{k} \bigcup_{Y \in \mathcal{C}_{Z_i}} Y. \end{aligned}$$

By Theorem 4.58(c),

$$\begin{aligned} \dim(X \cap \pi^{-1}(X(d))) &= \max_{1 \le i \le k} \left(\max_{Y \in \mathcal{C}_{Z_i}} \dim(Y) \right) \\ &= \max_{1 \le i \le k} (\dim(Z_i) + d) \\ &= \dim(X(d)) + d. \end{aligned}$$

□

Theorem 4.61. *(i) Let $X \subseteq M^{m+n}$ be a definable set, and let $\pi \colon M^{m+n} \to M^m$ be the projection on the first m coordinates. Then*

$$\dim(X) = \max_{0 \le d \le n} (\dim(X(d)) + d) \ge \dim(\pi(X)).$$

(ii) Let $X \subseteq M^m$ and $Y \subseteq M^n$ be definable. Then then $X \times Y \subseteq M^{m+n}$ is definable and

$$\dim(X \times Y) = \dim(X) + \dim(Y).$$

Proof. For (i), note that $\pi(X) = \bigcup_{i=0}^{n} X(d)$. This is because for each $\overline{a} \in \pi(X)$, $X_{\overline{a}} \subseteq M^n$, so $0 \le \dim(X_{\overline{a}}) \le n$. It follows that

$$X = \bigcup_{i=0}^{n} (X \cap \pi^{-1}(X(d))),$$

so we have the dimension computation by Proposition 4.60 and Theorem 4.58(c).

For the inequality,

$$\dim(\pi(X)) = \max_{0 \le d \le n} \dim(X(d)) \le \max_{0 \le d \le n} (\dim(X(d)) + d).$$

For (ii), if $\varphi(\overline{x})$ is the formula defining X and $\psi(\overline{y})$ is the formula defining Y, and the variables $\overline{x}$ and $\overline{y}$ are all distinct, then $\varphi(\overline{x}) \wedge \psi(\overline{y})$ defines $X \times Y$.

Now $\pi(X \times Y) = X$, and for all $\overline{a} \in X$, $(X \times Y)_{\overline{a}} = Y$. So if $d = \dim(Y)$, then $(X \times Y)(d) = X$ and $(X \times Y)(k) = \emptyset$ for all $k \neq d$. Thus, by (i), $\dim(X \times Y) = \dim((X \times Y)(d)) + d = \dim(X) + \dim(Y)$. □

We can leverage the previous results to understand the dimensional relationships between a definable set, its image, and its fibers, under any definable function, not just projections.

Exercise 23. Let $X \subseteq M^n$ be a definable set and $f\colon X \to M^m$ a definable function. For $d \in \mathbb{N}$, let $X_f(d) = \{\overline{a} \in f(X) \mid \dim(f^{-1}(\{\overline{a}\})) = d\}$.

(1) Show that $X_f(d)$ is definable and $\dim(f^{-1}(X_f(d))) = \dim(X_f(d)) + d$.

(2) Show that

$$\dim(X) = \max_{0 \leq d \leq n} (\dim(X_f(d)) + d) \geq \dim(f(X)).$$

Hint: Apply Theorem 4.61 to the definable set $\{(f(\overline{x}), \overline{x}) \mid \overline{x} \in X\} \subseteq M^{m+n}$, and note that this set is in definable bijection with X.

As a consequence of the result $\dim(X) \geq \dim(f(X))$ for all definable functions f, we conclude, for example, that there is no "space-filling definable curve", i.e., no definable surjective function $f\colon M \to M^2$.

Exercise 24. Suppose $f\colon X \to M^m$ is a definable finite-to-one function, i.e. for all $y \in M^m$, $f^{-1}(\{y\})$ is finite. Then $\dim(X) = \dim(f(X))$.

Exercise 25. Prove the following characterization of dimension, which is dual to Theorem 4.56. Let $X \subseteq M^n$ be a non-empty definable set with $\dim(X) = d$. Then d is the maximal integer such that there exists a definable function $f\colon X \to M^d$ where $f(X)$ has non-empty interior.

5 Further directions

I will end with a brief selection of further work around o-minimality which we did not have time to cover this semester.

Euler characteristic: In addition to dimension, another important invariant of definable sets in an o-minimal structure is the Euler characteristic χ.

For a cell C, we define

$$\chi(C) = (-1)^{\dim(C)}.$$

And for an arbitrary definable set X with disjoint cell decomposition $X = \bigcup_{C \in \mathcal{D}} C$, we define

$$\begin{aligned}\chi(X) &= \sum_{C \in \mathcal{D}} \chi(C) \\ &= k_0 - k_1 + k_2 - \cdots + (-1)^{\dim(X)} k_{\dim(X)}\end{aligned}$$

where k_d is the number of cells in $\mathcal{D}$ of dimension d. Now one can prove that the Euler characteristic $\chi(X)$ does not depend on the choice of cell decomposition of X, and that if $f\colon X \to Y$ is a definable bijection, then $\chi(X) = \chi(Y)$.

It turns out that in the context of an o-minimal field, the dimension and Euler characteristic provide a complete set of definable invariants for definable sets. That is, if $X \subseteq M^n$ and $Y \subseteq M^m$ are definable, $\dim(X) = \dim(Y)$, and $\chi(X) = \chi(Y)$, then there is a definable bijection $f\colon X \to Y$. The proof of this theorem is a starting point for the development of some basic algebraic topology for definable sets in o-minimal fields.

Reference: Tame topology and o-minimal structures by van den Dries.

o-minimal complex analysis: We have seen that every o-minimal field is real closed (so elementarily equivalent to $\mathbb{R}$), any for any real closed field R, we have $C = R[\sqrt{-1}]$ is algebraically closed (elementarily equivalent to $\mathbb{C}$). Identifying C with R^2, the field operations on C are definable in the field R.

Peterzil and Starchenko have developed a theory of o-minimal complex analysis for sets and functions in C which are definable in an o-minimal expansion of R. The usual tools of power series and integration are not available in this setting and have to be replaced by topological methods. But o-minimality allows them to obtain tameness and uniform finiteness results that strengthen classical results in complex analysis. In the other direction, they are able to show that certain important objects for complex analysis and algebraic geometry can be defined in o-minimal expansions of the reals.

Reference: A good starting place is the paper *Expansions of algebraically closed fields in o-minimal structures* by Peterzil and Starchenko.

Point counting and Diophantine applications: Many of the most celebrated applications of o-minimality flow from a theorem of Pila and Wilkie. Suppose $X \subseteq \mathbb{R}^n$ is definable in an o-minimal structure on $\mathbb{R}$, and let X^{trans} be the subset obtained from X by removing all connected semi-algebraic subsets of dimension at least 1 (this is the "transcendental" part of X —recall that "semi-algebraic" means definable in the ordered field language, and semialgebraic sets can have many rational points). Let $N(X^{\text{trans}}, t)$ be the number of points in $\mathbb{Q}^n \cap X^{\text{trans}}$ such that the numerator and denominator of each coordinate is bounded above by t when written in lowest terms. Then $N(X^{\text{trans}}, t)$ grows sub-polynomially in t: is it $O(t^\varepsilon)$ for all $\varepsilon > 0$.

For applications of this theorem to arithmetic geometry, the strategy is to show that some object of interest (like an elliptic curve) admits a transcendental covering map from a fundamental domain in $\mathbb{R}^n$, such that the fundamental domain and covering map are definable in an o-minimal expansion of $\mathbb{R}$, and such that that rational points are mapped to points of special interest on the object (like torsion points).

Reference: For a survey of this area, see *Counting special points: logic, Diophantine geometry and transcendence theory* by Scanlon.

Applications outside pure mathematics: In many areas of mathematical modeling, it is sufficient to consider objects (functions, regions in $\mathbb{R}^n$, probability distributions, etc.) which are definable in o-minimal structures on $\mathbb{R}$ (and often even definable in the real field). In such cases, the application can be carried out without worrying about pathologies from real analysis cropping up. And sometimes, techniques from o-minimality can strengthen results, e.g. via uniformity.

As an explicit example, in PAC learning theory, a "concept class" is a family of subsets of a given set. For most practical purposes, concept classes can be taken to be definable families of definable sets in an o-minimal structure on $\mathbb{R}$. As a consequence of cell-decomposition, every such concept class has finite VC-dimension, which implies that it is PAC learnable.

Reference: For an example application to PAC learning, see *On sample complexity in neural networks*, by Usvyatsov.

Variants of o-minimality: It is a common theme in model theory that understanding definable subsets of M^1 can help us understand definable subsets of M^n for all n. Strong minimality and o-minimality are the two most well-known examples, but model theorists study other kinds of "X-minimality".

One kind of generalization is to continue working in a linearly ordered setting, but weaken the hypotheses of o-minimality. For example, a weakly o-minimal theory is one in which every definable subset of M^1 is a finite union of convex sets (points and intervals whose endpoints need not be in the structure).

Another kind of generalization is to move to non-ordered structures, but require that every definable subset of M^1 is already definable in some smaller language (the empty language gives strong minimality, and the language $\{\leq\}$ gives o-minimality). Examples include C-minimality (which captures the definability behavior in algebraically closed valued fields) and P-minimality (which does the same for the p-adic field $\mathbb{Q}_p$).

References: Section 4 of *Notes on o-Minimality and Variations* by Macpherson, and the references therein.

NIP and distal theories: Finally, we turn to generalizations in "pure model theory". The o-minimal theories sit inside the class of NIP (not independence property) theories. These are exactly the theories in which definable families of definable sets have finite VC-dimension, as mentioned above. The NIP theories include both o-minimal theories and stable theories (roughly, those in which no infinite sequence can be definably ordered) like ACF.

Every o-minimal theory is distal (roughly, "purely unstable NIP"). Distal theories admit a kind of abstract cell decomposition, and recent work demonstrates that many combinatorial results about definable sets in the real field go through in the wider distal setting.

References: For NIP, Simon's book *A guide to NIP theories*. For an example combinatorial application of distality, see *Regularity lemma for distal structures* by Chernikov and Starchenko