The role of Hilbert problems in real algebraic geometry

Marie-Françoise Roy

IRMAR, CNRS, Université de Rennes I, Campus de Beaulieu 35042 Rennes cedex, France mfroy@maths.univ-rennes1.fr

Real algebraic geometry studies sets defined by systems of polynomial equations over the reals [5]. Three problems of Hilbert [18] are related to important aspects of real algebraic geometry.

We examine first these problems, what is known about their solution, and what developments they led to. We end with a short discussion on the role they played in the development of real algebraic geometry during this century.

In many cases, the references given in this text point to text books or survey papers where simple proofs and full references can be found rather than to the original papers.

1. Hilbert's 17'th problem

It is obvious that a polynomial which is a sum of square is everywhere nonnegative. A very natural question is the following.

• Is every polynomial everywhere non negative a sum of squares?

A polynomial everywhere non negative is not always a sum of squares of polynomials. This result is due to Hilbert [16] but Motzkin [34] was the first in 1967 to construct explicit examples of this situation. For example

$$P = Z^6 + X^4 Z^2 + X^2 Y^4 - 3X^2 Y^2 Z^2$$

is everywhere non negative and is not a sum of squares of polynomials. This is not too hard to prove since the degree of the polynomials to look for is at most 3.

Minkowski sugggested to Hilbert the following reformulation, which Is Hilbert's 17 th problem :

• Is every polynomial everywhere non negative a sum of squares of rational functions?

Now, since denominators are allowed, the space of search for an expression as a sum of squares is much bigger, and there are no a priori limitations on the degrees to consider.

Emil Artin's positive answer [2], in 1925, is one of the most convincing successes of modern algebra, which was starting at that time.

In order to prove a result about the reals, the method of the proof uses much more abstract objects, namely real closures of the field of rational functions.

Artin's proof goes this way:

- Consider a polynomial P which is not a sum of squares of rational functions with real coefficients.
- Since P is not a sum of squares, the set of sums of squares is a proper cone of the field of rational functions which does not contain P (a cone of a ring contains the squares, is closed under addition and multiplication and a proper cone also does not contain -1).
- Using Zorn's lemma, and taking a maximal proper cone which does not contain P, we get a total order on the field of rational functions for which P is negative.
- Taking the real closure of the field of rational functions for this order, which is the biggest possible ordered field extending the given order and algebraic over the field of rational functions, we get a field in which P takes negative values.
- It remains to prove that if P takes negative values in a real closed field containing the reals, P takes negative values over the reals.

So, for example, Motzkin's polynomial above is a sum of squares of rational functions, even it may not be easy to write it explicitly so.

Artin's proof we just sketched is the starting point of the abstract theory of the reals. Real closed fields originally defined by Artin and Schreier [3] as ordered fields with no algebraic ordered extension have been characeterized by Tarski [46] later as fields which are ordered, where every positive element is a square and every odd degree polynomial has a root. The real closure of an ordered field is the smallest real closed field containing it.

Many problems remain after Artin's proof. Among them:

- quantitative aspects: how many squares?
- effectivity problems: is there an algorithm checking that a given polynomial is everywhere non negative and providing the sum of squares?
- complexity problems : what are the best possible bounds on the degrees of the sum of squares ?

A bound on the number of squares was proved by Pfister in 1967 [36]: 2^n squares are enough if n is the number of variables of the polynomial P. It is remarkable that the degree of the polynomial plays no role in the bound on the number of squares needed. The proof uses Pfister's theory of multiplicative quadratic forms [36].

The exact bound is far from being known, since the only known lower bound on the number of squares needed is n + 2.

Since Artin's proof is based on Zorn's lemma, no explicit bound can be easily extracted from its inspection (see [10, 27] though). The explicit construction of the sum of squares is a difficult problem and has attracted much attention and many contributions (see [10, 11] for a detailed account of the litterature on this topic).

Hilbert's 17'th problem can be seen as part of a much more general problem, which is to provide a dictionary between algebra and geometry in the real case.

In complex algebraic geometry, this dictionary is very classical and given by the correspondance between algebraic sets and radical ideals (a particularly easy to read presentation can be found in [8]). Artin's result relates non negativity which is a geometric property to sums of squares, which is an algebraic notion. Stengle's positivstellensatz [45], proved in 1974, provides a quite general version of a dictionary between algebra and geometry in the real case. The version we present here can be found in [5].

[weak Positivstellensatz]

Let R be a real closed field, \mathcal{F}, \mathcal{G} and \mathcal{H} three finite subsets of $R[X_1, \ldots, X_n]$, \mathcal{C} the cone generated by \mathcal{F} , \mathcal{M} the monoid generated by \mathcal{G} and \mathcal{I} the ideal generated by \mathcal{H} .

The following are equivalent:

(i) The set

$$\{x \in \mathbb{R}^n \mid \forall f \in \mathcal{F} \ f(x) \ge 0 \ , \forall g \in \mathcal{G} \ g(x) \ne 0 \ , \ \forall h \in \mathcal{H} \ h(x) = 0\}$$

is empty.

(ii) There exists $f \in \mathcal{C}$, $g \in \mathcal{M}$ and $h \in \mathcal{I}$ such that $f + g^2 + h = 0$.

This result can be interpreted as follows: if a family of inequalities and equalities is incompatible, there exists an algebraic certificate testifying it.

Using a classical trick due to Rabinovitch the following result [45] follows.

[Positivstellensatz]

Let G a finite subset of $R[X_1, \ldots, X_n]$ and

$$W = \{ x \in R^n \mid \forall g \in \mathcal{G} \ g(x) \ge 0 \ \} \ .$$

Let C bethe cone of $R[X_1, \ldots, X_n]$ generated by G, and let $f \in R[X_1, \ldots, X_n]$. Then $\forall x \in W$ $f(x) > 0 \Leftrightarrow \exists g, h \in C$ fg = 1 + h.

As an immediate consequence of Positivstellensatz, when $\mathcal{G} = \emptyset$, one recovers that a nonnegative polynomial is a quotient of sums of squares, hence also a sum of squares.

The proof of these results rely on Zorn's lemma and gets clear using the notion of a prime cone [7, 5].

It is natural to wonder whether there exists an algorithm producing the algebraic identities announced. The answer is yes [26]; a lot remains to do to provide an efficient algorithm [28].

2. Hilbert's 16'th problem

The first part of this problem is to determine

• the different possible shapes of an algebraic curve, or more generally of a real hypersurface in projective space.

We do not discuss the second part of the problem here.

Consider a nonsingular algebraic hypersurface H of degree d in $\mathbb{R}P^n$. This polynomial has a set of zeroes $\mathbb{R}H$ in the real projective space $\mathbb{R}P^n$. The first part of Hilbert's 16'th problem asks what are the possible topological types of the painrs $(\mathbb{R}P^n, \mathbb{R}H)$ (for a given degree d). For the curves of $\mathbb{R}P^2$ the complete answer is known only up to degree 7. For the surfaces of $\mathbb{R}P^3$, the complete answer is known only up to degree 4.

To make progress on this problem it is needed to work in two main directions: first to find restrictions on the topological types of $(\mathbb{R}P^n, \mathbb{R}H)$, second to construct hypersurfaces with the types which are not forbidden.

In the study of the topology of real algebraic curves, it is traditional to pay a special attention to M-curves. An M-cufrve is a curve of $\mathbb{R}P^2$ such that its set of real points has the maximal number of connected components; this number is equal to (d-1)(d-2)/2+1 for degree d according to Harnack's theorem [14].

An *oval* of a real algebraic projective curve Γ is a connected component of Γ whose complement is not connected. The orientable connected component of this complement is the *interior* of the oval.

Harnack's theorem gives nor estriction on the relative position of the ovals. The relative position of ovals in the projective place can be described in terms of nest, when an oval contains another oval in its interior. The depth of an oval Ω of Γ is the number of obvals containing Ω in their interiors. Bezout's theorem creates restrictions. An M-curve of degree 4 with 4 ovals cannot have a nest, because otherwise a line having at least 6 poins of intersection with the curve could be easily constructed. So the only configuration for an M-courve of degree 4 is 4 ovals without any nest.

This simple arugment is no more sufficient in degree 6. This case has been completely solved in 1971 only. The M-curves of degree 6 have 11 ovals. Bezout's theorem proves that an oval of depth 2 is impossible and also that there exist at most one oval with other ovals in its interior. A construction due to Harnack gives a M-curve with the following configuration: an oval containing another oval in its interior and the nine other ovals outside, without any nest. This configuration is denoted by $1\langle 1\rangle \coprod 9$. A construction by Hilbert [17] gives a different configuration denoted $1\langle 9\rangle \coprod 1$: an oval containing nine ovals in its interior, the last oval being outside.

People wondered for a long time whether there are other configurations for an M-curve of degree 6, until Gudkov constructed such a curve with a configuration $1\langle 5 \rangle \coprod 5$ [13].

These three configurations are the only possible for the M-curves of degree 6. This is a consequence of a famous congruence conjectured by Gudkov and proved by Rokhlin [41, 48].

Let Γ be a non singular real algebraic curve of even degree 2k in $\mathbb{P}_2(\mathbb{R})$. It is possible to chose an homogeneous equation F of Γ , such that $F(x,y,z) \leq 0$, for every (x:y:z) outside all the ovals of Γ . Then, with

$$B_{+} = \{(x:y:z) \in \mathbb{P}_{2}(\mathbb{R}) \mid F(x,y,z) \geq 0\}$$
:
[Rokhlin-Gudkov's congruence]

The Euler-Poincaré characsteristic $\chi(B_+)$ is congruent to k^2 modulo 8.

The Euler-Poincaré characteristic of $\chi(B_+)$ can be computed in the following way. An oval of Γ is *even* (resp. *odd*) if its depth is even (resp. odd). The number of even (resp. odd) ovals of Γ is denoted by p (resp. n). Then $\chi(B_+) = p - n$.

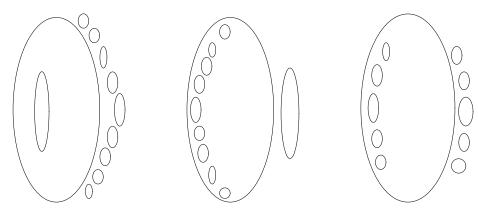


Figure 1

For an M-curve of degree 6, p+n=11 and, using true congruenc, $p-n\equiv 9\pmod 8$. The only possibilites are $(p=10\ ,\ n=1),\ (p=6\ ,\ n=5)$ or $(p=2\ ,\ n=9)$. Thus the three configurations described above are the only possible.

The njumbers p and n have been considered first by Virginia Ragsdale [37] who porposed the conjecture that for every curve of degree 2k,

$$p \le \frac{3k(k-1)}{2} + 1$$
, $n \le \frac{3k(k-1)}{2}$.

Virginia Ragsdale (1870-1945)

She graduated from Guilford College in 1892. She won the first scholarship established by Bryn Mawr College for a Guilford woman graduating with the highest degree average. She took her Bachelor degree in 1896 and was awarded the Bryn Mawr European Fellowphip. She chose to go to Gottingen for one year, to study under Hilbert and Klein. She completed her Ph D in Bryn Mawr in 1096. Her paper "On the arrangement of the real branches of plane algebraic curves" [37] was published immediately after her dissertation. She was instructor, assistant professor, professor and finally head of the department at Woman's College in North Carolina. She retired from teaching in 1928 and took care of her ill mother, taking the responsability of the home.

The spectacular development of the topology of real algebraic varieties in the years 1970 implies new restrictions on the topology of a real algebraic variety. V. Arnold [1], V. Rokhlin [41, 42, 43] et V. Kharlamov [22, 23, 24] have obtained important general obstructions. The discovery of new invariants for the varieties of dimension 4 by Seberg and Witten and the proof of Thom's conjecture in 1994 [25] have implied new restrictions in the topology of real algebraic curves [31].

If a lot of work had been done on obstructions, the methods of construction had not changed much from the 19th century. In 1980, Viro proposed a completely new method to construct real algebraic varieties [21, 20]. The combinatorial patchwork, which is a particular case of Viro's method, gives a recipee to construct hypersurfaces using a simple combinatorial

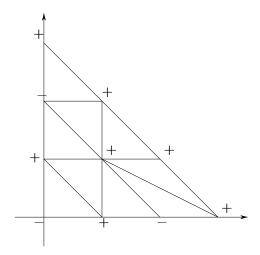


Figure 2

procedure. For simplicity the recipee is explained here for curves but the general case is completely similar.

In order to construct a real algebraic curve in $\mathbb{R}P^2$, the following combinatorial data are given.

Let d be a positive integer and T the triangle

$$\{(x, y) \in \mathbb{R}^2 : x \ge 0, y \ge 0, x + y \le d\}.$$

The number d will be the degree of the curve constructed and the triangle T will be the Newton polygon of the curve.

Suppose that T is triangulated in such a way that all the vertices of the triangulation have integer coordinates. Suppose also that a sign distribution, $a_{i,j} = \pm$ is given at the vertices of the triangulation, (see Figure 2). A piecewise linear curve is L in $\mathbb{R}P^2$ is constructed as follows.

Take copies $T_x = s_x(T)$, $T_y = s_y(T)$, $T_{xy} = s(T)$ of T, where $s = s_x \circ s_y$ and s_x , s_y are reflexions against the axis of coordinates. Extend the triangulation of T to a symmetric triangulation of $T \cup T_x \cup T_y \cup T_{xy}$, and extend the sign distribution to a sign distribution at the vertices of the extended triangulation with the following rule: when a vertex is transformed into its mirror image with respect to a coordinate axis, its sign is preserved when the distance to the axis is even and changed when this distance is odd. (see Figure 3).

If a triangle of the triangulation has vertices with different signs, a segment is drawnfrom the middle of the edges to isolate + from -. The union of these segments is denoted by L' and is contained in $T \cup T_x \cup T_y \cup T_{xy}$ (see Figure 3). Sides of $T \cup T_x \cup T_y \cup T_{xy}$ are glued using s. The space T_* so obtaine dis homeomorphic to $\mathbb{R}P^2$. The curve L is the image of L' in T_* .

A pair (T_*, L) is a chart of a real algebraic curve C in $\mathbb{R}P^2$, if there exists an homeomorphism between the pairs (T_*, L) and $(\mathbb{R}P^2, \mathbb{R}C)$.

Suppose now that the triangulation of T is convex. This means that there exist a piecewise convex function $\nu: T \longrightarrow \mathbb{R}$ linear on each triangle of the triangulation and not linear on the union of two triangles.

[Viro's theorem]

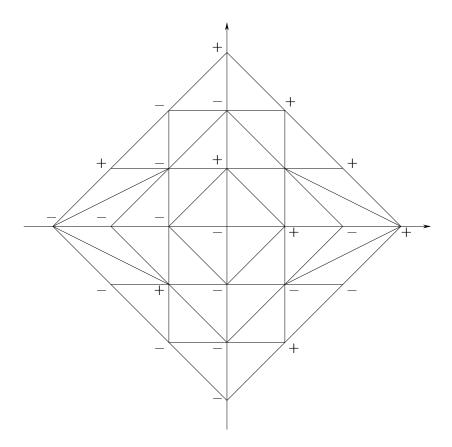


Figure 3

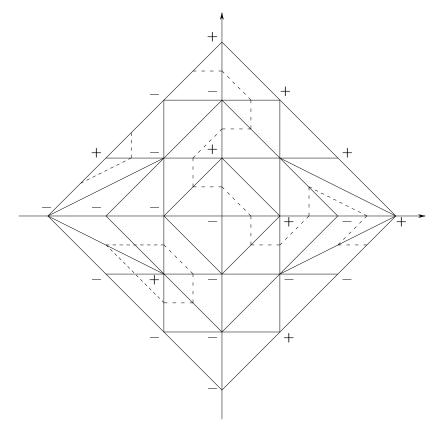


Figure 4

If the triangulation de T is convex, there exists a nonsingular real algebraic curve C of degree d in $\mathbb{R}P^2$ with chart (T_*, L) .

A curve with chart (T_*, L) is called a T-curve.

It is easy to verify that the triangulation of

$$\{(x, y) \in \mathbb{R}^2 : x \ge 0, y \ge 0, x + y \le 3\}$$

in Figure 2 is convex. Thus, Figure 3 is a topological picture of acurve of degree 3 in s $\mathbb{R}P^2$. The set of real points of this curve has two connected components in $\mathbb{R}P^2$.

The class of T-courbes is quite rich. For example all curve of $\mathbb{R}P^2$ up to degree 6 are T-curves. For every degree, there exists a T-curve which is a M-curve. Counter-examples to Ragsdale conjecture have been constructed as T-curves. [19, 20, 21]. There are also real algebraic curves of $\mathbb{R}P^2$ which are not T-curves [19].

3. Hilbert's 10'th problem

The problem is to answer to the following question:

• is there an algorithm deciding the existence of integer solutions to a set of Diophantine equations ?

this is to a set of polynomial equations with integer coefficients.

The answer is of course yes for univariate polynomials since there are easy bounds on the roots in terms of the coefficients.

The no answer for the general problem was proved by a young russian mathematician Matiyasevich [29, 30] in 1972. His work was using former contributions by Davis, Putnam and Robinson [9, 40].

Julia Robinson (1919-1985)

She took her PH. D. under Tarski in 1948. She made a major contribution to the solution of Hilbert's 10'th problem [9, 40]. In 1976 she became the first woman mathematician to be elected to the National Academy of Sciences. In 1982 she was nominated to the presidency of the American Mathematical Society. A book written by her sister Constance Reid is devoted to her [38].

Interestingly the corresponding problems when real solutions are looked for has a positive answer. The decision problem over the reals

• is there an algorithm deciding the existence of real solutions to a set of polynomial equation with integer coefficients?

was solved with a yes answer by Tarski [46] and Seidenberg [44].

Again the existence of an algorithm raises complexity questions. This is an active field of research [6, 12, 39, 15, 4]

4. Brief discussion

The role of Hilbert's problem in the development of real algebraic geometry was very important, as it can be understood clearly from the preceding developments.

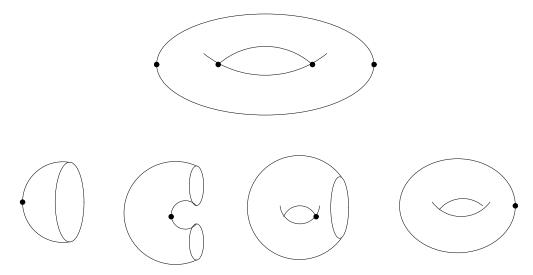
But some very important ideas were developed without any connection to these problems.

Morse for example related the change in topology to the existence and local behaviour of critical points of a function [32].

Morse theory plays a key role in the quantitative [35, 33, 47] and algorithmic aspects of real algebraic geometry [12, 4].

Olga Oleinik

A former professor at Moscow University, now retired, she lives in Russia. She worked in several fields of mathematics: topology of real algebraic varieties, PDE. Over 70 years old, she is still publishing papers.



Figure

References

- [1] V. Arnold, On the arrangement of ovals of the real algebraic curves, involutions of fourdimensional smooth manifolds, and arithmetic of integer quadratic forms. Funct. Anal. Appl. 5 3 1-9, 1971.
- [2] E. Artin, Über die Zerlegung definiter Funktionen in Quadrate. Hamb. Abh. 5, 100-115, 1927. The collected papers of Emil Artin, 273-288 Addison-Wesley, 1965
- [3] E. Artin, O. Schreier, Algebraische Konstruktion reeller Körper. Hamb. Abh. 5, 85-99, 1926. The collected papers of Emil Artin, 258-272 Addison-Wesley, 1965.
- [4] S. Basu, R. Pollack, M.-F. Roy, On the combinatorial and algebraic complexity of Quantifier elimination. J. Assoc. Comput. Machin., 43, 1002–1045, 1996.
- [5] J. Bochnak, M. Coste, M.-F. Roy, *Real algebraic geometry*. Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge, Bd. 36, Springer-Verlag, 1999.
- [6] G. Collins, Quantifier elimination for real closed fields by cylindrical algebraic decomposition. Autom. Theor. form. Lang., 2nd GI Conf., Kaiserslautern 1975, Lect. Notes Comput. Sci. 33, 134-183, 1975.
- [7] M. Coste, M.-F. Roy, La topologie du spectre réel. Contemp. Math. 8, 27-59, 1982.
- [8] D. Cox, J. Little, D. O'Shea, *Ideals, varieties and algorithms. An introduction to computational algebraic geometry and commutative algebra*. Undergraduate Texts in Mathematics. Springer-Verlag, 1997.
- [9] M. Davis, H. Putnam, J., Robinson, The decision problem for exponential Diophantine equations. Ann. Matg. 2 74 425-436, 1961.
- [10] C. N. Delzell, *Kreisel's unwinding of Artin's proof.* 113–245 in Kreiseliana: : About and Around Georg Kreisel, ed. P. Odifreddi, A K Peters, 1996.

- [11] C. N. Delzell, L. González-Vega, H. Lombardi A continuous and rational solution to Hilbert's 17'th problem and several Positivstellensatz cases. in: Computational Algerbraic Geometry. Eds. Eyssette F., Galligo A., Birkhäuser Progress in Math. 109, 61-76, 1993.
- [12] D. Grigor'ev, N. Vorobjov, Solving systems of polynomial inequalities in subexponential time. J. Symb. Comput. 5, No.1/2, 37-64, 1988.
- [13] D.A. Gudkov, Construction of a new series of M-curves. Sov. Math., Dokl. 12, 1559-1563, 1971; translation from Dokl. Akad. Nauk SSSR 200, 1269-1272, 1971.
- [14] A. Harnack, Über die Vieltheiligkeit der ebenen algebraischen Kurven. Math. Ann. 10, 189-198, 1876.
- [15] J. Heintz, M.-F. Roy, P. Solerno, Sur la complexité du Principe de Tarski-Seidenberg. (On the complexity of the Tarski-Seidenberg principle). Bull. Soc. Math. Fr. 118, No.1, 101-126, 1990.
- [16] D. Hilbert, Über die Darstellung definiter Formen als Summe von Formenquadraten. Math. Ann. 32, 342-350, 1888. Ges. Abh. vol. 2, 154-161. Chelsea Publishing Company, 1965.
- [17] D. Hilbert, Über die reellen Züge algebraischer Kurven. Math. Ann. 38, 115-138, 1891. Ges. Abh. vol. 2, 415-436. Chelsea Publishing Company, 1965.
- [18] D. Hilbert, *Mathematische Probleme*. Arch. Math. Physik 1, 44-63, 213-277, 1901. Ges.
 Abh. vol. 3, 290-323. Chelsea Publishing Company, 1965.
- [19] I. Itenberg, Counter-examples to Ragsdale conjecture and T-curves. Contemp. Math. 182, 55-72, 1995.
- [20] I. Itenberg, Viro's method and T-curves. Gonzalez-Vega, Laureano (ed.) et al., Algorithms in algebraic geometry and applications. Proceedings of the MEGA-94 conference, Santander, Spain, April 5-9, 1994. Birkhaeuser. Prog. Math. 143, 177-192, 1996.
- [21] I. Itenberg, O. Viro, Patchworking algebraic curves disproves the Ragsdale conjecture. Math. Intelligencer, 18 (4), 19-28, 1996.
- [22] V. Kharlamov, Generalized Petrovsky's inequality. Funct. Anal. Appl. 8 50-56, 1974.
- [23] V. Kharlamov, Additional congruences for the Euler characteristic of real algebraic variety of even degree. Funct. Anal. Appl. 9 51-60, 1975.
- [24] V. Kharlamov, Topological types of non-singular surfaces of degree 4 in $\mathbb{R}P^3$. Funct. Anal. Appl. 10 55-68, 1976.
- [25] P. Kronheimer, T. Mrowka, The genus of embedded surfaces in the projective plane. Math. Res. Lett. 1, No.6, 797-808, 1994.
- [26] H. Lombardi, Nullstellensatz réel effectif et variantes. C.R. Acad. Sci. Paris 310 635-640, 1990.
- [27] H. Lombardi, Relecture constructive de la théorie d'Artin-Schreier. Annals of Pure and Applied Logic 91, 1998. 59–92.

- [28] H. Lombardi, M.-F. Roy, Elementary complexity of positivstellensatz. In preparation
- [29] Y. Matiyasevich, On recursive unsolvability of Hilbert's tenth problem. Proceeding of Fourth Inernational Congress in Logic, Methodology and Philosophy of Science, Bucharest 1971. North Holland 89-110, 1973.
- [30] Y. Matiyasevich, *Hilbert's tenth problem*. MIT Press Series in the Foundations of Computing. MIT Press, 1993.
- [31] G. Mikhalkin, Adjunction inequality for real algebraic curves. Math. Res. Lett. 4, 1, 45–52, 1997.
- [32] J. Milnor, *Morse theory*. Annals of Mathematical Studies, Princeton University Press, 1963.
- [33] J. Milnor, On the Betti numbers of real varieties. Proc. Am. Math. Soc. 15, 275-280, 1964.
- [34] T. S. Motzkin, *The arithmetic-geometric inequality*. Inequalities, O. Shisha ed., 205-224. Academic Press, 1967.
- [35] O. A. Oleinik, Estimates of the Betti numbers of real algebraic hypersurfaces. Mat. Sb. (N.S.) 28 (70), 635-640, 1951.
- [36] A. Pfister, Quadratic forms with applications to algebraic geometry and topology. London Mathematical Society Lecture Note Series. 217. Cambridge Univ. Press, 1995.
- [37] V. Ragsdale, On the arrangement of the real branches of plane algebraic curves. Am. J. Math., 28, 377-404, 1906.
- [38] C. Reid, Julia, a life in mathematics. The Mathematical Association of America, 1996.
- [39] J. Renegar, On the computational complexity and geometry of the first-order theory of the reals, parts I, II and III. J. Symb. Comput., 13 (3) 255–352, 1992.
- [40] J. Robinson, Unsolvable Diophantine problems. Proc. Amer. Math. Soc. 22 534-538, 1969.
- [41] V. Rokhlin, Congruences modulo 16 in Hilbert's sixteenth problem. Funct. Anal. Appl. 6 58-64, 1972.
- [42] V. Rokhlin, Complex orientations of real algebraic curves. Funct. Anal. Appl. 8 71-75, 1974.
- [43] V. Rokhlin, Complex topological characteristics of real algebraic curves. Russian Math. Surveys 33 5 85-98, 1978.
- [44] A. Seidenberg, A new decision method for elementary algebra. Ann. Math. 60, 365-374, 1954
- [45] G. Stengle, A Nullstellensatz and a Positivstellensatz in semialgebraic geometry. Math. Ann. 207, 87-97, 1974.
- [46] A. Tarski, A decision method for elementary algebra and geometry. Prepared for publication by J.C.C. Mac Kinsey, Berkeley, 1951.

- [47] R. Thom, Sur l'homologie des variétés algébriques réelles. Differential and combinatorial topology, 255-265. Princeton Univ. Press, 1965.
- [48] G. Wilson, Hilbert's sixteenth problem. Topology 17, 53-73, 1978.