

A continuous and rational solution to Hilbert's 17th problem and several cases of the Positivstellensatz

CHARLES N. DELZELL¹ LAUREANO GONZÁLEZ-VEGA²
HENRI LOMBARDI

Abstract: From the Positivstellensatz we construct a continuous and rational solution for Hilbert's 17th problem and for several cases of the Positivstellensatz. The solutions are obtained using an especially simple method.

I. Introduction.

Let \mathbf{K} be an ordered field, and \mathbf{R} its real closure. Hilbert's 17th problem asks if an everywhere nonnegative polynomial $f \in \mathbf{K}[\mathbf{X}] := \mathbf{K}[X_1, \dots, X_n]$ can be expressed as a sum of squares of rational functions in $\mathbf{K}(\mathbf{X})$ with positive weights in \mathbf{K} . Since the answer is well known to be 'Yes,' we now seek more information, in particular, on the way the coefficients of the solution can vary in terms of the coefficients of f . The main tools we shall use to obtain this extra information are the notion of semipolynomial function and the Positivstellensatz for $\mathbf{K}[\mathbf{X}]$.

A semipolynomial function with coefficients in \mathbf{K} (a \mathbf{K} -semipolynomial) from \mathbf{R}^n to \mathbf{R} is a function obtained by a finite iteration of composition of polynomials in $\mathbf{K}[\mathbf{X}]$ and the function absolute value. A well-known proposition, not used here, assures that the set of \mathbf{K} -semipolynomials agrees with the minimal max-min stable set of functions containing polynomials in $\mathbf{K}[\mathbf{X}]$ (see, for example, [Del₄]).

By the name Positivstellensatz we shall refer to the more general version of this theorem, i.e., the one assuring that it is possible to associate to every incompatible finite conjunction of generalized sign conditions on a list of polynomials in $\mathbf{K}[\mathbf{X}]$ an algebraic identity in $\mathbf{K}[\mathbf{X}]$ making this incompatibility evident (see the beginning of §II).

¹ Supported by NSF, the Louisiana Board of Regents Research and Development Program (Education Quality Support Fund), and the Alexander von Humboldt Foundation.

² Partially supported by CICYT PB 89/0379/C02/01 and Esprit/Bra 6846 (Posso).

Using this Positivstellensatz we shall provide, in a constructive way, a continuous and rational solution for Hilbert’s 17th problem. More precisely, let $f_{n,d}(\mathbf{c}, \mathbf{X})$ be the general polynomial of degree d in the n variables $\mathbf{X} := (X_1, \dots, X_n)$ with coefficients $\mathbf{c} := (c_1, \dots, c_m)$, and consider the semialgebraic set

$$\mathbb{F}_{n,d} = \{\mathbf{c} \in \mathbb{R}^m : \forall \mathbf{x} \in \mathbb{R}^n \quad f_{n,d}(\mathbf{c}, \mathbf{x}) \geq 0\}.$$

Theorem II.5. (Main Theorem) *The general polynomial $f_{n,d}$ of degree d in n variables can be written as a weighted sum of squares of rational functions*

$$f_{n,d}(\mathbf{c}, \mathbf{X}) = \sum_j p_j(\mathbf{c}) \left(\frac{q_j(\mathbf{c}, \mathbf{X})}{k(\mathbf{c}, \mathbf{X})} \right)^2$$

(for all $\mathbf{c} \in \mathbb{R}^m$), where

- $k(\mathbf{c}, \mathbf{X})$ and the $q_j(\mathbf{c}, \mathbf{X})$ are polynomials in the variables \mathbf{X} whose coefficients are \mathbb{Q} -semipolynomial in the coefficients \mathbf{c} . Moreover, if $\mathbf{c} \in \mathbb{F}_{n,d}$, then $k(\mathbf{c}, \mathbf{X})$ vanishes only on the zeros of $f_{n,d}(\mathbf{c}, \mathbf{X})$, and
- each $p_j(\mathbf{c})$ is a \mathbb{Q} -semipolynomial which is nonnegative on $\mathbb{F}_{n,d}$. Moreover, under the hypothesis $\mathbf{c} \in \mathbb{F}_{n,d}$, the nonnegativity of $p_j(\mathbf{c})$ is ‘clearly’ evident.

A. Prehistory of Theorem II.5

Hilbert posed his 17th problem in 1900 [Hil]; E. Artin solved it in 1927 [Art] by a non-constructive method. In 1955 Artin asked Kreisel whether, from *Artin’s own* proof, one could extract bounds on the number and degrees of the rational functions, in terms of suitable elements of the data (n , d , and possibly \mathbf{c}); Kreisel gave a sketch of such an ‘unwinding’ in [Kre₂]; the bounds so constructed were in terms of n and d (not \mathbf{c}), and primitive recursive; in 1961 Daykin [Day] worked out Kreisel’s sketch, showing that, roughly, the bounds were obtained by applying primitive recursion at least twice to exponential functions of n and d .³ Independently, A. Robinson (see [Rob₁] and [Rob₂]) got (by definition total) general recursive bounds. These authors also expressed the weights and coefficients of the rational functions as \mathbb{Z} -piecewise-polynomial functions g of \mathbf{c} . All this handles the case in which \mathbf{K} is given with computable arithmetic operations and sign test (e.g., when $\mathbf{K} = \mathbb{Q}$)—and it was more than enough for Artin himself.

But the fact that those g were, *a priori*, discontinuous in \mathbf{c} for the usual, order topology on \mathbb{R} , meant that they were computationally inadequate for the case where $\mathbf{K} = \mathbb{R}$, since if we want to have computable arithmetic operations, elements of \mathbb{R} must be given by, say, rational approximations; this makes equality and, *a fortiori*, the order relation, undecidable. Intuitionistic logic gives small changes in the logical laws which do ensure

³ The statement of Daykin’s bounds was oversimplified in [Del₁] and [Del₂].

continuity of functions constructed (and for a wide range of topologies). So after the above contributions by *classical* logic, Kreisel asked in 1962 [Kre₃] whether intuitionistic logic could also contribute, by determining whether the g could be chosen to be continuous, or the rational functions $\in \mathbf{K}(\mathbf{X})$ could be chosen to be continuously extendible to \mathbf{R}^n . So far, intuitionistic logic has contributed little, and real algebraic geometry much:

(a) In 1978 Kreisel noticed [Kre₄] that Stengle’s 1974 “Positivstellensatz” [Ste] (which others have called a “Nichtnegativstellensatz”) easily represents a positive semidefinite (‘psd’) $f \in \mathbf{K}[\mathbf{X}]$ in the form $f = \sum_j p_j (q_j/k)^2$ with $0 \leq p_j \in \mathbf{K}$ and $k, q_j \in \mathbf{K}[\mathbf{X}]$ and, most importantly here, such that the functions q_j/k extend continuously to \mathbf{R}^n (see the end of §II below for details).

(b) In 1980 the first author showed [Del₂] that for even $d \geq 4$, the weights $p_j(\mathbf{c})$ in (II.5)(★) cannot be chosen to be rational functions ($\in \mathbf{R}(\mathbf{c})$); and in 1990 ([Del₆] and [Del₇]) he excluded even (germs of) real analytic functions when $\mathbf{R} = \mathbf{R}$.

(c) Also in 1980 he found his first positive result on this ([Del₃]): for all $d \geq 0$ the p_j and the coefficients of k and the q_j can be chosen to be (locally uniformly) continuous \mathbf{Q} -semialgebraic functions of \mathbf{c} ; functions with these properties *can* be effectively evaluated even over \mathbf{R} , in the—only reasonable—sense that they are computable on, say, the rationals, and they take approximations to approximations. And from the mere existence of this representation of f (and the completeness of the theory of real closed fields), the semialgebraic descriptions of these functions are given by general recursive functions of n and d .

There were two shortcomings in (c): (1) As Scowcroft [Sco₁] and the third author [Lom₂] remarked, it contained (only) one nonconstructive step, namely, the use of Stengle’s theorem, proved up until then by means of Zorn’s lemma. In 1956 Kreisel had observed [Kre₁] that by relativising a proof to the constructible universe, the axiom of choice (and even the generalized continuum hypothesis) can be eliminated from any proof of an *arithmetical* theorem, a fact which has been used to sanitize proofs of results on the order of homotopy groups (by Serre) and on p -adic fields (by Ax and Kochen). Thus, being arithmetic (at least if, say, $\mathbf{K} = \mathbf{Q}$, which was the only case used in (c)), Stengle’s theorem follows from ZF without further ado. It is clear that this ‘trick’ would not satisfy constructivists, however: often there is more to constructivity than purging the axiom of choice, namely purging the principle of the excluded middle. Note also that ZF with intuitionistic logic seems not constructively meaningful. Scowcroft [Sco₁] offered a sketch of a proof of Stengle’s theorem without AC or any other constructively dubious principle; if successful it would yield bounds which belong not only to the set of general recursive functions, but to the proper subset of those which are provably total in formal arithmetic (intuitionistic or classical, both having the same provably total recursive functions). The third author gave a direct, constructive proof of the Pos-

itivstellensatz [Lom₁], with explicit (primitive recursive) bounds [Lom₃]. Thus, by changing only one entry in the bibliography of [Del₃] (namely, from [Ste] to [Lom₁] and [Lom₃]), the proof in [Del₃] becomes constructive, and the semialgebraic descriptions of those functions also become primitive recursive. Thus, if all that we had wanted in a solution to Hilbert's 17th problem over \mathbf{R} were brutal constructivity, then we could have stopped here; in particular, there would have been no need for (II.5) (though the luxury of a simpler proof is welcome).

(2) The other shortcoming of (c), which was not to be overcome by the above kind of tinkering with the proof in [Del₃], was that the functions which it produced are 'only' semialgebraic, and therefore do not necessarily take \mathbf{K} -rational values at \mathbf{K} -rational arguments \mathbf{c} , as Hilbert asked for (unless \mathbf{K} is real-closed). So in the first author's thesis, and in [Del₁], [Del₃], and elsewhere, he proposed to construct continuous, \mathbf{Z} -piecewise-polynomial weights and coefficients; such functions obviously do take values in \mathbf{K} at \mathbf{K} -rational arguments \mathbf{c} , uniformly for all \mathbf{K} . And in [Del₄] he conjectured the full statement of Theorem II.5, that the functions could even be chosen to be \mathbf{Z} -semipolynomials, which would make their continuity 'evident'. This last conjecture/theorem appears to be 'strongest possible' by (b), and it contains the results in (a) and (c) above, as explained in more detail at the end of §II below.

B. History of Theorem II.5

Theorem II.5 is a special case of the Positivstellensatz for semipolynomials ('Pfs'). The first author proved II.5 and the Pfs in 1988 (see the abstracts [Del₅] and [Del₇]). The second and third authors, jointly, rediscovered these results in 1991, independently of the first author, and by a different method. This paper presents their proof of (II.5), along with their view (§IV) of it as being primarily a contribution to constructive mathematics (à la Bishop). Their proof of the Pfs will appear in [GV-L]. The first author's proof of both results will appear in [Del₈], along with his view of them as being primarily a contribution to (classical) topological algebra and to the *examination* of assumptions of the logical tradition, specifically, the assumption that logic contributes to this part of mathematics.

Perhaps the main difference between the two proofs of the Pfs is that the first author reduces it to the abstract Stellensätze for the real spectrum of an arbitrary commutative ring, while the second and third authors reduce it to the Positivstellensatz for polynomials. An advantage of the former proof is that it has some surprising facts about abstract semialgebraic (in particular, abstract 'semipolynomial') functions in f -rings as by-products. For example, while the absolute value function $c_1 \mapsto |c_1|$ is obviously positive semidefinite on \mathbf{R} , it is not psd on the real spectrum of the ring of semipolynomials. This example motivated a large part of N. Schwartz's recent investigation [Sch] into abstract piecewise-polynomial functions. An advantage of the latter proof of (II.5) is that it avoids even

the appearance of reliance on Zorn's lemma, simply by avoiding the abstract Stellensätze. This relies, instead, on—the third author's direct, constructive proof of—the Positivstellensatz for polynomials [Lom₁]; as a result, the functions produced by this proof are primitive recursive in n and d ; as in (A.c.1) and in Kreisel's and Daykin's results at the beginning of subsection A, these primitive recursive functions are of iterated exponential complexity in n and d ,⁴ while those of [Del₈] are only general recursive. Thus the latter proof provides another constructive solution to Hilbert's 17th problem over \mathbb{R} , simpler and more informative than that described in (A.c.1) above (see §IV for details).

The first author's proof of (II.5) was, at first, more complicated than necessary; A. Prestel simplified the proof, and incidentally rearranged it along the lines of the method below by the second and third authors. He also asked in [BP] whether the result extended to higher even powers; cf. the abstract in [Pre₂] when $n = 1$.

The underlying ideas used to prove (II.5) and the Pfs led all of us (again independently) to improve the continuous and semialgebraic variation in Scowcroft's Positivstellensatz [Sco₂] to semipolynomial variation.

Finally, unlike the proofs of (II.5) presented in this paper and in [Del₈], the proof presented in [GV-L] reduces the result to the Pfs.

II. A rational and continuous solution to Hilbert's 17th problem.

First we recall the definitions of strong incompatibility and the general form for the Real Nullstellensatz in the polynomial case (see [Lom₁] and [Lom₃]). We consider an ordered field \mathbf{K} , and \mathbf{X} denotes a list of variables X_1, X_2, \dots, X_n . We then denote by $\mathbf{K}[\mathbf{X}]$ the ring $\mathbf{K}[X_1, X_2, \dots, X_n]$. If F is a finite subset of $\mathbf{K}[\mathbf{X}]$, we let F^{*2} be the set of squares of elements in F , and $\mathcal{M}(F)$ be the *multiplicative monoid generated by* $F \cup \{1\}$. $\mathcal{Cp}(F)$ will be the *positive cone generated by* F (= the additive monoid generated by elements of type pPQ^2 , where $0 \leq p \in \mathbf{K}$, $P \in \mathcal{M}(F)$, and $Q \in \mathbf{K}[\mathbf{X}]$). Finally, let $I(F)$ be the ideal generated by F .

Definition II.1. Consider 4 finite subsets of $\mathbf{K}[\mathbf{X}]$: $F_>, F_{\geq}, F_=:, F_{\neq}$, containing polynomials for which we want respectively the sign conditions > 0 , ≥ 0 , $= 0$, and $\neq 0$: we say that $\mathbf{F} := [F_>, F_{\geq}, F_=:, F_{\neq}]$ is *strongly incompatible in* \mathbf{K} if we have in $\mathbf{K}[\mathbf{X}]$ an equality of the following type:

$$S + P + Z = 0 \quad \text{with} \quad S \in \mathcal{M}(F_> \cup F_{\neq}^{*2}), P \in \mathcal{Cp}(F_{\geq} \cup F_>), Z \in I(F_=).$$

It is clear that a strong incompatibility is a very strong form of incompatibility. In particular, it implies that it is impossible to give the indicated signs to the polynomials considered, in any ordered extension of \mathbf{K} . If one

⁴ However, they involve fewer iterations of the exponential function than Daykin's bounds.

considers the real closure \mathbf{R} of \mathbf{K} , the previous impossibility is testable by Hörmander’s algorithm, for example (see [BCR], chapter 1).

The different variants of the Nullstellensatz in the real case are a consequence of the following general theorem:

Theorem II.2. *Let \mathbf{K} be an ordered field and \mathbf{R} a real closed extension of \mathbf{K} . The three following conditions, concerning a generalized system of sign conditions on polynomials of $\mathbf{K}[\mathbf{X}]$, are equivalent:*

- *strong incompatibility in \mathbf{K} ;*
- *impossibility in \mathbf{R} ; and*
- *impossibility in all the ordered extensions of \mathbf{K} .*

This Nullstellensatz was first proved in 1974 [Ste]. Less general variants were given by Krivine [Kri], Dubois [Du], Prestel [Pre₁], Risler [Ris] and Efraymson [Efr]. All the proofs until [Lom₁] and the sketch in [Sco₁] ‘used’ the axiom of choice (recall (I.A.c.1)).

II.3 Parameterizing Hilbert’s 17th problem.

Let $f_{n,d}(\mathbf{c}, \mathbf{X})$ be the general polynomial of degree d in n variables (\mathbf{c} denotes the list of coefficients c_1, \dots, c_m and \mathbf{X} the list of variables X_1, \dots, X_n). It is a standard fact in real algebraic geometry that the set

$$\mathbf{F}_{n,d} = \{\mathbf{c} \in \mathbf{R}^m : \forall \mathbf{x} \in \mathbf{R}^n \quad f_{n,d}(\mathbf{c}, \mathbf{x}) \geq 0\}$$

is a closed \mathbf{Q} -semialgebraic set. So, applying the Finiteness Theorem, we have that $\mathbf{F}_{n,d}$ is a finite union of ‘basic’ closed \mathbf{Q} -semialgebraic sets:

$$\mathbf{F}_{n,d} = \bigcup_{i=1}^k \bigcap_{j=1}^{n_i} \{\mathbf{c} : R_{n,d,i,j}(\mathbf{c}) \geq 0\}.$$

Here the $R_{n,d,i,j}$ are polynomials in $\mathbf{Z}[\mathbf{c}]$.⁵ The last equation allows us to describe the set $\mathbf{F}_{n,d}$ in the following way:

$$\mathbf{F}_{n,d} = \left\{ \mathbf{c} : \max_{i=1, \dots, k} \left\{ \min \{ R_{n,d,i,j}(\mathbf{c}) : j = 1, \dots, n_i \} \right\} \geq 0 \right\}.$$

So, if for every i in $\{1, \dots, k\}$ we define

$$H_{n,d,i}(\mathbf{c}) = \min_{j=1, \dots, n_i} \{ R_{n,d,i,j}(\mathbf{c}) \}$$

⁵ Some of the published proofs of the Finiteness Theorem (e.g., [Del₁]) explicitly mention the fact that the coefficients of the $R_{n,d,i,j}$ may be chosen to be rational numbers (even integers, after clearing denominators), while others (e.g., [BCR]) assert only that they can be chosen in \mathbf{R} ; but even these other proofs actually do yield the integrality of the coefficients, if one merely pays attention to it. Likewise, in most of the proofs, the authors do not pay attention to the constructive character of their proofs; if one does, one sees that in fact most of the proofs are constructive; [Del₁] does pay attention, and [Sol] goes even further, by giving a careful complexity analysis.

and

$$H_{n,d}(\mathbf{c}) = \max_{i=1,\dots,k} \{H_{n,d,i}(\mathbf{c})\},$$

we have obtained the following description for the set $\mathbb{F}_{n,d}$:

$$\mathbb{F}_{n,d} = \{\mathbf{c} : H_{n,d}(\mathbf{c}) \geq 0\},$$

where $H_{n,d}(\mathbf{c})$ is a \mathbb{Q} -semipolynomial. Therefore we have shown the equivalence

$$\mathbf{c} \in \mathbb{F}_{n,d} \iff H_{n,d}(\mathbf{c}) \geq 0 \iff \forall \mathbf{x} \in \mathbf{R}^n \quad f_{n,d}(\mathbf{c}, \mathbf{x}) \geq 0.$$

II.4 The proof of the parameterized theorem.

The last equivalence allows us to conclude

$$\forall \mathbf{c} \in \mathbf{R}^m \quad \forall \mathbf{x} \in \mathbf{R}^n \quad \{H_{n,d}(\mathbf{c}) \geq 0 \implies f_{n,d}(\mathbf{c}, \mathbf{x}) \geq 0\},$$

or, what is the same, the incompatibility of the system of generalized sign conditions

$$H_{n,d}(\mathbf{c}) \geq 0, \quad f_{n,d}(\mathbf{c}, \mathbf{X}) < 0. \quad (1)$$

If $H_{n,d}(\mathbf{c})$ were a polynomial and not a semipolynomial, then, applying the classical Positivstellensatz (theorem II.2) to the incompatibility (1), we would get an algebraic identity in \mathbf{c} and \mathbf{X} making this incompatibility evident. This would give, for the polynomial $f_{n,d}(\mathbf{c}, \mathbf{X})$, a solution to Hilbert's 17th problem parameterized by polynomials in \mathbf{c} and so, a rational and continuous solution.

As this way of attacking the problem is not feasible, we shall try to translate our incompatible system to another one using only polynomials in order to be able to apply the Positivstellensatz (Theorem II.2). To achieve this goal we introduce new variables z_1, \dots, z_k , and for every $i \in \{1, \dots, k\}$ we consider the following polynomial system of generalized sign conditions:

$$\mathbb{H}_i = \begin{cases} (z_i - R_{n,d,i,1}(\mathbf{c}))(z_i - R_{n,d,i,2}(\mathbf{c})) \cdots (z_i - R_{n,d,i,n_i}(\mathbf{c})) = 0, \\ z_i - R_{n,d,i,1}(\mathbf{c}) \leq 0, \\ z_i - R_{n,d,i,2}(\mathbf{c}) \leq 0, \\ \vdots \\ z_i - R_{n,d,i,n_i}(\mathbf{c}) \leq 0. \end{cases}$$

It is clear from the definitions of $H_{n,d,i}$ and \mathbb{H}_i that, for fixed $\mathbf{c} \in \mathbf{R}^m$, if the system \mathbb{H}_i is verified then $z_i = H_{n,d,i}(\mathbf{c})$.

Next we consider a new variable z and the following polynomial systems of generalized sign conditions:

$$\mathbb{H} = \begin{cases} (z - z_1)(z - z_2) \cdots (z - z_k) = 0, \\ z - z_1 \geq 0, \\ z - z_2 \geq 0, \\ \vdots \\ z - z_k \geq 0. \end{cases} \quad \mathbb{K} = \{\mathbb{H}_1, \dots, \mathbb{H}_k, \mathbb{H}\}.$$

Clearly we have that, for fixed $\mathbf{c} \in \mathbf{R}^m$, if the system \mathbf{K} is verified, then $z = H_{n,d}(\mathbf{c})$.

After introducing in this way the variables z, z_1, \dots, z_k , what we have obtained is the following incompatible system of generalized sign conditions on polynomials in $\mathbf{K}[\mathbf{X}, z, z_1, \dots, z_k]$:

$$\mathbf{K}, z \geq 0, f_{n,d}(\mathbf{c}, \mathbf{X}) < 0.$$

Applying to this system the Positivstellensatz (theorem II.2) and replacing in the equality obtained every z_i by $H_{n,d,i}(\mathbf{c})$ and z by $H_{n,d}(\mathbf{c})$, we obtain an algebraic identity concerning polynomials in \mathbf{X} whose coefficients are \mathbf{Q} -semipolynomials in \mathbf{c} .

Next we study the different parts appearing in this algebraic identity:

- The strictly positive part in the initial identity (before the z_i 's replacement) was $f_{n,d}(\mathbf{c}, \mathbf{X})^{2r}$ and remains unchanged.
- The null part in the initial identity was a polynomial in the ideal of $\mathbf{K}[\mathbf{X}, z, z_1, \dots, z_k]$ generated by the polynomials with ' $= 0$ ' in \mathbf{K} , i.e.,

$$(z_i - R_{n,d,i,1}(\mathbf{c}))(z_i - R_{n,d,i,2}(\mathbf{c})) \cdots (z_i - R_{n,d,i,n_i}(\mathbf{c})) \quad 1 \leq i \leq k$$

$$(z - z_1)(z - z_2) \cdots (z - z_k).$$

After replacing z by $H_{n,d}(\mathbf{c})$ and every z_i by $H_{n,d,i}(\mathbf{c})$, this part becomes a function of \mathbf{c} identically 0 (zero for all $\mathbf{c} \in \mathbf{R}$), and this is the reason why it will not appear in the final identity which concerns polynomials in \mathbf{X} with coefficients \mathbf{Q} -semipolynomials in \mathbf{c} .

- The nonnegative part in the initial identity was a polynomial in the positive cone generated by the polynomials:

$$-f_{n,d}(\mathbf{c}, \mathbf{X}), z, z - z_1, \dots, z - z_k,$$

$$R_{n,d,i,1}(\mathbf{c}) - z_i, \dots, R_{n,d,i,n_i}(\mathbf{c}) - z_i \quad i \in \{1, \dots, k\}.$$

After the replacement only $-f_{n,d}(\mathbf{c}, \mathbf{X})$ remains unchanged, and the other generators of the cone become \mathbf{Q} -semipolynomials in \mathbf{c} which are clearly nonnegative for every \mathbf{c} (by the definition of the functions $H_{n,d,i}(\mathbf{c})$ and $H_{n,d}(\mathbf{c})$ in terms of max and min) or the \mathbf{Q} -semipolynomial $H_{n,d}(\mathbf{c})$ that is nonnegative if $f_{n,d}(\mathbf{c}, \mathbf{x})$ is nonnegative for all $\mathbf{x} \in \mathbf{R}^n$.

Summarizing, we have found an algebraic identity with the following structure:

$$f_{n,d}(\mathbf{c}, \mathbf{X})g(\mathbf{c}, \mathbf{X}) = f_{n,d}(\mathbf{c}, \mathbf{X})^{2r} + h(\mathbf{c}, \mathbf{X}), \quad (2)$$

where $g(\mathbf{c}, \mathbf{X})$ and $h(\mathbf{c}, \mathbf{X})$ are polynomials in \mathbf{X} whose coefficients are \mathbf{Q} -semipolynomials in \mathbf{c} . More precisely, $g(\mathbf{c}, \mathbf{X})$ and $h(\mathbf{c}, \mathbf{X})$ are sum of terms

$$p_j(\mathbf{c})q_j(\mathbf{c}, \mathbf{X})^2,$$

where the $q_j(\mathbf{c}, \mathbf{X})$ are polynomials in \mathbf{X} with coefficients \mathbb{Q} -semipolynomials in \mathbf{c} and the $p_j(\mathbf{c})$ are \mathbb{Q} -semipolynomials nonnegative under the hypothesis $H_{n,d}(\mathbf{c}) \geq 0$. More precisely we have that every $p_j(\mathbf{c})$ is a product whose factors have one of the following type:

- the \mathbb{Q} -semipolynomial $H_{n,d}(\mathbf{c})$,
- a \mathbb{Q} -semipolynomial $H_{n,d}(\mathbf{c}) - H_{n,d,i}(\mathbf{c})$,
- a \mathbb{Q} -semipolynomial $R_{n,d,i,j}(\mathbf{c}) - H_{n,d,i}(\mathbf{c})$, and
- a positive rational or the square of a \mathbb{Q} -semipolynomial in \mathbf{c} .

If we multiply by $f_{n,d}(\mathbf{c}, \mathbf{X})$ every member of equation (2), we get

$$f_{n,d}(\mathbf{c}, \mathbf{X}) = \frac{f_{n,d}(\mathbf{c}, \mathbf{X})^2 g(\mathbf{c}, \mathbf{X})}{f_{n,d}(\mathbf{c}, \mathbf{X})^{2r} + h(\mathbf{c}, \mathbf{X})},$$

and denoting by $k(\mathbf{c}, \mathbf{X})$ the denominator of this fraction, we obtain finally

$$f_{n,d}(\mathbf{c}, \mathbf{X}) = \frac{f_{n,d}(\mathbf{c}, \mathbf{X})^2 g(\mathbf{c}, \mathbf{X}) k(\mathbf{c}, \mathbf{X})}{k(\mathbf{c}, \mathbf{X})^2} = \frac{g_1(\mathbf{c}, \mathbf{X})}{k(\mathbf{c}, \mathbf{X})^2},$$

where $g_1(\mathbf{c}, \mathbf{X})$ is of the same type as $g(\mathbf{c}, \mathbf{X})$ and $h(\mathbf{c}, \mathbf{X})$. Moreover, $k(\mathbf{c}, \mathbf{X})$ vanishes only at the zeros of $f_{n,d}(\mathbf{c}, \mathbf{X})$ if $H_{n,d}(\mathbf{c}) \geq 0$, because then $h(\mathbf{c}, \mathbf{x})$ is positive for all $\mathbf{x} \in \mathbf{R}^n$.

Most of the following theorem has now been proved:

Theorem II.5. *The general polynomial $f_{n,d}$ of degree d in n variables can be written as a weighted sum of squares of rational functions*

$$f_{n,d}(\mathbf{c}, \mathbf{X}) = \sum_j p_j(\mathbf{c}) \left(\frac{q_j(\mathbf{c}, \mathbf{X})}{k(\mathbf{c}, \mathbf{X})} \right)^2, \quad (\star)$$

where

- the $q_j(\mathbf{c}, \mathbf{X})$ and $k(\mathbf{c}, \mathbf{X})$ are polynomials in the variables \mathbf{X} whose coefficients are \mathbb{Q} -semipolynomials in the coefficients \mathbf{c} . Moreover, if $\mathbf{c} \in \mathbb{F}_{n,d}$, then $k(\mathbf{c}, \mathbf{X})$ vanishes only on the zeros of $f_{n,d}(\mathbf{c}, \mathbf{X})$;
- each $p_j(\mathbf{c})$ is a product whose factors are $H_{n,d}(\mathbf{c})$, or one of the \mathbb{Q} -semipolynomials $H_{n,d}(\mathbf{c}) - H_{n,d,i}(\mathbf{c})$, or one of the \mathbb{Q} -semipolynomials $R_{n,d,i,j}(\mathbf{c}) - H_{n,d,i}(\mathbf{c})$, or a positive rational, or the square of a \mathbb{Q} -semipolynomial in \mathbf{c} . So, under the hypothesis $H_{n,d}(\mathbf{c}) \geq 0$, the nonnegativity of $p_j(\mathbf{c})$ is ‘clearly’ evident; and
- the equation

$$f_{n,d}(\mathbf{c}, \mathbf{X}) k(\mathbf{c}, \mathbf{X})^2 - \sum_j p_j(\mathbf{c}) q_j(\mathbf{c}, \mathbf{X})^2 = 0$$

is especially evident in the following sense: the first member of the equality, as polynomial in \mathbf{X} , has as coefficients \mathbb{Q} -semipolynomials in \mathbf{c} which are identically 0 (without assuming $\mathbf{c} \in \mathbb{F}_{n,d}$).

Equation (\star) provides a continuous, rational-valued solution to Hilbert's 17th problem, because

- all the coefficients (the $p_j(\mathbf{c})$ and the \mathbf{X} -coefficients of the $q_j(\mathbf{c}, \mathbf{X})$ and $k(\mathbf{c}, \mathbf{X})$) appearing in equation (\star) are continuous, rational-valued functions of \mathbf{c} ; more precisely, they are \mathbb{Q} -semipolynomials in \mathbf{c} ; and
- every summand in (\star)

$$p_j(\mathbf{c}) \left(\frac{q_j(\mathbf{c}, \mathbf{X})}{k(\mathbf{c}, \mathbf{X})} \right)^2$$

is a function which is rational in \mathbf{X} , and which can be continuously and semialgebraically extended to the closed semialgebraic set $\mathbb{F}_{n,d} \times \mathbf{R}^n$.

Proof:

The only statement not yet proved is the last: the semialgebraicity of the extension of $p_j(q_j/k)^2$ is obvious. To see its continuity, we use an argument of Kreisel [Kre₃] (with parameters \mathbf{c}): note that $k = f^{2r} + h$ (with h nonnegative over $\mathbb{F}_{n,d} \times \mathbf{R}^n$), which can vanish at some point $(\mathbf{c}; \mathbf{x}) \in \mathbb{F}_{n,d} \times \mathbf{R}^n$ only if f vanishes there, forcing each $p_j(q_j/k)^2$ to tend to 0 near $(\mathbf{c}; \mathbf{x})$, by (\star) . The fact that this pointwise continuity is actually locally uniform follows from the corresponding property of f ; we leave the ϵ 's and δ 's to the reader (see [GV-L] for details). ■

III. Rational and continuous solution to other cases of the classical Real Positivstellensatz.

The solution for Hilbert's 17th problem can be seen as a particular case of the Real Positivstellensatz, and for this case we have just proved, in the previous section, the existence of a solution depending on the parameters of the problem in a semipolynomial way. So what we shall do in this section is to generalize this result to other cases.

Let $\mathbb{H}(\mathbf{c}, \mathbf{X})$ be a system of generalized sign conditions on polynomials in $\mathbf{K}[\mathbf{c}, \mathbf{X}]$ where the X_i 's are considered as variables and the c_j 's as parameters. We denote by $\mathbf{S}_{\mathbb{H}}$ the semialgebraic set defined by

$$\mathbf{S}_{\mathbb{H}} = \{ \mathbf{c} : \forall \mathbf{x} \in \mathbf{R}^n \quad \mathbb{H}(\mathbf{c}, \mathbf{x}) \text{ is incompatible} \}.$$

If $\mathbf{S}_{\mathbb{H}}$ is locally closed (i.e., intersection of a closed and an open semialgebraic set), then, applying the Finiteness Theorem (see [BCR] or elsewhere) and the strategy followed in §II when dealing with the set $\mathbb{F}_{n,d}$, it is possible to construct two \mathbf{K} -semipolynomials $H_1(\mathbf{c})$ and $H_2(\mathbf{c})$ satisfying

$$\mathbf{c} \in \mathbf{S}_{\mathbb{H}} \iff \left[H_1(\mathbf{c}) \geq 0, H_2(\mathbf{c}) > 0 \right] \iff$$

$$\iff \forall \mathbf{x} \in \mathbf{R}^n \quad \mathbf{H}(\mathbf{c}, \mathbf{x}) \text{ is incompatible.}$$

We have obtained the incompatible system of sign conditions

$$\left[H_1(\mathbf{c}) \geq 0, H_2(\mathbf{c}) > 0, \mathbf{H}(\mathbf{c}, \mathbf{X}) \right],$$

but with $H_1(\mathbf{c})$ and $H_2(\mathbf{c})$ \mathbf{K} -semipolynomials. Now we proceed in the same way as in §II: we consider new variables $z_1, z_{1,i}$ ($i \in \{1, \dots, k_1\}$), z_2 and $z_{2,i}$ ($i \in \{1, \dots, k_2\}$), which are used to construct a system of polynomial sign conditions $\mathbf{K}(\mathbf{c})$ translating the definition of $H_1(\mathbf{c})$ and $H_2(\mathbf{c})$ as semipolynomials.

As the incompatible system of generalized sign conditions

$$\{\mathbf{K}(\mathbf{c}), z_1 \geq 0, z_2 > 0, \mathbf{H}(\mathbf{c}, \mathbf{X})\}$$

involves only polynomials, we can apply the Positivstellensatz (theorem II.2), obtaining an algebraic identity making this incompatibility evident. Finally, we replace in the algebraic identity obtained the variables $z_1, z_{1,i}$ ($i \in \{1, \dots, k_1\}$), z_2 and $z_{2,i}$ ($i \in \{1, \dots, k_2\}$) by the \mathbf{K} -semipolynomials they are representing. So we have obtained an algebraic identity concerning polynomials in the variables \mathbf{X} with coefficients \mathbf{K} -semipolynomials in \mathbf{c} .

The next theorem summarizes the results obtained in this section and provides a rational and continuous solution for some cases of the Real Positivstellensatz.

Theorem III.1. *Let $\mathbf{H}(\mathbf{c}, \mathbf{X})$ be a system of generalized sign conditions on polynomials in $\mathbf{K}[\mathbf{c}, \mathbf{X}]$, where the X_i 's are considered as variables and the c_j 's as parameters. If $\mathbf{S}_{\mathbf{H}}$ is the semialgebraic set defined by*

$$\mathbf{c} \in \mathbf{S}_{\mathbf{H}} \iff \forall \mathbf{x} \in \mathbf{R}^n \quad \mathbf{H}(\mathbf{c}, \mathbf{x}) \text{ is incompatible,}$$

and if $\mathbf{S}_{\mathbf{H}}$ is locally closed, then (Finiteness Theorem) there exist \mathbf{K} -semipolynomials $H_1(\mathbf{c})$ and $H_2(\mathbf{c})$ such that

$$\mathbf{c} \in \mathbf{S}_{\mathbf{H}} \iff \left[H_1(\mathbf{c}) \geq 0, H_2(\mathbf{c}) > 0 \right].$$

If $\mathbf{c} \in \mathbf{S}_{\mathbf{H}}$, then the incompatibility of $\mathbf{H}(\mathbf{X}) := \mathbf{H}(\mathbf{c}, \mathbf{X})$ inside \mathbf{R}^n is made obvious by a strong incompatibility of fixed type (independent of \mathbf{c}) and with coefficients given by \mathbf{K} -semipolynomials in \mathbf{c} . Moreover,

- the algebraic identity obtained, seen as a polynomial in \mathbf{X} , has an especially simple structure. More precisely, every \mathbf{X} -coefficient of this identity is a \mathbf{K} -semipolynomial in \mathbf{c} identically 0 (without assuming $H_1(\mathbf{c}) \geq 0$ and $H_2(\mathbf{c}) > 0$), and
- every coefficient $p(\mathbf{c})$ in the algebraic identity which must be nonnegative (resp. positive) is given by a \mathbf{K} -semipolynomial showing such

character in an especially clear way under the hypothesis $H_1(\mathbf{c}) \geq 0$ and $H_2(\mathbf{c}) > 0$.

In the same way that our rational and continuous solution for Hilbert's 17th problem (§II) improves the first author's result [Del₃], (III.1) improves Scowcroft's results [Sco₂] in four respects:

- a-. for us, the semialgebraic set $\mathbf{S}_{\mathbf{H}}$ need not be closed (Scowcroft knew this, but chose to use hypotheses involving the logical form of the implications implicit in \mathbf{H} , rather than topological hypotheses on $\mathbf{S}_{\mathbf{H}}$);
- b-. the coefficients of our solution are continuous, rational-valued functions (more precisely, \mathbf{K} -semipolynomials) in the parameters \mathbf{c} and not only continuous semialgebraic as in [Sco₂];
- c-. the algebraic identity obtained, seen as a polynomial in \mathbf{X} , has an especially simple structure: its coefficients are \mathbf{K} -semipolynomials in \mathbf{c} identically 0 (without assuming $H_1(\mathbf{c}) \geq 0$ and $H_2(\mathbf{c}) > 0$); and
- d-. the nonnegativity or positivity of those coefficients in the solution which must satisfy such conditions, is clearly evident under the hypothesis $H_1(\mathbf{c}) \geq 0$ and $H_2(\mathbf{c}) > 0$.

Finally, we note a strong converse of (III.1): the hypothesis that $\mathbf{S}_{\mathbf{H}}$ be locally closed is also *necessary* for the existence of a semipolynomially (or even continuously) varying Positivstellensatz; the proof is in [Del₈].

IV. Conclusion: the constructive content of the results.

In constructive mathematics (see [BB] and [MR_R]), the theorems presented in §§II and III are valid when the parameters \mathbf{c} take values in an ordered discrete field [LR], because in this setting we have a constructive proof of the Positivstellensatz ([Lom₁]).

An interpretation of the results admissible for everybody is the following one: all our proofs are effective, in particular without using the axiom of choice and, more precisely, providing uniformly primitive recursive algorithms if the structure of the field of parameters is given by an oracle giving the sign of every polynomial with integer coefficients on the parameters of the problem.

The only thing remaining to be mentioned is the constructive content of the results when dealing with the field \mathbf{R} of the real numbers in constructive analysis [BB], i.e., the real numbers defined as—equivalence classes of—Cauchy sequences of rational numbers. From an algorithmic point of view this means that the real parameters \mathbf{c} are given by oracles providing suitable rational approximations (depending on the request made to the oracle) of the real numbers involved, and that we are looking for a uniformly primitive recursive algorithm. More details on this question can be found in [GV-L].

The answer to Hilbert's 17th problem provided by theorem II.5 uses polynomials and semipolynomials with coefficients in \mathbf{Q} that can be computed explicitly. The nonnegativity of the weights is clear from a constructive point of view when dealing with real numbers “à la Cauchy” under

the hypothesis $H_{n,d}(\mathbf{c}) \geq 0$. This implies that, if the parameters \mathbf{c} satisfy the condition $H_{n,d}(\mathbf{c}) \geq 0$, then the polynomial $f_{n,d}(\mathbf{c}, \mathbf{X})$ is everywhere nonnegative. So, for the polynomial $f_{n,d}(\mathbf{c}, \mathbf{X})$, Hilbert's 17th problem is solved in a continuous, rational-valued way with respect to its coefficients. Moreover, since we can constructively prove (see [Lom₂] or [GV-L]) the converse

$$\forall \mathbf{x} \in \mathbb{R}^n \quad f_{n,d}(\mathbf{c}, \mathbf{x}) \geq 0 \implies H_{n,d}(\mathbf{c}) \geq 0,$$

we can conclude that also this continuous, rational-valued solution for Hilbert's 17th problem is complete and constructive for the field \mathbb{R} .

References.

- [Art] Artin E.: *Über die Zerlegung definitiver Funktionen in Quadrate*. Abh. Math. Sem. Hamburg, **5** (1927), 100–115.
- [BB] Bishop E., Bridges D.: *Constructive Analysis*. Springer-Verlag (1985).
- [BCR] Bochnak J., Coste M. and Roy M.-F.: *Géométrie Algébrique Réelle*. Ergebnisse vol. **12**, Springer-Verlag (1987).
- [BP] Bradley M., Prestel A.: *Representation of a real polynomial $f(X)$ as a sum of $2m$ -th powers of rational functions*. Ordered Algebraic Structures, J. Martinez, ed., Kluwer (1989), 197–207.
- [Day] Daykin D.: *Hilbert's 17th problem*. Ph.D. Thesis, unpublished (1961).
- [Del₁] Delzell C.N.: *A finiteness theorem for open semialgebraic sets, with applications to Hilbert's 17th problem*. Ordered Fields and Real Algebraic Geometry, Contemp. Math. **8**, AMS, D. Dubois and T. Recio, (eds.), Providence 1982, 79–97.
- [Del₂] Delzell C.N.: *Case distinctions are necessary for representing polynomials as sums of squares*. Proc. Herbrand Symp., Logic Coll. 1981, J. Stern (ed.), Amsterdam-Oxford-New York 1982, 87–103.
- [Del₃] Delzell C.N.: *A continuous, constructive solution to Hilbert's 17th problem*. Inventiones Mathematicae **76** (1984), 365–384.
- [Del₄] Delzell C.N.: *On the Pierce-Birkhoff conjecture over ordered fields*. Rocky Mountain Journal of Mathematics **19**(3) (Summer 1989), 651–668.
- [Del₅] Delzell C.N.: *A sup-inf-polynomially varying solution to Hilbert's 17th problem*. AMS Abstracts **10**(3), (Issue 63, April 1989), 208–209, #849-14-160.
- [Del₆] Delzell C.N.: *On analytically varying solutions to Hilbert's 17th problem*. Submitted to Proc. Special Year in Real Algebraic Geometry and Quadratic Forms at UC Berkeley, 1990–1991, (W. Jacob, T.-Y. Lam, R. Robson, eds.), Contemporary Mathematics.
- [Del₇] Delzell C.N.: *On analytically varying solutions to Hilbert's 17th problem*. AMS Abstracts **12**(1), (Issue 73, January 1991), page 47, #863-14-743.

- [Del₈] Delzell C.N.: *Continuous, piecewise-polynomial functions which solve Hilbert's 17th problem*. Submitted for publication (1992).
- [Du] Dubois D.W.: *A nullstellensatz for ordered fields*. Arkiv for Mat. **8** (1969), 111–114.
- [Efr] Efroymsen G.: *Local reality on algebraic varieties*. Journal of Algebra **29** (1974), 113–142.
- [GV-L] González-Vega L., Lombardi H.: *A Real Nullstellensatz and Positivstellensatz for the Semipolynomials over an Ordered Field*. Submitted to the Journal of Pure and Applied Algebra (1992).
- [Hil] Hilbert D.: *Mathematische Probleme*. Göttinger Nachrichten (1900), 253–297, and Archiv der Math. u. Physik (3rd ser.) **1** (1901), 44–53, 213–237.
- [Kre₁] Kreisel G.: *Some uses of metamathematics*. British J. Phil. Sci. **7**(26) (August 1956), 161–173.
- [Kre₂] Kreisel G.: *Sums of squares*. Summaries of Talks Presented at the Summer Institute in Symbolic Logic in 1957 at Cornell Univ., Institute Defense Analyses, Princeton (1960), 313–320.
- [Kre₃] Kreisel G.: *Review of Goodstein*, MR **24A**, 336–7, #A1821 (1962).
- [Kre₄] Kreisel G.: *Review of Ershov*, Zbl. **374**, 02027 (1978).
- [Kri] Krivine J.L.: *Anneaux préordonnés*. Journal d'Analyse Mathématique **12** (1964), 307–326.
- [Lom₁] Lombardi H.: *Effective real Nullstellensatz and variants*. Effective Methods in Algebraic Geometry. Editors T. Mora and C. Traverso. Progress in Mathematics **94** (1991), 263–288, Birkhauser. Detailed French version in *Théorème effectif des zéros réel et variantes (avec une majoration explicite des degrés)*, Memoire d'habilitation (1990).
- [Lom₂] Lombardi H.: *Une étude historique sur les problèmes d'effectivité en algèbre réelle*. Memoire d'habilitation (1990).
- [Lom₃] Lombardi H.: *Une borne sur les degrés pour le théorème des zéros réel effectif*. Proceedings of the conference on Real Algebraic Geometry held in La Turballe (1991). To appear as Lecture Notes in Mathematics.
- [LR] Lombardi H., Roy M.-F.: *Théorie constructive élémentaire des corps ordonnés*. Publications Mathematiques de Besançon, Théorie de Nombres, 1990–1991. English version in: *Constructive elementary theory of ordered fields*, Effective Methods in Algebraic Geometry. Editors T. Mora and C. Traverso. Progress in Mathematics **94** (1991), 249–262, Birkhauser.
- [MRR] Mines R., Richman F., Ruitenburg W.: *A Course in Constructive Algebra*. Universitext, Springer-Verlag (1988).
- [Pre₁] Prestel A.: *Lectures on Formally Real Fields*. IMPA Lecture Notes **22**, Rio de Janeiro 1975; reprinted in Lecture Notes in Mathematics **1093**

- (1983), Springer-Verlag.
- [Pre₂] Prestel A.: *Continuous representations of real polynomials as sums of 2^m -th powers*. Math. Forschungsinst. Oberwolfach Tagungsbericht (Reelle algebraische Geometrie, June 10–16, 1990) **25** (1990), 18.
- [Ris] Risler J.-J.: *Une caractérisation des idéaux des variétés algébriques réelles*. C.R.A.S. Paris, Série A, **271** (1970), 1171–1173.
- [Rob₁] Robinson A.: *On ordered fields and definite forms*. Math. Ann. **130** (1955), 257–271.
- [Rob₂] Robinson A.: *Further remarks on ordered fields and definite forms*. Math. Ann. **130** (1956), 405–409.
- [Sch] Schwartz N.: *Piecewise-polynomial functions*. Submitted.
- [Sco₁] Scowcroft P.: *A transfer theorem in constructive real algebra*. Ann. Pure and Appl. Logic **40**(1), 29–87 (1988).
- [Sco₂] Scowcroft P.: *Some continuous Positivstellensätze*. Journal of Algebra **124** (1989), 521–532.
- [Sol] Solernó P.: *Effective Lojasiewicz inequalities in semialgebraic geometry*. Applicable Algebra in Engineering, Communication and Computing **2**(1) (1991), 1–14.
- [Ste] Stengle G.: *A Nullstellensatz and a Positivstellensatz in semialgebraic geometry*. Math. Ann. **207** (1974), 87–97.

Charles N. Delzell
 Dept. Mathematics
 Louisiana State University
 Baton Rouge, LA 70803, U.S.A.
 (delzell@lsuvax.sncc.lsu.edu)

Laureano González-Vega
 Dept. Matemáticas
 Universidad de Cantabria
 Santander 39071, Spain
 (g_vega@ccucvx.unican.es)

Henri Lombardi
 Lab. de Mathématiques
 Université de Franche-Comté
 Besançon 25030, France
 (hl@math.univ-fcomte.fr)