Positive polynomials - Hilbert's 17th problem

Safdar Quddus
B.Math. Hons. IInd yr
Indian Statistical Institute
Bangalore.

This work was done as a part of a KVPY Project under the guidance of Professor B.Sury.

A famous theorem by Lagrange says that any positive integer a is expressible as a sum of four squares of integers. Positivity is a key concept implicitly used here. Thus, one could analyse questions of this nature over real numbers where there is a notion of positivity. For example, let us consider real polynomials possibly in more than one variable. We write $f \in \mathbb{R}[x_1, \cdots, x_n]$ to mean that it is a real polynomial in n variables. The analogue of the hypothesis that $a \geq 0$ in positive integers, would then be all f such that the function $f > 0$ for all values of $x_1, x_2, \cdots, x_n \in R$. Let us call such f as *positive semi-definite* (henceforth written psd for brevity). The 21-year old Minkowski presenting his Inaugural Dissertation in July 1885 on quadratic forms made the bold conjecture that there must exist homogeneous, real, psd polynomials of any degree > 2 in $n > 2$ variables which are *not* sums of squares of homogeneous real polynomials. At the public defense of this Dissertation, it was the task of Hilbert to attack it but the defense ended with Hilbert declaring that he "was convinced by Minkowski's exposition that already for $n = 3$ there may well be such remarkable forms, which are so stubborn as to remain positive without allowing themselves to submit to a representation as sums of squares of forms." In 1888, Hilbert proved Minkowski's 'conjecture' giving examples where a psd real polynomial f cannot be written as a sum of squares of polynomials. He studied it further and considered the problem of representing any psd $f \in \mathbb{R}[x_1, \cdots, x_n]$ as sums of squares of rational functions (elements of $\mathbb{R}(x_1, \cdots, x_n)$). In 1893, he proved that this does happen for $n = 2$ (this corresponds to the 3-variable homogeneous case of Minkowski's conjecture). In 1899, he proved the remarkable 'result' that, any segment of length $f(x_1, \cdots, x_n)$ which can be constructed from given lengths $x_1, \cdots, x_n$ by using a ruler and compass, can already be constructed without

a compass, provided $f(y_1 \cdots, y_n)$ is a totally real algebraic number for any $y_1, \cdots, y_n \in \mathbb{Q}$. His 'proof' of this result required the truth of the (at that time) unproved assertion that any psd rational function in $\mathbb{Q}(x_1, \cdots, x_n)$ is a sum of squares of rational functions in $\mathbb{Q}(x_1, \cdots, x_n)$. This was a further motivation for him to formulate in his famous 1900 address the 17th problem which is the question :
If $f \in \mathbb{R}[x_1, x_2....x_n]$ is psd, is it necessarily a sum of squares of rational functions in $\mathbb{R}(x_1, x_2, \cdots, x_n)$?
Note that, if instead of $\mathbb{R}$, we consider $\mathbb{C}$, which does not have an order, then we see that *every* polynomial in $\mathbb{C}[x_1, x_2....x_n]$ is a sum of squares of rational functions. The 17th problem was solved by E.Artin in 1926 in the affirmative. He proved it as an existence theorem. His remarkable proof opened up the new subject of model theory and nowadays his argument is viewed as a special case of Tarski's transfer principle. The proof also brought into the forefront the *real spectrum* of rings such as $\mathbb{R}[x_1, \cdots, x_n]$ and started the subject of real algebraic geometry. The proof works for more general 'real-closed' fields (to be defined below) and such fields were studied by Artin and Schreier in a paper in the same volume where Artin's solution of Hilbert's 17th problem was published. Indeed, the Artin-Schreier paper is from page 85 to 99 and is followed by Artin's paper from page 100 to 115.
We just mention a few words about the connection of Artin-Schreier's 1926 work with mathematical logic. Tarski proved later in 1948 that the theory of real-closed ordered fields admits 'quantifier elimination' in the language of ordered rings. That is, every 'formula' is equivalent to a quantifier-free formula. Thus, the theory of real-closed fields is 'model-complete.' In other words, if $E \subseteq F$ are real-closed fields, then an 'elementary sentence' about ordered fields with parameters in E, holds good in F if and only if it holds in E. This is crystallized as Tarski's transfer principle which asserts that every elementary sentence about ordered fields which holds in $\mathbb{R}$ also holds in every real-closed field.

The basic idea of Artin's theorem can be described roughly as follows. Although $\mathbb{R}$ (or more generally, any 'real-closed field') has a unique notion of positive elements, the function field $\mathbb{R}(x_1, \cdots, x_n)$ has several possible notions of positivity. If f is not a sum of squares in $\mathbb{R}(x_1, \cdots, x_n)$, then there would be some ordering (equivalently, a concept of positive elements) on $\mathbb{R}(x_1, \cdots, x_n)$, under which f would be negative. Then, there would be a 'specialisation' $f(a_1, \cdots, a_n)$ which would be negative. This last step is the

key one and is a form of the Tarski transfer principle. Our aim is to discuss Artin's proof in detail and the discussion will avoid the language of model theory etc. Some properties of polynomials over the real field like the intermediate value theorem and others were proved by Sturm in the 1880's and go through for all real-closed fields. These are used by Artin in his proof. For a more model-theoretic treatment of Artin's theorem, see $[DP]$. Our proof is based on the discussions in $[J]$ and $[N]$. See $[L]$ for a slightly different proof.

Hilbert's 17th problem for $n = 1$:
The proof for $n = 1$ is really simple and we give it now. In this case, in fact, we even have polynomials g and h such that every psd $f \in \mathbb{R}[x]$ can be written as $f = g^2 + h^2$.
Proof :
We first notice that every real root (if at all it exists) occurs with even multiplicity . To see this write: $f(x) = (x - \alpha)^n g(x)$ where $\alpha \in \mathbb{R}$ and $g(\alpha) \neq 0$. So if n is odd, then $(x - \alpha)^n$ and hence f would change sign in the neighbourhood of α, which contradicts the fact that f is psd. Hence each real root occur in with even multiplicity and we may write:

$$f(x) = c\prod(x - \alpha_i)^{2n_i}\prod(x - \beta_i)^{n_i}\prod(x - \bar{\beta_i})^{n_i}$$

where $\alpha_i \in \mathbb{R}$ and $\beta_i \notin \mathbb{R}, c \in \mathbb{R}$ Let

$$h(x) = \prod(x - \beta_i)^{n_i}\prod(x - \alpha_i)^{n_i}.$$

So, $f(x) = ch(x)h(\bar{x})$; Now, $h(x) = p(x) + iq(x)$ for $p, q \in \mathbb{R}$ So, $f(x) = c(p(x)^2 + q(x)^2) = (\sqrt{c}p(x))^2 + (\sqrt{c}q(x))^2$.

Necessity of rational functions - an example :
Here is an example to show that for general n, it is not sufficient to work with polynomials. Consider

$$f(x, y) = (x^2 + y^2 - 3)x^2y^2 + 1 \in \mathbb{R}[x, y].$$

Clearly f is psd, as $\frac{x^2+y^2+\frac{1}{x^2y^2}}{3} \geq 1$. Suppose, if possible, $f = f_1^2 + f_2^2 + \cdots + f_n^2$ with $f_i \in \mathbb{R}[x, y]$. For all $i \leq n$, deg $f_i \leq 3$, as the total degree of f is 6.
Since $f(x, 0) = f(0, y) = 1$, each $f_i(x, 0)$ and $f_i(0, y)$ is a constant. Thus $f_i(x, y) = a_i + xy(b_i + c_ix + d_iy) \forall 1 \leq i \leq n$. Now $f = \Sigma f_i^2$; so the coefficient of x^2y^2 in Σf_i^2 is -3. Hence Σb_i^2 = -3, which is a contradiction .

Hilbert's identities :
Let us contrast the above situation with the following identities involving higher powers over $\mathbb{R}$. We have :

$$(x^2+y^2)^3 = \frac{4}{5}(x^6 + (\frac{x+y}{\sqrt{2}})^6 + y^6 + (\frac{y-x}{\sqrt{2}})^6).$$

More generally , one has the remarkable identities due to B.Reznik :

$$(x^2+y^2)^s = \frac{2^{2s}}{v\binom{2s}{s}} \sum_{j=0}^{v-1} \{Cos(j\pi x/v) + Sin(j\pi y/v)\}^{2s}.$$

Interestingly, the analogues of the above identities over $\mathbb{Q}$ are unknown in explicit form (this is known as the champagne problem) but the existence of such identities over $\mathbb{Q}$ are already proved by Hilbert. In fact, in 1981-82 E.Becker proved the existence of identities of the form

$$(x_1^{2k} + \cdots + x_n^{2k})^l = f_1^{2kl} + \cdots + f_m^{2kl}$$

for some rational functions f_i over $\mathbb{Q}$.

Higher powers - examples :
Consider $f(x) = x^4 + nx^2 + 1$. It is a fact that it is a sum of 4-th powers in $\mathbb{R}(x)$. On the other hand we claim that f is not a sum of squares of real polynomials if n is large. To see this, suppose, if possible, that $f = \sum_{i=1}^{N}(a_i + b_i x)^4$. Comparing the coefficients of x^2 we have

$$n = 6\sum_{i=1}^{N} a_i^2 b_i^2.$$

Also, we have $1 = \sum_{i=1}^{N} a_i^4$ and $1 = \sum_{i=1}^{N} b_i^4$. Using the Cauchy-Schwarz inequality now, we see immediately that

$$n \leq 6\sum_{i=1}^{N} a_i^4 \sum_{i=1}^{N} b_i^4 = 6.$$

We now proceed towards the discussion of Artin's famous theorem which answers Hilbert's 17^{th} problem affirmatively. The method he used was developed together with Schreier and answers the question for more general fields than $\mathbb{R}$. We proceed to define these more general fields now.

Definitions and remarks :

(I) A field K is *formally real* if $\sum_{i=1}^{n} a_i^2 = 0$ implies that $a_i = 0$ for all $1 \leq i \leq n$. Equivalently, -1 cannot be written as sum of squares in K. As one can see immediately, $\mathbb{R}$, $\mathbb{Q}$ are formally real while $\mathbb{C}$ is not. Another example is $\mathbb{R}(x)$.
Note that a formally field must have characteristic zero because if it were a prime p, then $1^2 + \cdots + 1^2 = 0$ where the sum has p terms.
We shall see below that formally fields are exactly the class of fields for which there is a notion of positivity or ordering. This result of Artin and Schreier needs the notion of real-closed fields which is defined in what follows now.
(II) A formally real field K is said to be *real-closed* if no finite extension of K (other than itself) is formally real. Equivalently, if a formally real field has the properties : (i) the square roots of positive elements exist, and (ii) every odd degree polynomial over it has a root in it, then it is real-closed.
It is a very interesting fact that if the algebraic closure $\bar{R}$ of a field R is a finite, proper extension, then R is real-closed and $\bar{R} = R(\sqrt{-1})$. One can prove this by using the characterisation above.
(III) A subset P of a field K which is closed under the addition and the multiplication of K is called an *ordering* or a *positive cone* if $K = P \cup -P$ and $P \cap (-P) = \{0\}$. If such a P exists, then K is said to be an *ordered field* and that P and $-P$ are respectively the sets of all the positive and all the negative elements. Note that as each a in K is either positive or negative, its square a^2 is in $P.P \subset P$ or $(-P).(-P) = P.P \subset P$. Hence, nonzero squares are in P and, therefore, $\sum K^2$ (defined as the set of finite sums of squares) is a subset of P. Note K is obviously formally real because $1 \in \sum K^2 \subset P$ (and so $-1 \notin P$).
Moreover, clearly $P - \{0\}$ is a subgroup of index 2 in K^*. Conversely, for a field K, if K^* has a subgroup Q of index 2 which is additively closed, then $Q \cup \{0\}$ is an ordering in K.
Given an order P, one can define $a \geq b$ if $a - b \in P$.
(IV) An element a of a field K is said to be *totally positive* if $a \in \cap P_i$, for all orderings P_i of K.
(V) One defines Zariski open sets as the complements in $\mathbb{R}^n$ of the sets of roots of the polynomials in $\mathbb{R}[x_1, x_2, ..., x_n]$.

Examples of orderings :
(i) For any subfield K of $\mathbb{R}$, take $P = \{x \in K : x \geq 0\}$.

(ii) For $K = \mathbb{Q}(x)$, take $P = \{f/g : f(\pi)/g(\pi) \geq 0\}$. Notice that we could have taken any transcendental number instead of π as well.
(iii) For $K = \mathbb{R}(x)$, take $P = \{f/g : a > 0$ where $fg = ax^l +$ higher powers $\}$. In this ordering, note that x is infinitely smaller than every positive real number !
(iv) For $K = \mathbb{R}(x)$, take $P = \{f/g :$ top coefficient of fg is $> 0\} \cup \{0\}$. In this case, note that x is infinitely larger than every real number !

Lemma 1 :
(i) *If a field K is not formally real and has characteristic different from 2, then every element of K is a sum of squares.*
(ii) *Any formally real field K is an ordered field and conversely.*
(iii) *A real-closed field R has a unique ordering. Indeed, in this case $P = \sum R^2 = R^2$ (squares). Moreover, any automorphism f R is order-preserving.*
Proof :
(i) If $-1 = \sum a_i^2$, then any $t = ((1+t)/2)^2 + \sum (a_i(1-t)/2)^2$.
(ii) As we already saw in (III) above that an ordered field (with an ordering P) is formally real because all squares are in P and, in particular, $1 \in P$ and so $-1 \notin P$.
Conversely, let us suppose that K is a formally real field. Consider $S = \sum K^2$. We know $-1 \notin S$. Clearly, S is closed under addition and multiplication. By Zorn's lemma, there is a subset P of K containing S, closed under addition and multiplication, $-1 \notin P$ and P is maximal with respect to these properties. To show now that P is an ordering, we need only check that $P \cup -P = K$. Note that the other property $P \cap (-P) = \{0\}$ is automatically true; otherwise, $x = -y \in P \cap (-P)$ implies that $-1 = x/y \in P$ unless $x = y = 0$.
Let $x \in K, x \notin P$. Observing

$$(P - xP).(P - xP) \subseteq P - xP + x^2P \subseteq P - xP$$

we note that $-1 \notin P - xP$; otherwise, writing $-1 = p_1 - xp_2$, we have

$$x = (1 + p_1)p_2 \, \frac{1}{p_2^2} \in P,$$

a contradiction. Thus, $P - xP$ contains P and has all those properties which P has and with respect to which P was chosen to be maximal. This forces $P - xP = P$. Hence $-x \in P$; in other words, $K = P \cup (-P)$. So, P is an ordering on K. Thus (ii) is proved.

(iii) Of course $R^2 \subseteq P$ for any ordering P as $a^2 = (-a)^2 \subseteq P.P \subseteq P$. Conversely, given any P, if $a \in P$ were not in R^2, then $R(\sqrt{a})$ cannot be formally real. Writing $-1 = \sum(b_i + c_i\sqrt{a})^2$, we get $-a \in \sum R^2 \subseteq P$. Thus, $a \in P \cap (-P)$, a contradiction. So $R^2 = P$ is the unique ordering. Note that $R^2 \subseteq \sum R^2 \subseteq P$ implies all are equal. Finally, any automorphism of R preserves $R^2 = P$, and thus is order-preserving.

Now we can state Artin's theorem which solves Hilbert's 17th problem affirmatively.

Theorem (Artin) :
Let $\mathbb{R}$ be any real closed field and $f \in \mathbb{R}[x_1, x_2, ..., x_n]$. If f is psd (positive semi-definite), then $f = \Sigma g_i^2$ for some $g_i \in \mathbb{R}(x_1, x_2, ..., x_n)$.

The proof depends on many other results of independent interest as we shall see.

Proposition 1 (characterisation of squares) :
Let K be a field of characteristic $\neq 2$. Then, an element $a \in K^*$ is a sum of squares in $K \Leftrightarrow a > 0$ is totally positive.

An example :
Before proving the above proposition, let us discuss it for the field $F = \mathbb{Q}[\sqrt{2}]$. First, we claim that the only orders of F are:

$$P_1 = \{a + b\sqrt{2} | a + b\sqrt{2} \geq 0\}$$

and

$$P_2 = \{a + b\sqrt{2} | a - b\sqrt{2} \geq 0\}.$$

If P is an order then $\mathbb{Q}_{>0} \subset$ and $\sqrt{2} \in P$ or $-P$. If $\sqrt{2} \in P$, then $P_1 \subset P$, which implies that $\dot{P} = \dot{P_1}$ since both have index 2 in F^*. Hence $P = P_1$. If -$\sqrt{2} \in P$, then a similar argument shows that $P = P_2$. Now $P_1 \cap P_2 = \{a + b\sqrt{2} | a \geq 0, a^2 \geq 2b^2\}$. Now, $P_1 \cap P_2 = \Sigma F^2$ as, for any $c, d \in \mathbb{Q}$, we have $(c + d\sqrt{2})^2 = c^2 + 2d^2 + 2cd\sqrt{2} \in P_1 \cap P_2$. Hence $\Sigma F^2 \subset P_1 \cap P_2$.
Now suppose $a + b\sqrt{2} \in P_1 \cap P_2$. If $b = 0$, then $a \geq 0$ is in $\mathbb{Q}$; hence $a \in \Sigma\mathbb{Q}^2 \subset \Sigma F^2$.
If $b < 0$ and if we can write $a - b\sqrt{2}$ as a sum of squares, then by taking "conjugates" we can see that $a + b\sqrt{2}$ is also a sum of squares. So, we may

assume $b > 0$. Consider the square

$$(x + \frac{b\sqrt{2}}{2x})^2 = x^2 + \frac{b^2}{2x^2} + b\sqrt{2}.$$

Now $x^2 + \frac{b^2}{2x^2}$ attains its minimum at $x = \frac{b}{\sqrt{2}}$ and, this minimum value is $b\sqrt{2} \leq a$.
Hence we can find a rational q such that $q^2 + \frac{b^2}{2q^2} \leq a$. Then

$$a + b\sqrt{2} = (q + \frac{b\sqrt{2}}{2q})^2 + (a - q^2 - \frac{b^2}{2q^2}) \in \Sigma F^2.$$

So, we have checked the proposition in case of the field $\mathbb{Q}[\sqrt{2}]$.

Proof of proposition 1 :
If $a = \Sigma a_i^2; a \neq 0$, then $a > 0$ for all orderings of R. Conversely, assume that $a \neq 0$ is not a sum of squares in R. Let $\bar{R}$ be an algebraic closure of R and consider the set of subfields E of $\bar{R}$ in which a is not a sum of squares. By Zorn's Lemma, there is a maximal element (say F) such that F is formally real. We have used the fact that if a field is not formally real and has characteristic different from 2, then every element of K is a sum of squares.
We claim that $-a$ is a square in F. For, otherwise the subfield $F[\sqrt{-a}]$ of $\bar{R}$ properly contains R, and so a is a sum of squares in $F[\sqrt{-a}]$. Hence, we have b_i and $c_i \in F$, such that $a = \Sigma(b_i + c_i\sqrt{-a})^2$. This gives $\Sigma b_i c_i = 0$ and $a = \Sigma b_i^2 - a\Sigma c_i^2$. Therefore, $a(1 + \Sigma c_i^2) = \Sigma b_i^2$ and $1 + \Sigma c_i^2 \neq 0$ (since F is formally real). Then if $c = 1 + \Sigma c_i^2$; $a = \Sigma b_i^2 c^{-1} = \Sigma b_i^2(1 + \Sigma c_i^2)c^{-2}$. So, a is a sum of squares in F, contrary to the definition of F. Hence $-a = b^2; b \in F$, and so $a = -b^2$ is negative in an ordering of F. So $a \notin \cap P_i$, where P_i are all the orders.

Real-closures and Sturm's work :

A *real-closure* of an ordered field K is an algebraic extension L which is real-closed and its (unique) ordering extends the given one on K.
A key point is that a real-closure not only exists but it is unique in a sense to made clear below. These results follow from ideas due to Sturm from the 1880's which enables us determine the number of roots of a polynomial over a real-closed field (Sturm proved them for $\mathbb{R}$ but the same proofs go through

for any real-closed field.
To descibe these ideas and results of Sturm, we introduce some notations first.
Let R be a real-closed field. As it has a unique ordering, it makes sense to write notations like (a, b), $[a, b]$ etc. that one uses for real intervals and the notation $|a|$ for the positive element among $\pm a$. Let $f = c_0 + c_1 x + \cdots + c_{n-1}x^{n-1} + x^n \in R[x]$. We write $|a|$ in R to mean the obvious element - it is a or $-a$ according as $a \in P$ or $-a \in P$. The familiar algorithm to find the GCD of two polynomials, when applied to f and f', goes as follows :

$$f_0(x) = f(x), f_1(x) = f'(x),$$

$$f_{i-1}(x) = g_i(x)f_i(x) - f_{i+1}(x)$$

with deg $f_{i+1} <$ deg f_i for $i \geq 1$.
This sequence $f_0, f_1, \cdots$ has come to be known nowadays as the *standard sequence of* f. Looking at the smallest l for which $f_{l+1} = 0$, note that $f_l = GCD(f, f')$. Let V_r denote the number of changes of sign among $f_0(r), f_1(r), \cdots, f_l(r)$ - for counting this, one may simply drop all terms which are zero.
Then, we have the following easy :

Lemma 2 :
Let $M = max\ (1, |c_0| + \cdots + |c_{n-1}|)$. *Then every root of* f *in* R *belongs to* $[-M, M]$.
Proof :
If $a \in R$ is a root of f such that $|a| \leq 1$, there is nothing to prove. Suppose therefore that $|a| > 1$. Taking absolute values in

$$-a = c_{n-1}a^{-1} + \cdots + c_1 a^{-(n-2)} + c_0 a^{-(n-1)},$$

we get $|-a| \leq |c_{n-1}| + \cdots + |c_1| + |c_0|$. This proves the lemma.

The following result has the same proof as for $\mathbb{R}$ which Sturm gave in the 1880's. It is crucially used by Artin and Schreier in proving the existence of real-closures and their uniqueness properties.

Lemma 3 (Sturm) :
Let R *be a real-closed field and* $f \in R[x]$. *Then, the number of roots of* f *in any* (a, b) *such that* $f(a)f(b) \neq 0$, *equals* $V_a - V_b$.

Theorem (Artin-Schreier) :
Any ordered field has a real-closure. If K_1, K_2 are ordered fields with respective real-closures R_1, R_2, then any order isomorphism from K_1 onto K_2 uniquely extends to an isomorphism of the fields R_1 and R_2. Further, this extension is an order-isomorphism. In particular, an automorphism of a real-closure R of an ordered field K, which is identity on K must be the identity.

Finally, we state the main theorem.

Artin's theorem (in a slightly generalized form) :
Let K be an ordered field such that K has a unique ordering and such that K is dense in its real-closure K^*. Call Q the prime field of K. Let us write $L := K(x_1, \cdots, x_n)$ as $K(x)$ for short. Let $f \in L$ such that $f(a_1, a_2, ...a_n) \geq 0$ for all $a_i \in Q$ for which f is defined (f is defined over the Zariski open set corresponding to its denominator). Then, there exist $g_i(x) \in L$ such that $f = \sum g_i(x)^2$. The converse is true and obvious.

Towards the proof, we introduce a few notations and make a few preliminary observations.
A subfield K of an ordered field L is said to be *dense* if each interval (a, b) of L contains a point of K.
For an ordered field K, we denote by K^* its real-closure. Now, the ordering on K extends to an ordering on $L = K(x_1, \cdots, x_n)$ - to see this, just observe that since $K^*(x_1, \cdots, x_n)$ is formally real, it has an ordering which extends the unique ordering of K^*; but then its restriction to $K(x_1, \cdots, x_n)$ is an ordering extending the given ordering on K.
Denote by L^*, the real-closure of $L = K(x_1, \cdots, x_n)$. The whole trick now is to be able to go back and forth between K and L by means of the nice properties that their real-closures K^* and L^* have. With the above notations, we prove the following two lemmata which are key steps in the proof of the theorem :

Lemma 4 :
If $f(x_1, \cdots, x_n, Y) \in L[Y]$ has r distinct roots in L^*, then there are $g_1, \cdots, g_m$ in L with the property that whenever $a_1, \cdots, a_n \in K$ with

$$g_1(a_1, \cdots, a_n), \cdots, g_m(a_1, \cdots, a_n)$$

having the same signs (in K) as the sequence $g_1, \cdots, g_m$ (in L), there are

precisely r distinct roots of $f(a_1, \cdots, a_n, Y)$ *inside* K^*.

Proof :

We write $f(x_1, \cdots, x_n, Y) = c_0 + c_1 Y + \cdots + c_r Y^r$ with $c_i \in L$. The g_i's will be defined in terms of the c_j's and some specializations of the variable Y in f. Let us consider the standard sequence for f as a polynomial in Y; say $f_0, f_1, \cdots, f_l$. If m denotes $1 + \sum |c_i|$ (note that this is in L), we consider g_i's to be the sequence made up of :- (i) all the coefficients of $f_0, \cdots, f_l$, (ii) all $f_i(Y = \pm m)$, and (iii) all $f_i(Y = \pm m)/f_j(Y = \pm m)$ with $i < j$ and $f_i(Y = \pm m) \neq 0$.

Then, lemmata 2 and 3 imply that $V_{-m}(f) - V_m(f) = r$. But, if $a_1, \cdots, a_n \in K$ with the sequence g_i having the same signs (in L) as $g_i(a_1, \cdots, a_n)$ (in K) for the above choice of the g_i's shows that

$$V_{-m(a_1,\cdots,a_n)}(f(a_1, \cdots, a_n, Y) = V_{-m}(f),$$

$$V_{m(a_1,\cdots,a_n)}(f(a_1, \cdots, a_n, Y) = V_m(f).$$

Observing that the c_i's are also among the g_j's, all the roots of $f(a_1, \cdots, a_n, Y)$ in K^* are bounded between $-m(a_1, \cdots, a_n)$ and $m(a_1, \cdots, a_n)$. This completes the proof.

Lemma 5 :

Suppose $f_i(x_1, \cdots, x_n, Y) \in L[Y]; 1 \leq i \leq r$ *are monic polynomials. Suppose* $b_1 < \cdots < b_r$ *are roots in* L^* *of the corresponding* f_i*'s. Then, there are elements* $g_1, \cdots, g_m \in L$ *with the property that whenever* $a_1, \cdots, a_n \in K$ *with* $g_1(a_1, \cdots, a_n), \cdots, g_m(a_1, \cdots, a_n)$ *having the same signs (in* K*) as the sequence* $g_1, \cdots, g_m$ *(in* L*), there are* $b_1^* < \cdots < b_r^*$ *in* K^* *which are roots of corresponding* $f_i(a_1, \cdots, a_n, Y) \in K[Y]$.

Proof :

Consider the field $M = L(b_1, \cdots, b_r, \sqrt{b_2 - b_1}, \cdots, \sqrt{b_r - b_{r-1}})$, which is a finite extension of L.

Being in characteristic zero, we have $M = L(b)$ for some b.

Write $g(x_1, \cdots, x_n, Y) = c_0 + c_1 Y + \cdots + c_{s-1} Y^{s-1} + Y^s \in L[Y]$ with $c_i \in L$, for the minimal polynomial of b over L. The elements b_i, $\sqrt{b_{j+1} - b_j}$ and $1/\sqrt{b_{j+1} - b_j}$ have polynomial expressions in b over L; we write

$$b_i = b_i'(x_1, \cdots, x_n, b) \ ,$$

$$\sqrt{b_{j+1} - b_j} = e_j(x_1, \cdots, x_n, b) \ ,$$

$$1/\sqrt{b_{j+1}-b_j} = e'_j(x_1, \cdots, x_n, b)$$

where $b'_i(x_1, \cdots, x_n, Y), e_j(x_1, \cdots, x_n, Y), e'_j(x_1, \cdots, x_n, b) \in L[Y]$.
Now, as b_i is given to be a root of f_i, the polynomial
$F_i(x_1, \cdots, x_n, b'_i(x_1, \cdots, x_n, Y))$ vanishes at $Y = b_i$ which means that it must be divisible (in $L[Y]$) by the minimal polynomial $g(x_1, \cdots, x_n, Y)$.
Let us write, therefore,

$$F_i = gf'_i \in L[Y] \cdots\cdots\cdots (1)$$

for some polynomials $f'_i \in L[Y]$.
Similarly, on using the facts that the polynomials $b'_{j+1} - b'_j - (e'_j)^2$ and $e_j e'_j - 1$ in $L[Y]$ vanish at $Y = b$, we have some polynomials $h_j, h'_j \in L[Y]$ with

$$b'_{j+1} - b'_j - (e'_j)^2 = gh_j \cdots\cdots\cdots (2)$$

and

$$e_j e'_j - 1 = gh'_j \cdots\cdots\cdots (3)$$

in $L[Y]$.
Analogously to the previous lemma, let $g_1, \cdots, g_m \in L$ be the sequence of polynomials made up of : (i) all of the coefficients (in L) of all the $f_i, f'_i, g, h_j, h'_j, b'_i, e_j, e'_j$, and (ii) all the polynomials obtained by applying the previous lemma to $g(x_1, \cdots, x_n, Y)$.
Then, for $a_1, \cdots, a_n \in K$ as in our hypothesis here, the previous lemma implies that $g(a_1, \cdots, a_n, Y)$ has a root a^* in K^* since $g(x_1, \cdots, x_n, Y)$ has a root b in L^*.
Thus, (1) implies that $b^*_i := b'_i(a_1, \cdots, a_n, a^*)$ is a root of $f_i(a_1, \cdots, a_n, Y)$.
Using (2) and (3), we get that

$$e_j(a_1, \cdots, a_n, a^*)e'_j(a_1, \cdots, a_n, a^*) = 1$$

(so $e_j(a_1, \cdots, a_n, a^*) \neq 0$) and

$$b^*_{j+1} - b^*_j = (e_j(a_1, \cdots, a_n, a^*))^2 > 0 \ \ in \ \ K^*.$$

This completes the proof.

Finally, before proving the main theorem, we recall a few abbreviations we shall use for convenience. For a field K, we shall use the notation x to stand for n algebraically independent elements $x_1, \cdots, x_n$ over K and write

$f(x)$ to stand for a rational function in $x_1, \cdots, x_n$ over K. Further, if $L = K(x_1, \cdots, x_n)$, then we write a polynomial in a variable Y over L as $g(x; Y)$. If $a_1, \cdots, a_n \in K$, wealso write $f(a)$ and $g(a; Y)$ with the obvious meanings.

Proof of Artin's theorem :
Suppose there do not exist such $g_i(x)$'s. Then, we know from proposition 1 that $f(x) < 0$ in L for some ordering of L extended from the ordering of K. We claim that we can find $a_1, \cdots, a_n \in Q$ such that $f(a) < 0$. We proceed as follows.
The idea is to apply induction on the number n of variables. It is clearly true when there are no variables (that is if $n = 0$). Assuming the result holds for a fixed $n \geq 0$, we prove it for $n+1$ now. As mentioned before starting the proof, we will denote $(x_1, \cdots, x_n)$, $(a_1, \cdots, a_n)$ etc. by the symbols (x) and (a). The $(n+1)$-th variable x_{n+1} will be denoted by Y. L denotes $K(x_1, \cdots, x_n)$ as before. In other words, $f(x, Y)$ is our rational function; we may assume that it is in $L[Y]$, by multiplying with the square of the denominator if necessary. Further, it suffices to prove the theorem for the monic irreducible factors and the coefficient of the highest degree term of f in $L[Y]$, say $f_1, \cdots, f_r$. So, we may assume $f_i \in L$ or $f_i \in L[Y]$ is monic, irreducible for $i \leq r$. Suppose $b_1 < b_2 < \cdots < b_s$ are all the roots of $\prod_{i=1}^{r} f_i(x, Y)$ in L^*. Renumbering the f_i's and repeating them, if necessary, we may assume that b_i is a root of $f_i(x, Y)$ for $i \leq s$ and that $f_i \in L$ for $r \geq i > s$.
Let $g_1(x), \cdots, g_t(x)$ denote all those elements of L obtained as : (i) those elements of L obtained by applying lemma 4 to each f_i, or (ii) all those elements of L obtained by applying lemma 5 to the polynomials $f_i(x, Y)$ and roots b_i's for $i \leq s$, or (iii) the elements f_i for $r \geq i > s$, or (iv) the discriminants $d_i(x)$ of all the $f_i(x, Y)$ for $i \leq s$.
By the induction hypothesis, we can find $a_1, \cdots, a_n \in Q$ such that $g_i(a)$ and $g_i(x)$ have the same signs for $i \leq t$ and such that $\prod_{i=1}^{s} d_i(a) \neq 0$. Clearly, for $r \geq i > s$, we have that $f_i(a)$ and $f_i(x)$ have the same signs.
Let us look at the f_i's with $i \leq s$ now. Suppose $b_{i_1}, \cdots, b_{i_{n_i}}$ are the roots of f_i inside L^*. In our notation, this means that some f_j can coincide with f_i only if $j = i_k$ for some $k \leq n_i$. Applying lemma 5 to these roots, we have roots $b_1^* < \cdots b_s^*$ of $f_1(a, Y), \cdots, f_s(a, Y)$. Then $b_{i_1}^* < \cdots b_{i_{n_i}}^*$ are roots of $f_i(a, Y)$. But then lemma 4 implies that they are all the roots of f_i in K^*.
Writing $f_i(x, Y) = \prod_j q_{i_j}(x, Y).\prod_k (Y - b_{i_k})$ with $q_{i_j}(x, Y) \in L^*[Y]$ monic irreducible and of degree > 1. Therefore, q_{i_j} must be of the form $(Y+p)^2+q$ for some $p, q \in L^*$, $q > 0$. By the choice of a, since the discriminants do not

vanish, the $f_i(a, Y)$ have no multiple roots. Thus,

$$f_i(a, Y) = \prod_j q_{i_j}(a, Y). \prod_k (Y - b^*_{i_k})$$

with $q_{i_j}(a, Y)$ irreducible in $L^*[Y]$ since $b^*_{i_k}$ are the only roots of f_i in K^*. Hence, we must have $q_{i_j}(a, a_{n+1}) > 0$ for any a_{n+1}. In other words, the signs of $f_i(a, a_{n+1})$ and $f_i(x, Y)$ are determined by k and k' such that $b^*_u < a_{n+1}$ if and only if, $u < k$ and $b_u < Y$ if and only if, $u < k'$. Hence, it suffices to choose $a_{n+1} \in Q$ such that $k = k'$. This completes the proof.

Remarks in conclusion :

Tarski's model-theoretic method can be summed up in this context as the following theorem :
Let R_1 and R_2 be real closed fields having a common ordered subfield F; i.e., the orderings on F induced by R_1 and R_2 are identical. Suppose that we have a finite set S of polynomial equations, inequations (that is, statement of the form $f \neq 0$) and inequalities (that is, statements of the form $f > 0$) with coefficients in F. Then S has a solution in R_1 if and only if it has a solution in R_2.
Hilbert's 17th problem can also be restated in terms of Witt groups of quadratic forms and is intimated related to the so-called Pfister local-global principle. Roughly, this can be described as follows. For a field K, one has the commutative ring $W(K)$ of similarity classes of regular quadratic forms - f, g are similar if they are isomorphic upto adding copies of the hyperbolic plane. The operations are the orthogonal sum and the tensor product. It turns out that the possible positive cones in K are in bijective correspondence with those prime ideals of $W(K)$ which have characteristic zero. The bijection is through the so-called signature map which counts for a diagonal form $\sum a_i X_i^2$, the difference

$$|\{i : a_i > 0\}| - |\{i : a_i < 0\}|.$$

Using this terminology, Artin's theorem can be stated as :
Consider a real-closed field K, and $0 \neq f \in K[x_1, \cdots, x_n]$. Suppose that, for each anisotropic vector $a \in K^n$, the signature of the form $X^2 - f(a)Y^2$ is zero. Then, the form $X^2 - fY^2$ over $K(x_1, \cdots, x_n)$ has image in the Witt

group to be of finite order.
As a matter of fact, Pfister's local-global principle alluded to above says that the finite order elements in the above Witt group of the function field can be identified with the kernel of the signature homomorphism. In this language, it is easy to generalize Artin's theorem to the case of any even number of non-zero polynomials in place of the single polynomial f.

References :

[DP] C.N.Delzell & A.Prestel, *Positive polynomials*, Springer Monographs in Mathematics, 2001.
[J] N.Jacobson, *Basic algebra I,II.*
[L] S.Lang, *Algebra.*
[N] M.Nagata, *Field theory.*