FROM LOGIC TO GEOMETRY

PANTELIS E. ELEFTHERIOU

Logic is a scientific field traditionally practised within the disciplines of mathematics, philosophy and computer science. Model theory is a branch of mathematical logic which uses logical tools in order to study known and new mathematical structures (models). When those structures are of geometric nature, we tend to call this study *tame geometry.* This terminology is not so broadly used, since the branch is relatively new, and in this note we aim to explain its meaning.

The first person who used this terminology was the French geometer Grothendieck, who envisioned in his Esquisse d'un Programme [6] a '*topologie modérée*'. He asked whether there is a strict mathematical way to study restricted classes of geometric objects which, however, have better geometrical and topological properties. Among others, that class of objects should be closed under the usual set-theoretic operations, such as union, complement and projection.

Model theory, via tame geometry, offers one answer to Grothendieck's question. We risk the following definition:

Tame geometry is the study of those geometric objects that are ***definable*** *in some specific language from mathematical logic.*

That is, among all geometric objects one could possibly consider, we isolate and study only those that are definable in some specific mathematical language. The benefit of this intentional restriction is that tools from mathematical logic become available, which can then be applied to our class of objects in order to obtain new applications.

The key word in the above definition is that of *definability.* In mathematical logic, one first gives a definition of a *language*, then of a *structure* in that language, and finally of a set *definable* in that structure. The definition of definability is recursive and quite lengthy, and hence we omit it here. However, it is also very intuitive and we can capture it via some examples.

Example 0.1. The unit circle $S^1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ is definable in the structure $\langle \mathbb{R}, +, \cdot \rangle$. Indeed the equation of the unit circle uses $\cdot$ to express the squares, and $+$. It also uses $=$ and variables x, y, which are standard logical symbols and hence not mentioned explicitly, as well as 1, which is a constant of the universe $\mathbb{R}$ of our structure. Similarly, the unit disc $D^1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$ is definable in the structure $\langle \mathbb{R}, \leq, +, \cdot \rangle$.

Let us now try to identify the class of all definable sets in the structure $\langle \mathbb{R}, \leq, +, \cdot \rangle$. We claim that

$$X \subseteq \mathbb{R}^n \text{ definable in } \langle \mathbb{R}, \leq, +, \cdot \rangle \quad \Leftrightarrow \quad X \text{ `semialgebraic'}, \tag{1}$$

where the notion of a semialgebraic set can be defined in a purely geometrical way, see Definition 0.2 below. Before giving that definition and trying to argue for the above equivalence, let us point out that if we consider another structure, by changing the universe and the underlying language, the class of all definable sets may yield some other interesting class of geometric objects. For example:

$$X \subseteq \mathbb{C}^n \text{ definable in } \langle \mathbb{C}, +, \cdot \rangle \quad \Leftrightarrow \quad X \text{ constructible.}$$

In other words, via the notion of definability, model theory offers a uniform way to capture known classes of mathematical objects. Constructible sets are the objects of study in algebraic geometry, semialgebraic sets in *real* algebraic geometry. Here we focus on structures with universe $\mathbb{R}$.

Date: November 7, 2022.

Definition 0.2. A set $X \subseteq \mathbb{R}^n$ is *semialgebraic* if it is a Boolean combination of sets of the form

$$\{x \in \mathbb{R}^n : f(x) \geq 0\}, \tag{2}$$

where $f \in \mathbb{R}[X]$ is a polynomial. Sets of the form (2) are called *basic semialgebraic* sets.

It is straightforward from the definition that the class of all semialgebraic sets is closed under taking unions and complements. It is also closed under taking projections:

Theorem 0.3 (Tarski-Seidenberg 1950s)**.** *The class of semialgebraic sets is closed under taking projections.*

Therefore, the class of semialgebraic sets is closed under all set-theoretic operations envisioned by Grothendieck.

Where does logic appear in the above considerations? It appears in the expressions "Boolean combinations" and "projections". Let us convince ourselves of the validity of equivalence (1), even though we have not strictly defined the notion of definability. For the right-to-left direction, observe first that, since polynomials $f \in \mathbb{R}[X]$ are formed using $+, \cdot$, variables and coefficients from $\mathbb{R}$, just like the disc D^1, every basic semialgebraic set is definable in the structure $\langle \mathbb{R}, \leq, +, \cdot \rangle$. Moreover, Boolean combinations (that is, unions and complements) of two sets correspond to the standard logical symbols "or" and "negation". For example, if $A, B \subseteq \mathbb{R}^n$, then

$$A \cup B = \{x \in \mathbb{R}^n : x \in A \text{ or } x \in B\}.$$

Hence, the right-to-left direction of (1) is established. For the other direction, we need further to observe that the "existential quantifier", which we use in logic, corresponds to the set-theoretic operation of taking projections. For example, if $X \subseteq \mathbb{R}^2$ is defined via the formula $\varphi(x, y)$, then its projection onto the first coordinate is defined via the formula $\exists y\, \varphi(x, y)$. Hence, by the Tarski-Seidenberg theorem, it follows that every set definable in $\langle \mathbb{R}, \leq, +, \cdot \rangle$ is semialgebraic.

We may rephrase the Tarski-Seidenberg theorem in purely logical terms as follows:

Theorem 0.4. *The structure $\langle \mathbb{R}, \leq, +, \cdot \rangle$ eliminates quantifiers.*

We thus have a correspondence between logic and geometry: the notion of a semialgebraic set is captured via definable sets, and the content of the Tarski-Seidenberg theorem via quantifier elimination. This correspondence continues to a very extended level. Here we will only present the next step.

Quantifier elimination is a very powerful property in logic, since it ensures that any set we can define using quantifiers can also be defined without them. Thus, whenever we study a new structure in model theory, our first task is to examine whether it has quantifier elimination. But this property, being so strong, it almost always fails. Hence we try to replace it with some weaker one, again expressible in a logical way as well as capturing some geometric property. Let us take a second look at $\langle \mathbb{R}, \leq, +, \cdot \rangle$, and in particular at definable subsets of $\mathbb{R}$ (and not of any $\mathbb{R}^n$).

Fact 0.5. *$X \subseteq \mathbb{R}$ is definable in $\langle \mathbb{R}, \leq, +, \cdot \rangle \Leftrightarrow X$ is a finite union of points and intervals.*

Proof. Each basic semialgebraic subset of $\mathbb{R}$ is clearly of the desired form (namely, finite unions of points and intervals). Moreover, Boolean combinations of sets of the desired form remain of this form. Hence, inductively, every semialgebraic subset of $\mathbb{R}$ is a finite union of points and intervals. By (1), we are done. □

In the 1980s, van den Dries and Knight-Pillay-Steinhorn adopted the above property into a definition:

Definition 0.6 ([1, 7, 11])**.** An ordered structure $\mathcal{R} = \langle \mathbb{R}, \leq, \dots \rangle$ is *o-minimal* (where 'o' stands for 'order') if every definable set $X \subseteq \mathbb{R}$ in $\mathcal{R}$ is a finite union of points and intervals. Equivalently, X is definable in the structure $\langle \mathbb{R}, \leq \rangle$.

The novelty of the above definition is that, although it only requires a property for definable subsets of the universe, it has a number of non-trivial consequences for all definable subsets in any $\mathbb{R}^n$. These consequences are of geometrical/topological nature, much alike Grothendieck's vision. Here and below, the topology on $\mathbb{R}$ is taken to be the order-topology (generated by the open intervals), and on $\mathbb{R}^n$, the product topology.

Theorem 0.7. *Let $\mathcal{R} = \langle \mathbb{R}, \leq, \dots \rangle$ be o-minimal. We have:*

(a) If $X \subseteq \mathbb{R}^n$ is definable, then it can be partitioned into finitely many 'cells',

$$X = X_1 \cup \dots \cup X_k,$$

where each cell X_i is in particular connected.

(b) $f : X \to \mathbb{R}$ is definable (that is, its graph is a definable set), then the above partition can be done so that moreover each $f_{\restriction X_i}$ is continuous.

(c) There are no definable space-filling curves in $\mathbb{R}^2$, that is curves whose topological closure is the whole $\mathbb{R}^2$.

(d) The structure $\langle \mathbb{R}, \leq, \sin x \rangle$ is not o-minimal.

Proof. Properties (a) - (c) are non-trivial, and an extended account on o-minimality containing their proofs can be found in van den Dries [2]. Property (d) is easy to see: in the structure $\langle \mathbb{R}, \leq, \sin x \rangle$ one can define the set of integers

$$\mathbb{Z} = \{x \in \mathbb{R} : \sin x = 0\},$$

which is not a finite union of points and intervals. □

Property (d) may strike us as a weakness of o-minimality, since it excludes trigonometric functions from our study. However, this exclusion is not because of the 'shape' or topology of the sin function but because of its periodicity. It is exactly that periodicity that o-minimalists tried to avoid, as being 'wild behavior', in their attempt to realize Grothendieck's vision. As it turns out, if we restrict trigonometric functions to some bounded domain and add them to the real field, the resulting structure remains o-minimal. In fact, one can add all restricted real analytic functions along with the exponential and stay o-minimal. This theorem was one of the biggest breakthroughs in the early years of o-minimality, established in a series of works:

Theorem 0.8 ([3, 4, 13])**.** *Let $\mathbb{R}_{an,exp} := \langle \mathbb{R}, \leq, +, \cdot, \text{res. analytic}, \exp \rangle$ be the expansion of the real field by all analytic functions restricted to bounded boxes $[0,1]^n$, and the exponential map $\exp : \mathbb{R} \to \mathbb{R}$. Then $\mathbb{R}_{an,exp}$ is o-minimal.*

With a strong toolbox from o-minimality on the one hand, and an abundance of o-minimal structures on the other, it is reasonable to expect some new applications. We describe one trend of such applications, namely, of o-minimality to Diophantine geometry. These applications are often obtained by reducing big conjectures from Diophantine geometry (such as Manin-Mumford, André-Oort) to the following theorem.

Theorem 0.9 (Pila-Wilkie [10])**.** *Let $\mathcal{R} = \langle \mathbb{R}, \leq, +, \cdot, \dots \rangle$ be an o-minimal structure, and $X \subseteq \mathbb{R}^n$ a definable set. Assume that X contains 'many' rational points. Then X contains an infinite semialgebraic set A.*

We define the notion of containing 'many' rational points briefly, as follows: the set X may contain infinitely many rational points (that is, tuples with all coordinates being rational). If we bound the enumerators and denominators of those rationals by some number T, then X contains only finitely many such points, say $N(X,T)$ many. We say that X contains *many rational points* if $N(X,T)$ increases at least polynomially in terms of T. That is, for every $\epsilon \in \mathbb{R}^{>0}$,

$$N(X,T) > O(T^\epsilon).$$

The gist of the Pila-Wilkie theorem is that, knowing our definable set X contains many rational points, one can recover an infinite semialgebraic subset; that is, a set definable only in the real field $\langle \mathbb{R}, \leq, +, \cdot \rangle$. This statement is of Diophantine nature and matches up with the following statement.

Theorem 0.10 (Manin-Mumford Conjecture, roughly)**.** *Let $V \subseteq (\mathbb{C}^*)^d$ be an algebraic variety. Assume V contains 'many' torsion points (for example, those may be Zariski dense in V). Then V contains a coset aH of an algebraic subgroup $H \leqslant (\mathbb{C}^*)^d$.*

That is, again, knowing V contains 'many' special points (torsion points, in this case), we can recover (a coset of) an algebraic group in it.

Sketch of Manin-Mumford. [1] The reduction to Pila-Wilkie is done via the map

$$\theta : \mathbb{C} \to \mathbb{C}^*, z \mapsto e^{2\pi i z}.$$

It is easy to see that

$$\theta^{-1}(V) \text{ has many rational points } \Leftarrow V \subseteq \mathbb{C}^* \text{ many torsion points}$$

Moreover,

$$X = \theta^{-1}(V) \subseteq \mathbb{R}^2 \text{ is definable in } \mathbb{R}_{an,exp}!$$

Let $A \subseteq X$ be as in the Pila-Wilkie theorem. Then it is not so hard to show that $\theta(A)$ contains a coset aH of an algebraic group, as needed. For details, see Marker [8]. □

Currently, a whole group of conjectures from Diophantine geometry are being tackled using (methods around) the Pila-Wilkie theorem. The Pila-Wilkie theorem is also being extended to more general structures, beyond the o-minimal framework (and beyond the scope of this note), such as in [5]. Applications of the extended Pila-Wilkie theorems are pending to be explored.

Concluding, in tame geometry, instead of studying all geometric objects one could possibly consider, we focus only on those that can be defined using a specific mathematical language. Such a restriction often yields new tools from logic that we can be applied on the class of all definable sets in order to obtain new applications. The idea of using logic in order to restrict the universe of our interest down to objects better manipulated is not new. It goes back to Gödel, who, in his celebrated Incompleteness Theorem (1931), worked with the language of arithmetic, and instead of considering all subsets of $\mathbb{N}^n$, he only dealt with those that can be defined in $\langle \mathbb{N}, +, \cdot \rangle$. Within that restricted fragment, he was able to 'code' the Liar sentence (which says 'I am lying') and produce the first sentence in arithmetic that cannot be proved nor disproved from the Peano Axioms, refuting Hilbert's dream of axiomatizing the whole of mathematics. Notably, Gödel's coding functions are definable in an even more restricted fragment of Arithmetic, that of *recursive functions*, which later on gave rise to the Turing machines, the prodromes of the current computers. Tame geometry is a distant descendant of Gödel's logical considerations and his exploitation of the logical power of restricted languages to produce striking results. Of course, in every mathematical study one restricts their focus to a specific class of objects, but when this restriction becomes the object of study itself, and is relevant to geometry, we call that study *tame geometry.*

References

[1] L. van den Dries, *Remarks on Tarski's problem concerning* $(R, +, -, exp)$, Logic Colloquium (1982), 97-121, (G. Lolli, G. Longo and A. Marcja, eds.), North-Holland, 1984.

[2] L. van den Dries, TAME TOPOLOGY AND O-MINIMAL STRUCTURES, Cambridge University Press, Cambridge (1998).

[3] L. van den Dries, A. Macintyre, and D. Marker, *The elementary theory of restricted analytic fields with exponentiation*, Ann. Math. 140 (1994), 183-205.

[4] L. van den Dries and C. Miller, *On the real exponential field with restricted analytic functions*, Israel J. Math. 85 (1994), 19–56.

[1] We note that Manin-Mumford Conjecture was previously solved without logical methods. The André-Oort Conjecture (or rather certain cases of it) are indeed only proved using o-minimal methods, by Pila [9].

[5] P. Eleftheriou, *Counting algebraic points in expansions of o-minimal structures by a dense set*, arXiv:1708.03936v3 (2018), 14 pages.
[6] A. Grothendieck, *Esquisse d'un Programme (1984 manuscript)*, [Sketch of a program], With an English translation on pp. 243–283. London Math. Soc. Lecture Note Ser., 242, Geometric Galois actions, 1, 5–48, Cambridge Univ. Press, Cambridge, 1997.
[7] J. Knight, A. Pillay and C. Steinhorn, *Definable sets in ordered structures II*, Trans. Amer. Math. Soc. 295 (1986), 593-605.
[8] D. Marker, *The Pila-Zannier proof of the Manin-Mumford Conjecture*, online notes, http://homepages.math.uic.edu/∼marker/orsay/mm.pdf
[9] J. Pila, *O-minimality and the André-Oort conjecture for* $\mathbb{C}^n$, Ann. Math. 173 (2011), 1779-1840.
[10] J. Pila, A. J. Wilkie, *The rational points of a definable set*, Duke Math. J. 133 (2006), 591–616.
[11] A. Pillay, C. Steinhorn, *Definable sets in ordered structures I*, Trans. Amer. Math. Soc. 295 (1986), 565-592.
[12] T. Scanlon, *Counting special points: logic, Diophantine geometry, and transcendence theory*, Bull. Amer. Math. Soc. (N.S.), 49(1):51–71, 2012.
[13] A. J. Wilkie, *Model Completeness Results for Expansions of the Ordered Field of Real Numbers by Restricted Pfaffian Functions and the Exponential Function*, J. Amer. Math. Soc. 9 (1996), 1051–1094.

Department of Mathematics and Statistics, University of Konstanz, Box 216, 78457 Konstanz, Germany

Email address: `panteleimon.eleftheriou@uni-konstanz.de`