Definable Sets in Mathematics

Coven–Wood Lecture I

David Marker

Mathematics, Statistics, and Computer Science
University of Illinois at Chicago

November 1, 2023

Ethan Coven & Carol Wood

Logic is the beginning of wisdom not the end

Main Goal Use tools from mathematical logic (particularly formal langauges) to better understand classical mathematical structures.

Exploit the interplay of semantics and syntax

Semantics = truth in mathematics structures
Syntax = formal expressions in symbolic first order logic

Mathematical Structures

Sets with distinguished functions, relations and elements we want to study

- algebraic structures $\mathcal{L} = \{+, \cdot, 0, 1\}$.
  - $\mathbb{N}$;
  - (rings) $\mathbb{Z}$;
  - (fields) $\mathbb{Q}, \mathbb{R}, \mathbb{C}$.
- structures with a binary relation $\mathcal{L} = \{R\}$.
  - an equivalence relation;
  - a graph ($R(x, y)$ if there is an edge between x and y);
  - an ordering.
- ordered algebraic structures $(\mathbb{Q}, +, <, 0)$, $(\mathbb{R}, +, \cdot, <, 0, 1)$;
- fields with exponentiation $\mathbb{R}_{\mathrm{exp}} = (\mathbb{R}, +, \cdot, \exp, 0, 1)$, $\mathbb{C}_{\mathrm{exp}} = (\mathbb{C}, +, \cdot, \exp, 0, 1)$.

Atomic Formulas

Fix a language $\mathcal{L}$–for example $\mathcal{L}_{\text{exp}} = \{+, \cdot, -, \exp, <, 0, 1\}$
• Build simple formulas using symbols of $\mathcal{L}_{\text{exp}}$, variables $x, y, z, x_1, x_2, \ldots$ and parenthesis (and)

For example

- $0 + 1 = 1$
- $(1 + 1) \cdot (1 + 1 + 1) = (1 + 1 + 1 + 1 + 1 + 1)$ $\qquad 2 \cdot 3 = 6$
- $y \cdot y = x$ $\qquad y^2 = x$
- $x \cdot x + y \cdot y = 1$ $\qquad x^2 + y^2 = 1$
- $\exp(x + y) = \exp(x) \exp(y)$.

Formulas

• We build up more complicated formulas using Boolean connectives: $\wedge$ ("and"), $\vee$ "or", $\neg$ "not" $\rightarrow$ "implies"

For example

- $x + y = z \wedge x \cdot x + (1 + 1) \cdot y = 0$
- $x < y \rightarrow x + z < y + z$
- $\neg(x \cdot y = 0) \rightarrow \neg(x = 0)$ If $xy \neq 0$, then $(x \neq 0 \wedge y \neq 0)$

• Quantifiers $\exists$ (there exists) and $\forall$ (for all)

for example:

- $\exists x \; x \cdot x + x + 1 = 0$
- $\exists y \; y \cdot y = x$ x is a square
- $\forall x \exists y \; y \cdot y = x$ every element is a square
- $\forall \epsilon > 0 \; \exists \delta > 0 \; \forall x \; (|x - a| < \delta \rightarrow |\exp(x) - b| < \epsilon)$

$$\lim_{x \rightarrow a} \exp(x) = b$$

Formulas v. Sentences

An important technical point: A *sentence* is a formula where all of the variables are bound in the scope of a quantifier.

Sentences:

$$\forall x \exists y \; y^2 = x$$

$$\exists x \; x^2 = 1 + 1$$

Non Sentences:

$$\exists y \; y^2 = x$$

$$\exists x \; x^2 + y \cdot x + z = 0$$

Theories

Sentences are declarative statements. In any particular structure they are either true or false.

- $\exists x \forall y \; x \cdot y = y$
  - True in $\mathbb{N}$, $\mathbb{Z}$,$\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ (take $x = 1$).
- $\forall x \exists y \; x \cdot y = 1$
  - False in $\mathbb{N}$, $\mathbb{Z}$ (take $x = 2$)
  - True in $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$.
- $\forall x \exists y \; y^2 = x$
  - False in $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ (no $\sqrt{2}$)
  - False in $\mathbb{R}$ (no $\sqrt{-1}$)
  - True in $\mathbb{C}$

The *Theory* of a structure $\mathcal{M}$ is the set of all sentences true in $\mathcal{M}$ and denoted $\mathrm{Th}(\mathcal{M})$.

Definable Sets

Formulas with free variable assert a property of the free variables.
$\exists y\ y^2 = x$ asserts "x is a square"

- in $\mathbb{Z}$ or $\mathbb{Q}$ it is true for $x = 9$, but false for $x = 3$
- in $\mathbb{R}$ it is true of any $x \geq 0$ but false for $x = -3$
- in $\mathbb{C}$ it is true for every x.

Suppose $\phi(x_1, \ldots, x_n)$ is a formula with free variables $x_1, \ldots, x_n$ and $\mathcal{M}$ is a structure. We say that

$$\{(a_1, \ldots, a_n) : \phi \text{ holds in } \mathcal{M} \text{ of } a_1, \ldots, a_n\}$$

is *definable*.

We also allow parameters.

For example, $(0, \pi)$ is definable in $\mathbb{R}$.

Examples of Definable Sets in $(\mathbb{R}, +, \cdot, 0, 1$

Some definable sets in $\mathbb{R}^2$.

- $\{(x, y) : x < y\}$ is defined by

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

- the closed unit disk defined by $\exists z\ x^2 + y^2 + z^2 = 1$
- $\{(x, y) : 0 < x < 1 \wedge y^2 > y^3\}$

Lemma

Suppose $A \subset \mathbb{R}^2$ is definable, then $\overline{A}$ the closure of A is definable

Let $\phi(x, y)$ define A. Then $\overline{A}$ is defined by
$\forall \epsilon > 0 \exists x_0 \exists y_0\ (\phi(x_0, y_0) \wedge (x - x_0)^2 + (y - y_0)^2 < \epsilon)$.

More definable sets

- $\mathbb{N}$ is definable in $(\mathbb{Z}, +, \cdot)$
  (Lagrange) $x \geq 0 \Leftrightarrow \exists y_1 \exists y_2 \exists y_3 \exists y_4 \; x = y_1^2 + y_2^2 + y_3^2 + y_4^2$
- $\mathbb{Z}_p$ is definable in $\mathbb{Q}_p$
  for $p \neq 2$ use Hensel's Lemma to show $\mathbb{Z}_p = \{x : \exists y \; y^2 = px^2 + 1\}$
- (J. Robinson 1950s) $\mathbb{Z}$ is definable in $(\mathbb{Q}, +, \cdot)$.
  Recently Koenigsmann $\mathbb{Z}$ can be defined by a formula $\forall x_1 \ldots \forall x_m \psi$ where ψ has no quantifiers

An undefinability result

Proposition

$\mathbb{R}$ *is not definable in* $\mathbb{C}$*.*

Suppose $\mathbb{R}$ is defined by $\phi(x, \overline{a})$.
Let k be the field generated by $\overline{a}$.

Fact: If σ is an automorphism of $\mathbb{C}$ that fixes k then $\phi(x, \overline{a}) \Leftrightarrow \phi(\sigma(x), \overline{a})$.

Let $x \in \mathbb{R}$ and $y \in \mathbb{C} \setminus \mathbb{R}$ be transcendental over k. Then there is an automorphism σ of $\mathbb{C}$ such that $\sigma(x) = y$.
But $\phi(x, \overline{a})$ and $\neg\phi(y, \overline{a})$, a contradiction.

Our Main Goals Restated

Let $\mathcal{M}$ be one of our classical mathematical structures.

- Try to understand $\mathrm{Th}(\mathcal{M})$, the complete theory of $\mathcal{M}$.
- Try to understand the definable subsets of $\mathcal{M}^n$.

A bad example-Hilbert's Program

Understand $\mathrm{Th}(\mathbb{N})$.

- (Axiomatization Problem) Can we give a simple set of axioms T true about $\mathbb{N}$ such that all true statements can be derived from T by simple logical rules?
- (Decidability Problem) Is there an algorithm which when given a sentence ϕ as input will decide if ϕ is true in $\mathbb{N}$?

Good candidate for axiomatization: *Peano Axioms*

- Basic properties of $+$ and $\cdot$ like $\forall x \forall y \; x(y+1) = xy + x$
- Induction axioms

$$[\phi(0) \land \forall x \; (\phi(x) \to \phi(x+1)] \to \forall x \; \phi(x)$$

Gödel's Incompleteness Theorem

In 1931 Kurt Gödel left Hilbert's Program in ruins.

Theorem (Gödel)

i) There are true sentences about the natural numbers that can not be derived from the Peano axioms.
ii) The same is true for any other possible simple set of axioms
iii) There is no algorithm which when input a sentence ϕ will halt and tell you if ϕ is true in $\mathbb{N}$.

Because we can define $\mathbb{N}$ in $\mathbb{Z}$ and $\mathbb{Q}$, $\mathrm{Th}(\mathbb{Z})$ and $\mathrm{Th}(\mathbb{Q})$ are also undecidable.

It gets worse-Hilbert's Tenth Problem

Let $P_0, P_1, P_2, \ldots$ list all computer programs in your favorite language.

Theorem (Matiyasevich-J. Robinson-Davis-Putnam 1949-70)

There is a polynomial $p(X, Y, Z_1, \ldots, Z_n) \in \mathbb{Z}[X, Y, \overline{Z}]$ such that P_e halts on input x iff and only if

$$\exists z_1 \in \mathbb{Z} \ldots \exists z_n \in \mathbb{Z} \ p(e, x, \overline{z}) = 0.$$

Solving Diophantine equations is as hard as deciding if a computer program halts, which was shown undecidable by Turing.
So there is no algorithm which can decide if a polynomial over the integers has an integer zero.

Open Question: Is the same true for $\mathbb{Q}$?

Key Lesson: *Quantifiers lead to complexity.*

A Good Example–Tarski

Consider the ordered real field $(\mathbb{R}, +, \cdot, <)$

Theorem (Tarksi–Quantifier Elimination)

Every formula is equivalent to a formula without quantifiers and there is an algorithm that converts every formula to an equivalent quantifier free formula.

Note: $<$ is necessary as otherwise not eliminate the quantifier from

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

Familiar examples of quantifier elimination:

$$\exists x\ x^2 + bx + c = 0 \ \Leftrightarrow b^2 - 4c \geq 0.$$

$$\exists x \exists y \exists u \exists v\ (ax + bu = 1 \wedge ay + bv = 0 \wedge cx + du = 0 \wedge cy + dv = 1) \Leftrightarrow$$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \text{ is invertible } \Leftrightarrow ad - bc \neq 0.$$

Decidability and Axiomatizbility

Corollary

$\mathrm{Th}(\mathbb{R})$ *is decidable*

To decide if a sentence ϕ is true, convert it to a quantifier free sentence ψ. It is easy to check if ψ is true. $(1+1)\cdot(1+1+1) > (1+1+1+1+1)$

Tarski also showed that $\mathrm{Th}(\mathbb{R})$ can be axiomatized by:
• axioms for ordered fields
• saying that if p is a polynomial, $a < b$ and $p(a) < 0 < p(b)$, then there is $a < c < b$ such that $p(c) = 0$.

Tarski also showed that $\mathbb{C}$ has quantifier elimination and can be axiomatized by saying its an algebraically closed field of characteristic 0

Tarski's Problem–exponentiation?

Open Problem Suppose we consider the structure $\mathbb{R}_{\exp} = (\mathbb{R}, +, \cdot, \exp)$, where $\exp(x) = e^x$. Is $\mathrm{Th}(\mathbb{R}_{\exp})$ decidable?

A positive answer would show the decidability of hyperbolic geometry.

Even deciding equality of terms is difficult. Is

$$e^e = 9e^3 - 6e^2 - 121? \text{ Probably not}$$

A New Paradigm

Decidability is the wrong problem.
Even the theories we know are decidable are provably intractable.
Our goal should be understanding definable sets.

Semialgebraic Sets

Definition

We say that $X \subseteq \mathbb{R}^n$ is *semialgebraic* if it is a finite Boolean combination of sets of the form

$$\{x \in \mathbb{R}^n : p(x) = 0\} \text{ and } \{x \in \mathbb{R}^n : q(x) > 0\}$$

where $p, q \in \mathbb{R}[X_1, \ldots, X_n]$.

Semialgebraic $\Leftrightarrow$ Quantifier-free definable $\Leftrightarrow$ Definable

Corollary

If $X \subseteq \mathbb{R}$ is definable, X is a finite union of points and intervals.
In particular, $\mathbb{Z}$ is not definable.

o-minimality

Definition

We say that $(\mathbb{R}, +, \cdot, <, \dots)$ is *o-minimal* if every definable subset of $\mathbb{R}$ is a finite union of points and intervals.

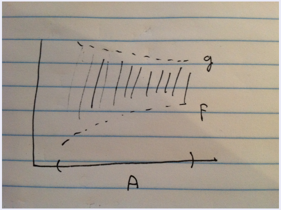
Intuition: In an o-minimal structure the definable subsets of $\mathbb{R}$ are exactly the ones that can be defined using the only the ordering.

Although o-minimality talks about definable subsets of $\mathbb{R}$, it has strong topological and geometric consequences about definable subsets of $\mathbb{R}^n$.

Cells

Definition

- points and intervals in $\mathbb{R}$ are *cells*;
- if $A \subseteq \mathbb{R}^n$ is a cell and $f : A \to \mathbb{R}$ is continuous and definable, then
  $\{(x, f(x)) : x \in A\}$ is a cell;
- if $A \subseteq \mathbb{R}^n$ is a cell and $f, g : A \to \mathbb{R}$ are continuous and definable with $f < g$ on A, then $\{(x, y) : x \in A, f(x) < y < g(x)\}$ is a cell.



Cells are definable homeomorphic to $(0, 1)^n$ for some n.

Cell Decomposition

Charles Steinhorn 

Theorem (van den Dries/Knight-Pillay-Steinhorn)

Suppose $(\mathbb{R}, +, \cdot, <, \ldots)$ *is o-minimal.*

1. *If* $X \subseteq \mathbb{R}^n$ *is definable, then* X *is a finite union of cells. In particular definable sets have finitely many connected components.*
2. *If* $f : X \to \mathbb{R}$ *is definable, we can partition* X *into cells* $X_1 \cup \cdots \cup X_m$ *such that* $f|X_i$ *is continuous (indeed* $\mathcal{C}^r$*).*

Is $\mathbb{R}_{\exp}$ o-minimal?

We knew early on that we do not have quantifier elimination.
Wilkie proved the next best thing.

Theorem (Wilkie)

If $X \subseteq \mathbb{R}^n$ is definable in $\mathbb{R}_{\exp}$, then there is $V \subseteq \mathbb{R}^{n+m}$ an exponential-algebraic variety, such that

$$X = \{x \in \mathbb{R}^n : \exists y\ (x, y) \in V\}.$$

Exponential varieties are quantifier free definable.
Thus every definable set is of the form $\{\overline{x} \in \mathbb{R}^n : \exists \overline{y} \in \mathbb{R}^m\ \phi(\overline{x}, \overline{y})\}$ where ϕ is quantifier free.

Theorem (van den Dries–Macintyre–Marker)

We can eliminate quantifiers if we add ln *and all analytic functions on* $[0, 1]^n$, $n \in \mathbb{N}$.

o-minimality of $\mathbb{R}_{\exp}$

Wilkie: definable = projection of exponential-algebraic variety.

Theorem (Khovanski)

Exponential-algebraic varieties have finitely many connected components.

$X \subset \mathbb{R}$ definable $\Rightarrow$ finitely many connected components.

Corollary (Wilkie)

$\mathbb{R}_{\exp}$ *is o-minimal.*

Macintyre and Wilkie were able to show decidability BUT assuming Schannuel's Conjecture in transcendental number theory.

Open Problem: We can axiomatize $\mathrm{Th}(\mathbb{R}_{\exp})$ using only axioms of the form $\forall \overline{x} \exists \overline{y} \phi$ where ϕ is quantifier free. What are the axioms?

The trouble with $\mathbb{C}_{\text{exp}}$

What can we say about definability in $\mathbb{C}_{\text{exp}} = (\mathbb{C}, +, \cdot, \exp, 0, 1)$?

The first thing you notice:

$$\mathbb{Z} = \{x : \forall y\ (\exp(y) = 1 \rightarrow \exp(xy) = 1)$$

Thus $\mathbb{C}_{\text{exp}}$ is undecidable and all of the Gödel pheonomena arise.

Is this the end of the story?

Open questions about definability in $\mathbb{C}_{\exp}$

- Is $\mathbb{R}$ definable in $\mathbb{C}_{\exp}$?

- (quasiminimality) Is every definable subset of $\mathbb{C}$ countable or co-countable?

- Does $\mathbb{C}_{\exp}$ have nontrivial automorphisms other than $z \mapsto \overline{z}$?

Zilber's Approach

Zilber described a class $\mathcal{K}$ of *pseudoexponential* fields where:

- For each uncountable cardinal κ there is, up to isomorphism, a unique K in $\mathcal{K}$ of cardinality κ;
- Every $K \in \mathcal{K}$ is quasiminimal;
- If $K \in \mathcal{K}$ has size $\kappa > \aleph_0$, then $|Aut(K)| = 2^\kappa$.

Is $\mathbb{C}_{\exp} \in \mathcal{K}$?

Zilber's axioms for $(K, +, \cdot, E) \in \mathcal{K}$

- K is an algebraically closed field of characteristic 0.
- $E : K^+ \to K^\times$ is a surjective homomorphism.
- There is a transcendental η such that the kernel of E is $\mathbb{Z}\eta$.
- (Schanuel's Condition) If $x_1, \ldots, x_n \in K$ are $\mathbb{Q}$-linearly independent, then

$$\mathrm{td}\ \mathbb{Q}(x_1, \ldots, x_n, E(x_1), \ldots, E(x_n)) \geq n.$$

- (Strong exponential closure) For "reasonable" algebraic varieties $V \subset K^{2n}$, there is $x \in K^n$ such that $(x, E(x)) \in V$
- (Countable closure) systems as above have at most countably many "generic" solutions.

Evidence for $\mathbb{C}_{\exp} \in \mathcal{K}$?

• Zilber showed countable closure is true for $\mathbb{C}_{\exp}$.

• Some success has been made showing strong exponential closure in special cases.

Theorem (Marker)

If $p(X, Y) \in \mathbb{C}[X, Y]$ is irreducible and both variables occur, then $p(z, \exp(z)) = 0$ has infinitely many solutions. If Schanuel's Conjecgture is true and $p \in \mathbb{Q}[X, Y]$, then there are infinitely many algebraically independent solutions..

Thank you!