TAME TOPOLOGY AND O-MINIMAL STRUCTURES

Outline of Lectures

An Introduction to Definability

For which x do we have

$$\exists y \begin{bmatrix} ax + by > e_1 \\ cx + dy \leq e_2 \end{bmatrix} \qquad (*)$$

where $a, b, c, d, e_1, e_2 \in \mathbb{R}$. Want

$$\{x \in \mathbb{R} \mid \text{statement } (*) \text{ holds for } x\}.$$

We also can replace the constants a, b, c, d, e_1, e_2 by variables u, v, w, z, s_1, s_2 and ask for which $u, v, w, z, s_1, s_2 \in \mathbb{R}$ do we have

$$\exists x \exists y \begin{bmatrix} ux + vy > s_1 \\ wx + zy \leq s_2 \end{bmatrix}. \qquad (**)$$

That is, we want to know

$$\{(u, v, w, z, s_1, s_2) \in \mathbb{R}^6 \mid (u, v, w, z, s_1, s_2) \text{ satisfies } (**)\}$$

These are simple examples of *definable sets*.

Before getting more precise, some further

Examples

1. $S = \{x \in \mathbb{R} \mid \sin \pi x = 0\}$

   Observe $S = \mathbb{Z}$.

2. Let $L = \{(m, n) \in \mathbb{R}^2 | m, n \in \mathbb{Z}\}$ be the integer lattice in $\mathbb{R}^2$. View L as a two-place relation on pairs of real numbers:

$$L(m, n) \quad \Leftrightarrow \quad m, n \in \mathbb{Z}.$$

   Define

$$C = \Big\{(x, y) \in \mathbb{R}^2 \mid \exists u \exists v \big[L(u, v) \\ \wedge\ (x - u)^2 + (y - v)^2 \leq 1/16\big]\Big\}.$$

   The set C is the set of all closed discs of radius $1/4$ whose centers are points in L.

Structures on the real numbers $\mathbb{R}$

- Fix a stock
  i. $\mathcal{F}$ of *basic functions* $f\colon \mathbb{R}^k \to \mathbb{R}$, for $k = 1, 2, 3, \ldots$
  ii. $\mathcal{R}$ of *basic relations (subsets)* $R \subseteq \mathbb{R}^k$, for $k = 1, 2, 3, \ldots$, such that

  $$\text{the ``less than'' relation } < \subset \mathbb{R}^2$$

  (and equality) always is included in $\mathcal{R}$.

  Refer to $\mathcal{F}$ and $\mathcal{R}$ together as a *language* $\mathcal{L}$.

- The real numbers $\mathbb{R}$ together with the functions and relations included in $\mathcal{L}$ is called an *$\mathcal{L}$-structure* that in general we denote by $\mathfrak{R}_{\mathcal{L}}$.

  Special structures will have special names.

Examples

1. $\mathcal{F} = \{+, \cdot, -\}$ where $-\colon \mathbb{R} \to \mathbb{R}$ is given by $x \mapsto -x$, and $\mathcal{R} = \{<\}$. This is the language $\mathcal{L}_{\mathbf{alg}}$ of the ordered field of real numbers, the structure denoted by $\mathbb{R}_{\mathbf{alg}}$.

2. $\mathcal{F}$ contains $+$ and, for each $r \in \mathbb{R}$, the scalar multiplication function $\mu_r\colon \mathbb{R} \to \mathbb{R}$ given by $x \mapsto rx$, and $\mathcal{R} = \{<\}$. This is the ordered real vector space language $\mathcal{L}_{\mathbf{lin}}$ for the real numbers, whose corresponding structure is denoted by $\mathbb{R}_{\mathbf{lin}}$

3. $\mathcal{F} = \{+, \cdot, -, \exp\}$ where $\exp\colon \mathbb{R} \to \mathbb{R}$ is the exponential function $x \mapsto e^x$ and $\mathcal{R} = \{<\}$. This is the language $\mathcal{L}_{\mathbf{exp}}$ of the ordered exponential field of real numbers; the structure is denoted by $\mathbb{R}_{\mathbf{exp}}$.

Terms

- Construct an expanded class of functions by repeated composition starting from the functions in $\mathcal{F}$. These are called *terms*.

Examples

1. The $\mathcal{L}_{\mathbf{alg}}$-terms are all integer cofficient polynomial functions

$$p(x_1, \dots x_k) \text{ for } k = 1, 2, 3, \dots.$$

2. The $\mathcal{L}_{\mathbf{lin}}$-terms are all $\mathbb{R}$-linear functions

$$L(x_1, \dots, x_k) = \mu_{r_1} x_1 + \cdots + \mu_{r_k} x_k.$$

Formulas

- First the *basic formulas*
  i. $t_1 = t_2$ for terms t_1, t_2.
  ii. $R(t_1, \ldots, t_k)$ where R is a k-place relation and $t_1, \ldots, t_k$ are terms. Note that $t_1 < t_2$ is a special case.

- Recursively, from the basic formulas construct
  i. by boolean operations from formulas φ and ψ:

     $\varphi \wedge \psi$ read as "φ and ψ"
     $\varphi \vee \psi$ read as "φ or ψ"
     $\neg\varphi$ read as " not φ"

  ii. by existential quantification over $\mathbb{R}$ from a formula φ:

     $\exists v\, \varphi$ read as "there exists v such that φ"

Caution To get to the point more quickly I have mixed syntax and semantics.

$\mathcal{L}$-definable sets

For our purposes these are subsets of $\mathbb{R}^k$, for $k = 1, 2, 3, \ldots$, specified as follows.

For each $\mathcal{L}$-formula φ, certain variables are bound to quantifiers and others are not. Call the latter *free variables*, and it is for these that we can substitute real numbers.

For an $\mathcal{L}$-formula φ list its free variables as $x_1, \ldots, x_k, z_1, \ldots, z_m$. Choose $c_1, \ldots, c_m \in \mathbb{R}$ and substitute them for $z_1, \ldots, z_m$, respectively. Write $\bar{c}$ for $(c_1, \ldots, c_m)$. The set

$$D_{\varphi,\bar{c}} = \{(x_1, \ldots, x_k) \in \mathbb{R}^k \mid \varphi(x_1, \ldots, x_k, \bar{c}) \text{ is true}\} \subseteq \mathbb{R}^k$$

is an *$\mathcal{L}$-definable* subset of $\mathbb{R}^k$.

If $\mathcal{L}$ is clear from context, we usually shall drop the $\mathcal{L}$.

Comments

- We call the $\bar{c}$ *parameters*; the division of the free variables in a formula into those which are parameter variables and those that are not can be made in arbitrarily.
- A function $F\colon A \subseteq \mathbb{R}^m \to \mathbb{R}^n$ is said to be $\mathcal{L}$-definable if its graph, as a subset of $\mathbb{R}^m \times \mathbb{R}^n$, is $\mathcal{L}$-definable.
- If a function $F(x_1, \ldots, x_m)$ is definable, we can treat it as if it is a term.
- Similarly, we can treat a definable subset of $\mathbb{R}^k$ as a basic relation in the language.

Some Further Examples

1. Let $f\colon \mathbb{R} \to \mathbb{R}$ be $\mathcal{L}$-definable, where $\mathcal{L}$ contains $\mathcal{L}_{\mathbf{alg}}$. Then

$$\{x \in \mathbb{R} \mid f \text{ is convex in an interval around } x\}$$

is $\mathcal{L}$-definable.

2. Let $f\colon A \subset \mathbb{R}^k \to \mathbb{R}$ be $\mathcal{L}$-definable, where $\mathcal{L}$ contains $\mathcal{L}_{\mathbf{alg}}$. Then $\{\bar{x} \in A \mid f \text{ is continuous at } \bar{x}\}$ is $\mathcal{L}$-definable. Similarly for differentiability.

3. Let $f\colon \mathbb{R} \to \mathbb{R}$ be $\mathcal{L}$-definable, where $\mathcal{L}$ contains $\mathcal{L}_{\mathbf{alg}}$. If f is differentiable, then f' is $\mathcal{L}$-definable. Same for functions of several variables.
4. Let $f\colon \mathbb{R} \to \mathbb{R}$ be $\mathcal{L}$-definable, where $\mathcal{L}$ contains $\mathcal{L}_{\mathbf{alg}}$. Then

$$\{x \in \mathbb{R} \mid f \text{ is Lipschitz in an interval around } x\}$$

is $\mathcal{L}$-definable.

5. Let $A \subset \mathbb{R}^k$ be $\mathbb{R}_{\mathbf{alg}}$-definable. Then there is a formula that expresses that "A is convex." Same for any $\mathcal{L}$ containing $\mathcal{L}_{\mathbf{alg}}$.
6. Let $A \subset \mathbb{R}^k$ be $\mathcal{L}$-definable, where $\mathcal{L}$ is arbitrary. Then the topological closure of A in $\mathbb{R}^k$ is $\mathcal{L}$-definable. Many other basic notions from point-set topology also are definable.

6. Let $A \subset \mathbb{R}^k$ be $\mathcal{L}$-definable. Then all level sets of A are $\mathcal{L}$-definable.

7. Closed polyhedra in $\mathbb{R}^k$ are $\mathcal{L}_{\mathbf{lin}}$-definable.

Guidelines for which sets are definable:

- Defining formulas cannot be infinitely long.
- Quantification over real numbers only is allowed.

Definability set-theoretically

The $\mathcal{L}$-definable subsets of $\mathbb{R}^n$ for $n = 1, 2, 3, \ldots$ is the smallest collection $\mathfrak{D} = \{\mathcal{D}_n \,|\, n \geq 1\}$ such that

1. Each $D \in \mathcal{D}_n$ is a subset of $\mathbb{R}^n$;
2. $\mathbb{R}^n \in \mathcal{D}_n$;
3. The graph of each $f\colon \mathbb{R}^n \to \mathbb{R}$ in $\mathcal{F}$ is in $\mathcal{D}_{n+1}$
4. Each $R \subseteq \mathbb{R}^n$ in $\mathcal{R}$ is in $\mathcal{D}_n$
5. For all $1 \leq i, j \leq n$, $\{(x_1, \ldots, x_n) \in \mathbb{R}^n \,|\, x_i = x_j\} \in \mathcal{D}_n$;
6. Each $\mathcal{D}_n$ is closed under intersection, union, and complement;
7. If $\pi\colon \mathbb{R}^n \to \mathbb{R}^m$ is a projection map $(x_1, \ldots, x_n) \mapsto (x_{i_1}, \ldots, x_{i_m})$ and $X \in \mathcal{D}_n$ then $\pi(X) \in \mathcal{D}_m$;
8. If π is as above and $Y \in \mathcal{D}_m$, then $\pi^{-1}(Y) \in \mathcal{D}_n$;
9. If $X \in \mathcal{D}_{n+m}$ and $\bar{b} \in \mathbb{R}^m$, then $\{\bar{a} \in \mathbb{R}^n \,|\, (\bar{a}, \bar{b}) \in X\} \in \mathcal{D}_n$.

Finer structure of definable sets

A definable set typically has several different definitions. We have the $\mathcal{L}_{\mathbf{alg}}$-definitions of the unit interval $[0,1]$:

$$\{x \in \mathbb{R} \mid 0 < x \wedge x < 1\}$$
$$\{x \in \mathbb{R} \mid \exists y \, \exists z \; x^2 + y^2 + z^2 = 1\}$$
$$\{x \in \mathbb{R} \mid \forall y \, \exists z \; x^2 + (y - z)^2 = 1\}.$$

Maxim Quantification adds complexity.

One goal of model theory is to attempt to analyze definable sets in a specific context by showing that definable sets can be defined by simple formulas.

Quantifier elimination All $\mathcal{L}$-definable sets can be defined by quantifier-free $\mathcal{L}$-formulas.

Quantifier elimination for $\mathcal{L}_{\mathbf{lin}}$-definable sets

Theorem. *For every $n = 1, 2, 3, \ldots$, every $\mathcal{L}_{\mathbf{lin}}$-definable subset of $\mathbb{R}^n$ can be defined by a quantifier-free $\mathcal{L}_{\mathbf{lin}}$-formula.*

For convenience, write rx instead of $\mu_r(x)$, where $r \in \mathbb{R}$.

The theorem tells us that every $\mathcal{L}_{\mathbf{lin}}$-definable subset of $\mathbb{R}^n$ is a finite boolean combination (i.e., finitely many intersections, unions, and complements) of sets of the form

$$\{(x_1, \ldots, x_n) \in \mathbb{R}^n \mid a_1x_1 + \cdots + a_nx_n > b\}$$

where $a_1, \ldots, a_n, b$ are fixed but arbitrary real numbers.

The $\mathcal{L}_{\mathbf{lin}}$-definable sets are called the *semilinear* sets.

By routine set theoretic manipulation, semilinear sets can be written as a finite union of the intersection of finitely many sets defined by conditions of the form

$$a_1x_1 + \cdots + a_nx_n + b = 0$$
$$c_1x_1 + \cdots + c_nx_n + d > 0.$$

Thus $\mathcal{L}_{\mathbf{lin}}$-definability reduces (basically) to linear algebra, which we understand well.

$\mathcal{L}_{\mathbf{lin}}$-definable subsets of $\mathbb{R}$

All $\mathcal{L}_{\mathbf{lin}}$-definable subsets of $\mathbb{R}$ are finite boolean combinations of sets of the form

$$\{x \in \mathbb{R} \mid ax > b\}.$$

Geometrically these are the union of finitely many (possibly unbounded) open intervals and points.

Consequence Neither $\mathbb{Z}$ nor $\mathbb{Q}$ is $\mathcal{L}_{\mathbf{lin}}$-definable.

Idea of the proof

Eliminate quantifiers one at a time (proceed inductively).

Example

Eliminate the quantifier for the $\mathcal{L}_{\mathbf{lin}}$-definable set

$$\{x \in \mathbb{R} \mid \exists y\, [2x - 3y > 2 \,\wedge\, 4x - 2y \leq 0]\}$$

High school algebraic elimination.

Question: What about $\mathcal{L}_{\mathbf{alg}}$-definability?

Next time

Some General References

D. Marker, *Model Theory: An Introduction* (Graduate Texts in Mathematics, vol. 217, New York: Springer-Verlag, 2002.

Two articles in

D. Haskell, A. Pillay, and C. Steinhorn (editors), *Model Theory, Algebra, and Geometry* (Mathematical Sciences Research Institute Publications, vol. 39), Cambridge: Cambridge University Press, 2000:

> Haskell, Pillay, and Steinhorn, "Overview."
>
> Marker, "Introduction to Model Theory."