I radical: $p^n \in I \Rightarrow p \in I$

I real: $p_1^2 + \dots + p_k^2 \in I \Rightarrow p_1, \dots, p_k \in I$

I radical iff $p^2 \in I \Rightarrow p\ I : p^n \in I$

$n = 2m$: $p^{2m} \in I \Rightarrow (p^m)^2 \in I \Rightarrow p^m \in I \overset{\text{ind.}}{\Rightarrow} p \in I$

$n = 2m+1$: $p^{2m+1} \in I \Rightarrow p^{2m+2} \in I \Rightarrow p^{m+1} \in I \overset{\text{ind.}}{\Longrightarrow} p \in I$

REAL NULLSTELLENSATZ AND SUMS OF SQUARES

MARIA MICHALSKA

Abstract. In this paper we highlight the foundational principles of sums of squares in the study of Real Algebraic Geometry. To this aim the article is designed as mainly a self-contained presentation of a variation of the standard proof of Real Nullstellensatz, the only relevant omission being the (long) proof of the Tarski-Seidenberg theorem. On the way we see how the theory follows closely developments in algebra and model theory due to Artin and Schreier. This allows us to present on the way Artin's solution to Hilbert's 17th Problem: whether positive polynomials are sums of squares. These notes are intended to be accessible to math students of any level.

Introduction

Any sum of squares of real numbers is equal zero if and only if the numbers are zero themselves; this is not true anymore over the algebraic closure of the real field. These fundamental facts underlie a host of subtle differences of Algebraic Geometry over the Real and the Complex numbers. The first and foremost difference is the Nullstellensatz, a theorem which describes the relation between algebraic objects and their vanishing sets. The complex Nullstellensatz asks the defining ideal of a set to be radical, whereas the Real Nullstellensatz demands more: for the ideal to be real, that is to have the property that if a sum of squares is an element of this ideal, then all summands are elements of the ideal also.

This may come as surprise, but the Real Nullstellensatz was unknown until the paper [Risler, 1970] of Jean-Jacques Risler in 1970. By all means, the sums of squares were already a very prominent element in the study of Algebraic Geometry over the reals. In 1900 among the famous problems of David Hilbert was the following, the 17th Problem: is any nonnegative polynomial a sum of squares? This questions lies naturally in Hilbert's general predisposition to formalize mathematics, since being a sum of squares is an algebraic certificate for nonnegativity. It was already known to Hilbert that one cannot demand the positive polynomial to be a sum of squares of polynomials. Nevertheless, one had to wait for Emil Artin to present in 1927 a positive solution for rational functions, [Artin, 1927]. (As a sidenote, one would like to remark that by [Delzell, 1984] a nonnegative polynomial is even a sum of squares of regulous functions which currently are quite intensively studied, compare [Fichou et al., 2016]. The aforementioned fact can be seen as basis of Nullstellensatz for regulous functions which again demands the defining ideals to be simply radical, as in the complex case.). This comes therefore as no surprise that elements of Artin-Schreier Theory are useful in the proof of Real Nullstellensatz. Thus we will present some elements of the theory and use the opportunity to present a full proof of Artin's solution to Hilbert's 17th Problem.

Real Radical $\sqrt[R]{I} := \{ p \mid p^{2m} + q_1^2 + \dots + q_k^2 \in I \}$

$m, k \geq 0$

$I \subseteq \sqrt{I} \subseteq \sqrt[R]{I}$

This note was designed foremost as a self-contained presentation of a variation of the standard proof of Real Nullstellensatz, we will omit only the (long) proof of the Tarski-Seidenberg theorem. These notes are intended to be accessible to math students of any level. Notes are organized as follows: presentation of the Real Nulsellensatz is given in Section 1 followed by explanation of notation and notions, as well as essential properties and proofs of intermediate results in Sections 2, 3 and 4. In Section 5 one finds the presentation and Artin's solution of Hilbert's 17th Problem and the paper ends with presentation of proof of Real Nullstellensatz over real closed fields in Section 6.1. On first lecture it is advised for a novice reader to prove Propositions and Properties left without proof.

1. Real Nullstellensatz

Every real algebraic set in $\mathbb{R}^n$ is defined to be the vanishing set of an ideal $I \triangleleft \mathbb{R}[X_1, \dots, X_n]$, i.e. it is a set of the form

$$V(I) = \{x \in \mathbb{R}^n : \forall_{f \in I} f(x) = 0\}.$$

Note that every polynomial ideal I is finitely generated by ,say, $f_1, \dots, f_k$, hence any real algebraic set can be given by one equation $f_1^2 + \dots + f_k^2 = 0$. We say an ideal I is real if from $\sum a_j^2 \in I$ follows all $a_j \in I$, see Section 2.

On the other hand, for a set $V \subset \mathbb{R}^n$ denote the defining ideal

$$\mathcal{I}(V) = \{f \in R[X] : \forall_{x \in V} f(x) = 0\}$$

i.e. $\mathcal{I}(V)$ is the largest ideal in $\mathbb{R}[X_1, \dots, X_n]$ such that all its elements vanish on V. Obviously, always $I \subset \mathcal{I}(V(I))$.

Real Nulstellensatz ties the geometric meaning of ideals with the algebraic meaning of sets in the real euclidean space in the following way:

Theorem 1 (Real Nullstellensatz). *Let $I \triangleleft \mathbb{R}[X_1, \dots, X_n]$.*

$$I = \mathcal{I}(V(I)) \iff I \text{ is real}$$

Proof of Real Nullstellensatz is given in the last Section. Reader is advised to start with the proof and go back to relevant sections when needed.

2. Basic algebra

Throughout this section let R be a commutative ring (with unity) and $I \triangleleft R$ an ideal.

Definition 2.1. *I is real if*

$$a_1^2 \dots + a_k^2 \in I \Rightarrow a_1, \dots, a_k \in I$$

for any $a_1, \dots, a_k$.

Property 2.2. *(1) If an ideal is prime, then it is radical.*
(2) If an ideal is real, then it is radical.

(1) I prime: $p \cdot q \in I \Rightarrow p \in I$ or $q \in I$, I radical: $p^n \in I \Rightarrow p \in I$.
Prime $\Rightarrow$ Radical: $p^n \in I, n \geq 2 \Rightarrow p \cdot p^{n-1} \Rightarrow p \in I$ or $p^{n-1} \in I$.

(2) I real: $p_1^2 + \dots + p_k^2 \in I \Rightarrow p_1, \dots, p_k \in I$.
Real $\Rightarrow$ Radical: $p^{2m+1} \in I \Rightarrow p^{2m+2} \in I \Rightarrow (p^{m+1})^2 \in I \Rightarrow p^{m+1} \in I \overset{ind.}{\Rightarrow} p \in I$

Property 2.3. *I is prime iff the quotient ring R/I is an integral domain i.e. has no zero divisors.*

Property 2.4. (1) *Field R embeds naturally into $R[X_1,\dots,X_n]/I$ if $I \neq R[X_1,\dots,X_n]$.*
(2) *Integral domain R embeds naturally into its field of fractions* $\mathrm{Quot}(R)$.

Definition 2.5. *I is primary if*

$$ab \in I \Rightarrow a \in I \text{ or } b^m \in I \text{ for some } m \in \mathbb{N}.$$

Definition 2.6. *We say that the commutative ring is noetherian if every ascending chain of ideals stabilizes.*

The above is equivalent to saying that every ideal is finitely generated. Note that every field is noetherian, because it contains only two ideals (0) and (1).

Theorem 2.7 (Hilbert's basis theorem). *If R is a noetherian ring, then the ring of polynomials $R[X_1,\dots,X_n]$ is also noetherian.*

Theorem 2.8 (Noether-Lasker Theorem). *Assume ring is noetherian. Every ideal is an intersection of finitely many primary ideals.*

Proof: We divide the proof into two steps.

• *Every ideal is a finite intersection of irreducible ideals.*

We say that an ideal I is irreducible if for any two ideals J, K if $I = J \cap K$, then $I = J$ or $I = K$. The proof is standard for noetherian rings:

Let A be the set of all ideals which are not a finite intersection of irreducible ideals. Take $I \in A$. If I cannot be expressed as an intersection of two ideals different from I, then I is irreducible. Therefore $I \notin A$. Hence $I = J_1 \cap K_1$. Obviously, either $J_1 \in A$ or $K_1 \in A$. Set $I_1 = J_1$ if $J_1 \in A$ or $I_1 = K_1$ otherwise. Proceed inductively, given $I_k \in A$ we have $I_k = J_k \cap K_k$ and $I_k \neq J_k$, $I_k \neq K_k$. Put

$$I_{k+1} = \begin{cases} J_k & \text{if } J_k \in A \\ K_k & \text{otherwise} \end{cases}$$

We get an ascending sequence

$$I \subset I_1 \subset \dots$$

of ideals. Since R is noetherian, we get $I_k = I_N$ for all $k \geq N$ and some $N \in \mathbb{N}$. But then $I_N = I_{N+1}$ contrary to assumption. Therefore $A = \emptyset$. This ends the proof.

• *Every irreducible ideal is primary*

Take an irreducible ideal I and take $ab \in I$. We will use quotients of ideals to prove that $a \in I$ or $b^m \in I$.

Define $J_k = I : (b^k) = \{c \in R : cb^k \in I\}$. We have that J_k are ideals and

$$I = J_0 \subset J_1 \subset J_2 \subset \dots$$

Since R is noetherian, the sequence stabilizes. Let J_N be such that $J_k = J_N$ for all $k \geq N$.

Put $J = J_N$ and $K = I + (b^N)$. Then obviously $I \subset J \cap K$. Moreover, if $c \in J \cap K$, then

$$c = i + f b^N,\ i \in I \tag{1}$$

and

$$b^N c \in I.$$

Multiplying both sides of (1) above by b^N we get

$$cb^N - i = f b^{2N}.$$

Hence $f b^{2N} \in I$. Therefore, $f \in J_{2N} = J_N$. Hence $f b^N \in I$ and from the form (1) we see $c \in I$. Therefore, $I = J \cap K$.

Since I is irreducible, we get either $I = K = I + (b^m)$ and $b^m \in I$ or $I = J_N$. In the latter case we have $I = J_N \supset J_1 \supset J_0 = I$, hence $J_1 = I$. Since $ab \in I$, hence $a \in I : (b) = I$. ■

Corollary 2.9 (Prime decomposition of a radical). *Assume ring is noetherian. Every radical ideal is a finite intersection of minimal prime ideals.*

Here a prime ideal p is minimal with respect to I if $I \subset p$ and for any p' prime: $I \subset p' \subset p \Rightarrow p' = p$.

Proof: Three easy steps.

• *The radical of primary ideal is prime*

Let I be primary and $\sqrt{I} = \{a \in R : a^m \in I \text{ for some } m\}$ be its radical. Take $ab \in \sqrt{I}$. Then $(ab)^m \in I$. Since I is primary, we get $a^m \in I$ or $b^{km} \in I$. From definition of radical, either $a \in \sqrt{I}$ or $b \in \sqrt{I}$.

• Since $I = p_1 \cap \cdots \cap p_k$ with p_i primary ideals due to Noether-Lasker Theorem and I is radical, then

$$I = \sqrt{I} = \sqrt{p_1 \cap \cdots \cap p_k} = \sqrt{p_1} \cap \cdots \cap \sqrt{p_k},$$

where every $\sqrt{p_i}$ is prime.

• *The prime ideals in decomposition can be taken as minimal.*

We have $I = p_1 \cap \cdots \cap p_k$ with all p_i prime. Fix $p_i =: p$. Consider any chain $(P_\alpha)_\alpha$ with respect to inclusion of prime ideals P_α such that $p \supset P_\alpha \supset I$ and $P_\alpha \subset P_\beta$ for $\alpha \geq \beta$. Then $P := \cap_\alpha P_\alpha$ is a prime ideal. Indeed, let $ab \in P$. Then $ab \in P_\alpha$ for every α. Assume $a, b \notin P$, then $a, b \notin P_\alpha$ for some α (α can be chosen in common for a, b because of inclusions). But this is contrary to assumption that P_α is prime. Hence every chain has a lower bound. Therefore by Kuratowski-Zorn Lemma[1] there exists a minimal element P_i. The prime ideal P_i is a minimal prime containing I by its definition.

One has $I = p_1 \cap \cdots \cap p_i \cap \cdots \cap p_m = p_1 \cap \cdots \cap P_i \cap \cdots \cap p_m$. Apply above reasoning to every ideal p_i in the representation. ■

[1]Kuratowski-Zorn Lemma: If every chain in a partially ordered set is bounded from below, then there exists a minimal element in the set.

Proposition 2.10. *Assume ring is noetherian. All minimal prime ideals containing a real ideal are real.*

Proof: Let I be a real ideal. Since real ideal is radical, from Corollary 2.9 we can write $I = p_1 \cap \cdots \cap p_r$ with p_i minimal prime ideals containing I. Assume p_1 is not real. Then we can take $a_1^2 + \cdots + a_k^2 \in p_1$ such that $a_1 \notin p_1$. Since p_l are minimal, we can choose $b_l \in p_l \setminus p_1$ for $l = 2, \ldots, r$. Put $b = \Pi_{l=2,\ldots,r} b_l$. We have $b \notin p_1$ by definition of b, because p_1 is prime. Then

$$(a_1 b)^2 + \cdots + (a_k b)^2 = (a_1^2 + \cdots + a_k^2) b^2 \in p_1 \cap \bigcap_{l=2,\ldots,r} p_l = I$$

and since I is real, we have $a_1 b \in I \subset p_1$. Since p_1 is prime, we get $a_1 \in p_1$ or $b \in p_1$. This gives a contradiction. Hence $a_1, \ldots, a_k \in p_1$ and p_1 is real. ■

Knowing there exists prime decomposition of radical ideals, we can reformulate Proposition 2.10 in a following way.

Corollary 2.11 (Real prime decomposition of real ideal). *Assume ring is noetherian. Every real ideal is a finite intersection of minimal real prime ideals.*

Now, the following paragraph is not necessary for proof of RN, but is basic and of interest in view of Artin-Lang homomorphism theorem.

Proposition 2.12. *Let R be a commutative ring. An R-algebra A is finitely generated iff it is isomorphic to a quotient ring $R[X]/I$ for some polynomial ring over R and an ideal $I \lhd R[X]$.*

Proof: Suppose A is finitely generated as an R-algebra, this means there exist polynomials $f_1, \ldots, f_k \in R[X_1, \ldots, X_n]$ such that $A = R[f_1, \ldots, f_k]$. Then put $\Phi : R[X_1, \ldots, X_k] \to A$ as $\Phi(f) = f(f_1, \ldots, f_k)$. Without doubt Φ is a surjective homomorphism. Take $I := \ker\Phi$. Then $R[X_1, \ldots, X_k]/I$ is isomorphic to A.

Now suppose that $R[X_1, \ldots, X_k]/I$ is isomorphic to A. Since the natural homomorphism $\Phi : R[X_1, \ldots, X_k] \ni f \to f + I \in A$ is surjective and $\Phi(f) = f(\Phi(X_1), \ldots, \Phi(X_k))$, we get $A = \Phi(R[X_1, \ldots, X_k]) = R[\Phi(X_1), \ldots, \Phi(X_k)]$. ■

3. Elements of Artin-Schreier Theory

One property that separates complex and real numbers is zeros of sums of squares.

Definition 3.1. *A field R is real if*

$$a_1^2 + \cdots + a_k^2 = 0 \Rightarrow a_1, \ldots, a_k = 0$$

(or satisfies any of the equivalent conditions of Theorem 3.5).

You can see that complex numbers cannot be a real field since $i^2 + 1^2 = 0$. The Artin-Schreier Theory deals with this in a model-theoretic way.

Another thing that sets apart real and complex numbers is the ordering.

Definition 3.2. *Let R be a ring. We say that $\leq$ is a total (linear) ordering of R if it is an ordering*

(i) $a \leq a$
(ii) $(a \leq b \wedge b \leq c) \Rightarrow a \leq c$ TRANSITIVE
(iii) $(a \leq b \wedge b \leq a) \Rightarrow a = b$ ANTISYMMETRIC

which is total (linear)

(iv) $a \leq b \vee b \leq a$

and consistent with addition and multiplication

(v) $a \leq b \Rightarrow (\forall_c \quad a + c \leq b + c)$
(vi) $(0 \leq a \wedge 0 \leq b) \Rightarrow 0 \leq ab$

We write $a < b$ when $a \leq b$ and $a \neq b$.

Property 3.3. *If ring R is ordered, then*

(1) $0 \leq a^2$, *in particular* $0 < 1$
(2) $0 \leq a \Rightarrow -a \leq 0$

Moreover, if R is a field, then

(3) $0 < a < b \Rightarrow 0 < \frac{1}{b} < \frac{1}{a}$
(4) $0 < ab \iff 0 < \frac{a}{b} \wedge b \neq 0$

Corollary 3.4. *If the ring R is ordered, then $\mathbb{N} \subset R$.*

If a field R is ordered, then $\mathbb{Q} \subset R$. In particular, $\operatorname{char} R = 0$.

Denote by R^2 all squares of elements of R.

Let us denote by $\sum R^2$ all finite sums of squares of elements of R.

If ring R is ordered, then $0 \leq a$ for all $a \in \sum R^2$. Not all rings can be ordered: note that for complex numbers -1 is a square, so ordering would imply all complex numbers to be zero.

First Artin-Schreier Theorem gives characterization of ordered fields as real fields.

Theorem 3.5 (Artin-Schreier Theorem for real fields). *Let R be a field. Following conditions are equivalent*

(1) *R is real i.e.* $a_1^2 + \cdots + a_k^2 = 0 \Rightarrow a_1, \ldots, a_k = 0$
(2) *-1 is not a sum of squares*
(3) *R can be ordered*

Proof: (1) $\iff$ (2) If $-1 \in \sum R^2$, then $-1 = a_1^2 + \cdots + a_k^2$. Hence $0 = 1^2 + a_1^2 + \cdots + a_k^2$ and R is not real. If $\sum_{j=1,\ldots,k} a_j^2 = 0$ and $a_1 \neq 0$, then $\sum_{j \neq 1} \left(\frac{a_j}{a_1}\right)^2 = -1$.

(3) $\Rightarrow$ (2) Assume R is ordered. If $-1 = \sum_{j=1,\dots,k} a_j^2$, then $0 \leq -1$. Hence $0 < 1+(-1) = 0$ which gives a contradiction.

(1), (2) $\Rightarrow$ (3) To prove this we will introduce a set defining the ordering.

We say $P \subset R$ is a proper cone, if
- (a) $\sum R^2 \subset P$
- (b) $P + P \subset P$, $P \cdot P \subset P$ CLOSED UNDER ADDITION AND MULTIPLICATION
- (c) $-1 \notin P$ PROPER
- (d) $-P \cap P = \{0\}$ ANTISYMMETRIC

A proper cone P is said to be a positive cone if
- (e) $P \cup -P = R$ TOTAL

Naturally, $-P := \{a \in R : -a \in P\}$. Note that if $\sum R^2$ is a positive cone, then it is the unique positive cone of R.

• *There is a one-to-one correspondence between total orderings of R and positive cones P of R. The correspondence is given by*

$$a \leq b \iff b - a \in P.$$

Indeed, every total ordering defines a positive cone and every positive cone defines a total ordering in this way.

• *For R real there exists a maximal proper cone in R. A maximal proper cone is a positive cone.*

The set $\sum R^2$ is a proper cone by assumption that -1 is not a sum of squares. Consider any chain (P_α) of proper cones. Then $P := \bigcup(P_\alpha)$ is a proper cone. Indeed, it is obvious that P satisfies points (a)-(c) of the definition. To prove (d) it suffices to note that $P_\alpha \cap -P_\beta = \{0\}$ for all α, β. Hence $0 \in P \cap -P \subset \{0\}$. Therefore, every chain is bounded from above and by Kuratowski-Zorn Lemma there exists a maximal proper cone $P_\leq$ in R.

Assume $P_\leq$ is not a positive cone. Then for $c \notin P_\leq \cup -P_\leq$ we have that c is not a sum of squares and $P_c := P_\leq + cP_\leq$ is the smallest proper cone containing $P_\leq \cup \{c\}$. Since $P_\leq$ is maximal, we get $P_\leq = P_c$. Hence $c \in P_\leq$. Contradiction.

Therefore, every real field contains a positive cone, hence it can be ordered. ∎

Proposition 3.6. *Let R be a ring and $I \lhd R$ a prime ideal. Field of fractions* $\mathrm{Quot}(R/I)$ *is real iff I is real.*

Proof: Note that $(a+I)/(b+I) = 0$ in $\mathrm{Quot}(R/I)$ iff $a \in I$ and $b \notin I$. In particular

$$\sum_{i=1,\dots,k} \left(\frac{f_i + I}{g_i + I}\right)^2 = 0 \iff \sum_{i=1,\dots,k} \left(\frac{f_i g_1 \cdots g_k}{g_i}\right)^2 \in I.$$

So if we assume I is real, then for $\sum_{i=1,\dots,k} \left(\frac{f_i+I}{g_i+I}\right)^2 = 0$ we get $\frac{f_i g_1 \cdots g_k}{g_i} \in I$ for every i. Therefore $f_i/g_i = 0$ for every i. On the other hand, if $\mathrm{Quot}(R/I)$ is real and we take

$f_1^2 + \cdots + f_k^2 \in I$, then $(f_1 + I)^2 + \cdots + (f_k + I)^2 = 0$ and it follows $f_i + I = 0$ for all i. Therefore, $f_i \in I$. ■

Definition 3.7. *A field R is algebraically closed if any univariate polynomial over R has a root in R.*

Theorem 3.8. *For any field R if a field C is an algebraic extension of R and every polynomial $R[t]$ has a root in C, then C is algebraically closed.*

This characterization of extensions is classic for field theory, for proof you can look up [Isaacs, 1980].

Definition 3.9. *A real field R is real closed if its algebraic extension $R[\sqrt{-1}] = R[X]/(X^2+1)$ is proper and algebraically closed.*
(or when R satisfies any of the equivalent conditions of Theorem 3.11)

Note $R(a) = R[a]$ for algebraic extension of field R.

Remark 3.10. The field $\mathbb{R}$ is a real closed field.

Theorem 3.11 (Artin-Schreier Theorem for real closed fields). *Let R be a field. Following conditions are equivalent:*

1. *(1) R is real closed i.e. its algebraic extension $R[\sqrt{-1}]$ is proper and algebraically closed.*
2. *(2) R is real and has no (proper) algebraic extension which is real*
3. *(3) the positive cone of R is the squares R^2 and any odd-degree polynomial has a root in R*

Proof: In the proof we will use a following remark

- *For a field $R \neq R[\sqrt{-1}]$ we have*

$$\left(R = R^2 \cup -R^2 \ \wedge\ R^2 = \sum R^2\right) \iff R[\sqrt{-1}] = \left(R(\sqrt{-1})\right)^2$$

Indeed, assume $R = R^2 \cup -R^2$ and $R^2 = \sum R^2$. Take any $a + \sqrt{-1}b$ with $a, b \in R$. The discriminant of $f = 4X^2 - 4aX - b^2$ is $(4a)^2 + (4b)^2 \in R^2$, hence a root c of f lies in R. Since $R = R^2 \cup -R^2$, we get $c = \alpha^2$ or $c = (\sqrt{-1}\alpha)^2$. Put $x = \alpha$ and $y = \frac{b}{2\alpha}$ in first case or $x = \sqrt{-1}\alpha$ and $y = \frac{b}{\sqrt{-1}\alpha}$ otherwise. Then $x + \sqrt{-1}y \in R[\sqrt{-1}]$ and $(x + \sqrt{-1}y)^2 = a + \sqrt{-1}b$.

Assume $R[\sqrt{-1}] = (R[\sqrt{-1}])^2$. Take $a \in R$. There is $b + \sqrt{-1}c$, $b, c \in R$, such that $a = (b + \sqrt{-1}c)^2 = b^2 - c^2 + 2\sqrt{-1}bc$. Hence $b = 0$ or $c = 0$ and $a = -c^2$ or $a = b^2$ respectively. This proves $R = R^2 \cup -R^2$. To prove $R^2 = \sum R^2$ it suffices to show $a^2 + b^2$ is a square. Take $c, d \in R$ such that $a + \sqrt{-1}b = (c + \sqrt{-1}d)^2$. Then $a = c^2 - d^2$, $b = 2cd$ and $a^2 + b^2 = (c^2 + d^2)^2$.

(1)⇒(2) Since $R[\sqrt{-1}]$ is a proper algebraic closure of R, in particular we have $\sqrt{-1} \notin R$ and $R[\sqrt{-1}] = (R[\sqrt{-1}])^2$. Hence $R = R^2 \cup -R^2$, $R^2 = \sum R^2$ and $R^2 \cap -R^2 = \{0\}$. Therefore R has a positive cone, hence is real.

Any proper algebraic extension of R contains an element $a+\sqrt{-1}b \in R[\sqrt{-1}] \setminus R$. Since $b \neq 0$ we have $R[a+\sqrt{-1}b]$ equals

$$R[X] \,/\, (x^2 - 2ax + a^2 + b^2),$$

thus $a - \sqrt{-1}b \in R[a+\sqrt{-1}b]$. Hence $(a+\sqrt{-1}b + a - \sqrt{-1}b)/2b = \sqrt{-1}$ and $R[a+\sqrt{-1}b] = R[\sqrt{-1}]$. Hence any proper algebraic extension of R is algebraically closed. Algebraically closed field is never real.

(2)⇒(3) Suppose $a \in R \setminus R^2$. Then $R[\sqrt{a}]$ is an algebraic extension of R, by assumption it is not real. Hence

$$-1 = \sum_j (b_j + c_j\sqrt{a})^2 = \sum b_j^2 + a\sum c_j^2 + \sqrt{a}\sum 2b_jc_j.$$

Therefore $\sum 2b_jc_j = 0$ and $a = -(1^2 + \sum b_j^2)/\sum c_j^2$. Hence $a \leq 0$. Therefore every positive element is a square.

Now we need to show every odd-degree polynomial has a root in R. Any polynomial of degree 1 is linear and has a root in R. Assume all odd-degree polynomials of degree $< d$ have a root in R. Let $f \in R[X]$ be of odd degree d and suppose f does not have a root in R. If f was reducible, then one of the factors would be an odd-degree polynomial of degree lower than f, hence f would have a root in R. Therefore f is irreducible over R. Then $R[f] = R[X]/(f)$ is an algebraic extension of R. By assumption the field of fractions is not real. Therefore there exist $g_1, \dots, g_k$ of degrees $< d$ such that

$$-1 = \sum \Big(g_j + (f)\Big)^2 = \sum g_j^2 + (f).$$

Note that $\deg\Big(\sum g_j^2\Big) \leq 2d-2$ and is even (because the leading coefficient is a sum of squares in R and R real, hence it does not vanish). Hence $-1 = \sum g_j^2 + fh$ for some h of odd degree $\leq d-2$. By inductive assumption, h has a root a in R. We get $-1 = \sum g_j^2(a) + f(a)h(a) = \sum \Big(g_j(a)\Big)^2$, so $-1 \in \sum R^2$. Contradiction.

(3)⇒(1) Under assumption (3) we have $-1 \notin R^2$, hence $R[\sqrt{-1}] \neq R$.

We will show any polynomial over R of degree $d = 2^m n$, n odd, has a root in $R[\sqrt{-1}]$ by induction on m. When $m = 0$ we get the claim from assumption (3). Assume for any $m' < m$ the assumption holds. Consider polynomial f of degree $d = 2^m n$. Let $a_1, \dots, a_d$ be roots of f in the algebraic closure of R. For $N \in \mathbb{N}$ put

$$g_N(X) = \Pi_{i<j}(X - a_i - a_j - Na_ia_j).$$

The polynomial g_N is of degree $d(d-1)/2 = 2^{m-1}(2^m n - 1)$ and it is symmetric in a_j, the roots of f. From fundamental theorem of symmetric polynomials, see [Macdonald, 1979], we get that coefficients of g_N can be expressed in terms of coefficients of f, hence $g_N \in R[X]$. From inductive assumption every g_N has a root in $R[\sqrt{-1}]$. Hence there exist i,j and $N, N' \in \mathbb{N}$, $c, c' \in R[\sqrt{-1}]$ such that $a_i + a_j + Na_ia_j = c = c' + (N - N')a_ia_j$. Therefore a_ia_j and $a_i + a_j$ are elements of $R[\sqrt{-1}]$.

We have $(X - a_i)(X - a_j) = X^2 - (a_i + a_j)X + a_ia_j$ is a quadratic polynomial over $R[\sqrt{-1}]$ with roots a_i, a_j and its discriminant is $(a_i + a_j)^2 - 4a_ia_j = (a_i - a_j)^2$. Since

$R = R^2 \cup -R^2$, then $R[\sqrt{-1}] = (R[\sqrt{-1}])^2$. Hence exists $c \in R[\sqrt{-1}]$ such that $c^2 = (a_i - a_j)^2$. Therefore from formulæ for solving quadratic equations we get a_i or $a_j \in R[\sqrt{-1}]$ and f has a root in $R[\sqrt{-1}]$. This ends the inductive proof. ∎

Definition 3.12. *We say that a real field $\overline{R}$ is an extension of an ordered ring R if R embeds into $\overline{R}$ with its ring operations and ordering.*

Theorem 3.13. *Every real field has a (unique) minimal extension to a real closed field.*

Proof: Note that algebraically closed field is not a real field, because -1 is a square.

Take a real field R with ordering $\leq$ and its algebraic closure C. Consider any chain $(R_\alpha, \leq_\alpha)$ of algebraic extensions of R (contained in C) with consistent orderings. The field $\bigcup R_\alpha$ is an algebraic extension of R (because it is contained in the algebraic closure). Moreover, if $a_1^2 + \cdots + a_k^2 = 0$ for $a_1, \ldots, a_k \in \bigcup R_\alpha$, we get $a_1, \ldots, a_k \in R_\alpha$ for some α. Since R_α is real, then $a_1 = \cdots = a_k = 0$ and $\bigcup R_\alpha$ is real. Hence by Kuratowski-Zorn Lemma there exists a maximal real field $\overline{R} \subset C$ that is an algebraic extension of R with consistent ordering. The only algebraic extension of $\overline{R}$ is C and C is not real. Hence $\overline{R}$ is a real closed field. Obviously, if $R \subset R' \subset \overline{R}$ and R' is real closed, then $R' = \overline{R}$. ∎

Uniqueness in the theorem is up to an order-preserving isomorphism. For instance, one can define infinitely many orderings in $\mathbb{R}(t)$ and some of them have non-isomorphic extensions if we ask the isomorphism to respect the order.

4. Tarski's Transfer Principle

Definition 4.1. *We say that a formula is a boolean combination in variables $X_1, \ldots, X_n$ over an ordered ring R if it is a (syntax correct) finite combination of formulas of the form $f(X_1, \ldots, X_n) \geq 0$ with $f \in R[X_1, \ldots, X_n]$ and the logic operators $\vee, \wedge$ and $\neg$.*

Note that a polynomial is a finite (syntax correct) combination of elements of the field, variables $X_1, \ldots, X_n$, addition and multiplication.

Definition 4.2. *A first order formula over an ordered ring is a (syntax correct) finite combination of $\wedge, \vee, \neg$, boolean combinations over the ordered field and quantifiers $\forall, \exists$. The variables which are not under range of any of the quantifiers are called free variables (and the formula is in fact a sentential function in the free variables).*

The two definitions above are far from precise, for more exact formulation see [Robinson, 1963, Chapter VIII].

For instance $\Phi(X,Y): \; X^2 + 2Y^2 \leq 0 \Rightarrow Y = 0$ is a boolean combination with free variables X, Y. Then $\Phi_1(Y): \; \exists_x \, \Phi(x,Y)$ is a first order formula with free variable Y and $\Phi_2: \; \forall_y \, \Phi_1(x,y)$ is also a first order formula without free variables, Φ_2 is a true statement. The formula $\psi(X): \; \exists_y \sum_{j=1}^{\infty} X^j < y$ is not a first order formula.

We treat a first order formula Φ over R as a formula over an extension R_1 of R by taking the range in the quantifiers as R_1.

Remark 4.3. Formulas without free variables are either true or false.

We will now state and leave without proof the Tarski's Quantifier Elimination Theorem known in real algebraic geometry as Tarski-Seidenberg Theorem, see [Bochnak et al., 1998], [Tarski, 1951] or [Robinson, 1963] for different presentations of its proof.

Theorem 4.4 (Tarski-Seidenberg Theorem). *Let R be an ordered ring. Let $b(X_0, X_1, \ldots, X_n)$ be a boolean combination. There exists a boolean combination $B(X_1, \ldots, X_n)$ such that for any real closed field R_1 extending R we have*

$$\left\{x \in R_1^n :\ \exists_{x_0 \in R_1}\ b(x_0, x)\right\} = \{x \in R_1^n :\ B(x)\}$$

i.e. the projection of a semialgebraic set is semialgebraic.

This is equivalent to the following

Theorem 4.5 (Quantifier Elimination). *Let R be an ordered ring. For every* **first order** *formula $\Phi(X)$ over R there* **exists a boolean** *combination $B(X)$ over R such that for any real closed field R_1 extending R we have*

$$\forall_{x \in R_1} \left(\Phi(x) \iff B(x)\right).$$

It is important to note that quantifier elimination holds in the class of algebraically closed fields for constructible sets (see discussion of Lefschetz Principle and Minor Lefschetz Principle in [Seidenberg, 1958] or [Eklof, 1973]).

Now we can prove Tarski's transfer principle

Theorem 4.6 (Tarski's Transfer Principle). *Let R be an ordered ring. Let R_1, R_2 be real closed extensions of R and $B(X_1, \ldots, X_n)$ a boolean combination over R. Then*

$$\exists_{x \in R_1^n}\ B(x) \iff \exists_{x \in R_2^n}\ B(x)$$

Tarski's Transfer Principle can be equivalently stated as follows: theory of real closed fields is model-complete.

Proof: Note that since R is ordered, it is an integral domain and by Proposition 3.6 and Theorem 3.13, there exist real closed extensions of R.

Take $B(X_1, \ldots, X_n)$ a boolean combination over R. By Tarski-Seidenberg Theorem and finite induction we can eliminate the quantifier in the formula $\exists_{x_1, \ldots, x_n} B(x_1, \ldots, x_n)$ i.e. there exists a boolean combination $\tilde{B}$ such that for any real closed extension R_1 of R we have

$$\exists_{x \in R_1^n}\ B(x) \iff \forall_{y \in R_1} \exists_{x \in R_1^n}\ B(x) \iff \forall_{y \in R_1}\ \tilde{B} \iff \tilde{B}.$$

The formula $\tilde{B}$ does not have free variables, therefore it is either true or false. Due to Tarski's Quantifier elimination it has uniform logical value over all real closed fields extending R, in particular over R_1 and R_2. ■

5. Artin's solution of Hilbert's 17th Problem

Following theorems are not necessary for the proof of RN, but of interest partly because Artin-Schreier Theory was developed to answer the following question:

> **Hilbert's 17th Problem**
> Is every positive polynomial a sum of squares of rational functions?

In fact, the problem dates back to Minkowsky and Hilbert was considering mainly polynomials with rational coefficients. Hilbert already proved that there exist polynomials positive on $\mathbb{R}^n$ such that they are not sums of squares of polynomials. On the other hand, all nonnegative polynomials of degree d in n variables are sums of squares of polynomials if and only if $d \le 2$ or $n = 1$ or $d = 4$ and $n = 2$ (see for instance [Bochnak et al., 1998, Section 6.3]).

Theorem 5.1 (Solution to Hilbert's 17th Problem). *Let R be a real closed field and Q its subfield with the positive cone $P = Q \cap R^2$. Take $f \in Q[X_1, \dots, X_n]$ which is nonnegative i.e.*

$$\forall_{x \in R^n}\ f(x) \ge 0.$$

Then

$$f \in \sum P \cdot (Q(X))^2$$

i.e. $f = \sum a_j q_j^2(X)$ with $a_j \in P$ and $q_j \in (Q(X))^2$.

Proof: Take $f \in Q[X_1, \dots, X_n]$ nonnegative and suppose $f \notin \sum P \cdot (Q(X))^2$. Hence either $-f \in \sum P \cdot (Q(X))^2$ or not. In both cases, we can extend the proper cone $\sum P \cdot (Q(X))^2$ to a positive cone P' of $Q(X_1, \dots, X_n)$ such that $-f \in P'$ (compare page 7).

Write $f = \sum a_\alpha X^\alpha$ with $a_\alpha \in Q$. Consider the first order variable-free formula Φ with coefficients in Q of the form

$$\Phi : \quad \exists_{x_1, \dots, x_n} \sum a_\alpha X_1^{\alpha_1} \cdots X_n^{\alpha_n} < 0.$$

Note that Φ is equivalent to

$$\exists_{x_1, \dots, x_n} f(x_1, \dots, x_n) < 0$$

From the choice of ordering of $Q(X)$, the statement Φ is true over the real closure of $Q(X)$. By Tarski's Transfer Principle, Φ is also true over R. Therefore, there exists $x \in R^n$ such that $f(x) < 0$ which is against nonnegativity of f. ■

In particular the above theorem is the desired solution to Hilbert's problem: every nonnegative real polynomial is a sos of real rational functions ($R = Q = \mathbb{R}$). Moreover, every polynomial with rational coefficients is a sos of functions in $\mathbb{Q}(X)$ ($R = \mathbb{R}, Q = \mathbb{Q}$).

In the original solution of Hilbert's problem by Artin an important tool was:

Theorem 5.2 (Artin-Lang Homomorphism Theorem). *Let $R \subset R_1$ be real closed fields and A a finitely generated R-algebra. If there is a homomorphism $\phi_1 : A \to R_1$, then there exists a homomorphism $\phi : A \to R$.*

Proof: We may assume $A = R[X_1, \dots, X_n]/I$ by Proposition 2.12. Take a homomorphism $\phi_1 : A \to R_1$ and put $y = (\phi_1(X_1), \dots, \phi_1(X_n)) \in R_1^n$. Since $R[X_1, \dots, X_n]$ is noetherian, consider finitely many generators $f_1, \dots, f_k$ of I. For any polynomial $f = \sum a_\alpha X_1^{\alpha_1} \cdots X_n^{\alpha_n}$ we have

$$\phi(f+I) = \phi\left(\sum a_\alpha (X_1+I)^{\alpha_1} \cdots (X_n+I)^{\alpha_n}\right) = \sum a_\alpha \phi(X_1+I)^{\alpha_1} \cdots \phi(X_n+I)^{\alpha_n} = f(y).$$

Therefore $f_1(y) = \cdots = f_k(y) = 0$. By Tarski's Transfer Principle we get there exists $x \in R$ such that $f_1(x) = \cdots = f_k(x) = 0$. Now we see the homomorphism $\phi : A \to R$ given by assignment $X_i \to x_i$ is well-defined. ■

6. Proof of Real Nullstellensatz

In this section we prove Real Nullstellensatz. Careful reader may note that the proof remains the same if we replace $\mathbb{R}$ by a real closed field. Hence the following more general statement is true:

> Let R be a real closed field and $I \triangleleft R[X_1, \dots, X_n]$. We have
> $$I = \mathcal{I}(V(I)) \iff I \text{ is real}$$

6.1. **Proof of RN in easy direction.** Note that for any arbitrary set $V \subset R^n$, the ideal $\mathcal{I}(V)$ is real. Assume $I = \mathcal{I}(V)$. Take $a_1^2 \cdots + a_k^2 \in I$. Hence $a_1^2(x) \cdots + a_k^2(x) = 0$ at every point $x \in V$. Therefore, $a_1 = \cdots = a_k \equiv 0$ on V. Hence $a_1, \dots, a_k \in \mathcal{I}(V) = I$ and I is real. This holds in particular when $I = \mathcal{I}(V(I))$.

6.2. **Proof of RN for prime ideals.** Take a prime real ideal $I \subsetneq R[X]$. To prove RN it suffices to show that $I \supset \mathcal{I}(V(I))$.

Take $f \notin I$ and denote $g_1, \dots, g_k$ the generators of I. Due to Proposition 3.6 and Theorem 3.13 we can take R_1, the real closure of the real field $\mathrm{Quot}(R/I)$. Naturally R embeds into R_1, see Property 2.4, and one can check the natural embedding preserves the order. Note that 0 in R_1 is the image of I.

Consider elements $y_1 = X_1 + I, \dots, y_n = X_n + I$ of R_1 and the boolean combination

$$B(Y_1, \dots, Y_n): \; g_1(Y) = \cdots = g_k(Y) = 0 \wedge f(Y) \neq 0$$

defined over R.

Since f is polynomial i.e. $f = \sum a_\alpha X^\alpha$ a finite sum, we get

$$f(y) = \sum a_\alpha y^\alpha = \sum a_\alpha (X_1+I)^{\alpha_1} \cdots (X_n+I)^{\alpha_n} = \left(\sum a_\alpha X^\alpha\right) + I = f + I$$

Hence $f(y) = f + I \neq I = 0$ since $f \notin I$.

Analogously we show $g_1(y) = \cdots = g_k(y) = 0$.

The fields R_1 and R are both real closed fields over R. Therefore, from Tarski's Transfer Principle we get

$$\exists_{y \in R_1^n} B(y) \Rightarrow \exists_{x \in R^n} B(x).$$

Since the left-hand is true, there exists $x \in R^n$ such that $g_1(x) = \cdots = g_k(x) = 0$ and $f(x) \neq 0$. Hence $f(x) \neq 0$ for $x \in V(I)$. Therefore, $f \notin \mathcal{I}(V(I))$ and this ends the proof.

6.3. **Proof of RN for any ideals.** We will show that if RN is true for prime ideals, then it is true for any ideal.

Assume the left implication of RN holds for real prime ideals. Take any real ideal I. Hence I is radical and from prime decomposition of Corollary 2.9 we have

$$I = \bigcap_{i=1,\dots,r} p_i \tag{2}$$

where p_i are minimal prime ideals and from Proposition 2.11 follows that the ideals p_i are real.

Hence p_l in equality (2) are real. Since we have RN is true for prime ideals (and from Property **??**) we get

$$\mathcal{I}(V(I)) = \mathcal{I}\left(V(\bigcap_{l=1\dots,r} p_l) \right) = \mathcal{I}\left(\bigcup_{l=1\dots,r} V(p_l) \right) = \bigcap_{l=1\dots,r} \mathcal{I}(V(p_l)) = \bigcap_{l=1\dots,r} p_l = I.$$

This ends the proof. ■

References

[~~Artin,~~ 1927] Artin, E. (1927). Über die Zerlegung definiter Funktionen in Quadrate. *Abh. Math. Sem. Univ. Hamburg*, 5(1):100–115.

[Bochnak et al., 1998] Bochnak, J., Coste, M., and Roy, M.-F. (1998). *Real algebraic geometry*, volume 36 of *Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)]*. Springer-Verlag, Berlin. Translated from the 1987 French original, Revised by the authors.

[Delzell, 1984] Delzell, C. N. (1984). A continuous, constructive solution to Hilbert's 17th problem. *Invent. Math.*, 76(3):365–384.

[Eklof, 1973] Eklof, P. C. (1973). Lefschetz's principle and local functors. *Proc. Amer. Math. Soc.*, 37:333–339.

[Fichou et al., 2016] Fichou, G., Huisman, J., Mangolte, F., and Monnier, J.-P. (2016). Fonctions régulues. *J. Reine Angew. Math.*, 718:103–151.

[Isaacs, 1980] Isaacs, I. M. (1980). Roots of polynomials in algebraic extensions of fields. *Amer. Math. Monthly*, 87(7):543–544.

[Macdonald, 1979] Macdonald, I. G. (1979). *Symmetric functions and Hall polynomials*. The Clarendon Press, Oxford University Press, New York. Oxford Mathematical Monographs.

[Risler, 1970] Risler, J.-J. (1970). Une caractérisation des idéaux des variétés algébriques réelles. *C. R. Acad. Sci. Paris Sér. A-B*, 271:A1171–A1173.

[Robinson, 1963] Robinson, A. (1963). *Introduction to model theory and to the metamathematics of algebra*. North-Holland Publishing Co., Amsterdam.

[Seidenberg, 1958] Seidenberg, A. (1958). Comments on Lefschetz's principle. *Amer. Math. Monthly*, 65:685–690.

[Tarski, 1951] Tarski, A. (1951). *A decision method for elementary algebra and geometry*. University of California Press, Berkeley and Los Angeles, Calif. 2nd ed.