Hilbert's 17th Problem for Real Closed Fields à la Artin

Aaron Crighton

MAT 477

February 4, 2014

Main Theorem: $f \in \mathbb{R}[x]$, $f(x) \geq 0$ imply $f = \sum_i f_i^2$,

where $x = (x_0, \ldots, x_n)$, $\{f_j\}_j \subset \mathbb{R}(x)$, $\mathbb{R}[x]$ and $\mathbb{R}(x)$ are the ring of polynomials and the field of their fractions. We'll use fields and models.

Def 1: Field $\mathbb{F}$ is ordered with order $<_{\mathbb{F}}$ (or "$<$" if clear) when

i) $\forall x, y, z \in \mathbb{F}$, $x < y \Longrightarrow x + z < y + z$ (implies $char(\mathbb{F}) = 0$);

ii) $\forall x, y, z \in \mathbb{F}$, $(x < y$ and $0 < z) \Longrightarrow xz < yz$ (implies $x^2 > 0$ for $x \neq 0$).

Def 2: A real closed field is an ordered field $(\mathbb{F}, <_{\mathbb{F}})$ such that:

i) Every positive element of $\mathbb{F}$ has a square root in $\mathbb{F}$;

ii) Every odd degree polynomial of $\mathbb{F}$ has a root in $\mathbb{F}$.

Fact: Real closed fields admit quantifier elimination.

With p a prime number we'll also use the following

Easy Fact: Groups of size p^k have normal subgroups of index $= p$.

Lemma 1: If -1 and $b \in \mathbb{F}$ are not sums of squares in a field $\mathbb{F}$

then -1 is not a sum of squares (shortly ss) in $\mathbb{F}(\sqrt{-b})$.

Proof: Case $\sqrt{-b} \notin \mathbb{F}$ suffices, equivalently $\dim_{\mathbb{F}} \mathbb{F}(\sqrt{-b}) = 2$. Then

$$-1 = \sum_{i=1}^{m} (x_i + y_i\sqrt{-b})^2 \;\Rightarrow\; b = \frac{1+\sum_i {x_i}^2}{\sum_i {y_i}^2} = \sum_i {w_i}^2 \text{ since}$$

$(\sum_i {y_i}^2)^{-1} = \sum_i (y_i / \sum_j {y_j}^2)^2$, contrary to the assumption. $\square$

Question: Why $\mathbb{R}(x)$ and not $\mathbb{R}[x]$?

Proposition: The function $f(x, y) = x^4y^2 + x^2y^4 - x^2y^2 + 1$ is positive

but not a sum of squares in $\mathbb{R}[x, y]$.

Pf: Easy calculus $\Rightarrow$ f is positive $(min(f(x, y)) = 26/27)$. Suppose

$f = \sum q_i^2$ with $q_i \in \mathbb{R}[x]$. Notice that $deg(q_i) \leq 2$ w.r.t both x and y.

Then q_i is of the form:

$q_i = a_0^i + a_1^i x + a_2^i y + a_3^i xy + a_4^i x^2 + a_5^i y^2 + a_6^i x^2 y + a_7^i y^2 x + a_8^i x^2 y^2$.

Comparing coefficients of in the equation $f = \sum q_i^2$ shows $\sum(a_8^i)^2 = 0$

Hence $a_i^9 = 0$ for each i. Similarly, $a_4^i = a_5^i = 0$ for each i. Then

coefficients with $a_9^i = a_4^i = a_5^i = 0$ show that $a_2^i = a_1^i = 0$ as well. Finally,

Looking at the coefficent of x^2y^2 in the new equation

$$x^4y^2 + x^2y^4 - x^2y^2 + 1 = \sum(a_0^i + +a_3^i xy + a_6^i x^2 y + a_7^i y^2 x)^2$$

we obtain $-1 = \sum(a_3^i)^2$, which is impossible. $\square$

More Algebraic Results on Ordered Fields

Lemma 2: If $\mathbb{F}$ is a field where -1 is not a ss and $b \in \mathbb{F}$ is not a ss then $\mathbb{F}$ can be ordered so that $b < 0$.

Proof: Let $\mathbf{F} = \{\text{fields } \mathbb{K} : \mathbb{F}(\sqrt{-b}) \subset \mathbb{K} \subset \overline{\mathbb{F}} \text{ and -1 is not a ss in } \mathbb{K}\}$

By Zorn's Lemma, $\mathbf{F}$ has a maximal element $\mathbb{K}$. By Lemma 1, if c is not a ss in $\mathbb{K}$, then $\mathbb{K}(\sqrt{-c}) \in \mathbf{F}$. So, by maximality, $\sqrt{-c} \in \mathbb{K}$. Order $\mathbb{K}$ as follows:

$$x < y \iff y - x \neq 0 \text{ and } y - x \text{ is a square in } \mathbb{K}$$

This is easily checked to be well-defined. Then both $\mathbb{F}(\sqrt{-b})$ and $\mathbb{F}$

inherit this order as subfields and $-b = (\sqrt{-b})^2 > 0$ so $b < 0$. $\square$

Corollary 1: A field $\mathbb{F}$ can be ordered iff -1 is not a sum of squares in $\mathbb{F}$

Pf: Lemma 2 implies "$\Longleftarrow$" . For "$\Rightarrow$" note $1 = 1^2 > 0 \iff -1 < 0$ $\square$

Fund. Thm. Alg. If $\mathbb{F}$ is a real closed field, then $\mathbb{F}(\sqrt{-1})$ is alg. closed.

Proof: If a , $b \in \mathbb{F}$ then, $\left(\sqrt{\frac{a+\sqrt{a^2+b^2}}{2}} \pm \sqrt{\frac{-a+\sqrt{a^2+b^2}}{2}}\sqrt{-1}\right)^2 =$

$$\frac{a+\sqrt{a^2+b^2}}{2} - \frac{-a+\sqrt{a^2+b^2}}{2} \pm 2\sqrt{\frac{a+\sqrt{a^2+b^2}}{2}\frac{-a+\sqrt{a^2+b^2}}{2}}\sqrt{-1} = a \pm |b|\sqrt{-1} \ ,$$

where $|b| := \max\{b; -b\}$, i.e. elements in $\mathbb{F}(\sqrt{-1})$ have square roots.

Proof of Fund. Thm of Alg. for Real Closed Fields.

$\mathbb{F}(\sqrt{-1})$ has no quadratic extensions, i.e. $P \in \mathbb{F}(\sqrt{-1})[x]$ of deg 2 factor.

For any finite Galois extension $\mathbb{K}$ of $\mathbb{F}(\sqrt{-1})$ write $dim_{\mathbb{F}}\mathbb{K} = 2^n m$ (m odd).

Sylow Thm: exists subgroup H of $G := Gal(\mathbb{K}/\mathbb{F})$ with $|H| = 2^n$.

Then $[G : H] = m$. Say β generates over $\mathbb{F}$ the field $\mathbb{L}$ fixed by H . Then

minimal degree $f(x) \in \mathbb{F}[x]$ with $f(\beta) = 0$ are irreducible and $\deg f = m$,

but m being odd and $\mathbb{F}$ a real closed field $\Rightarrow$ $m = 1$ $\Rightarrow$ G is a p-group

with $p = 2$, i.e. $|G| = 2^k$.

$\mathbb{F}(\sqrt{-1})$ is Galois, so $J = Gal(\mathbb{F}(\sqrt{-1})/\mathbb{F}) \trianglelefteq G$, i.e. G/J is a group.

Basic Fact: $Gal(\mathbb{K}/\mathbb{F}(\sqrt{-1})) \cong G/J \Rightarrow |Gal(\mathbb{K}/\mathbb{F}(\sqrt{-1}))| = 2^{n-1}$.

If $n \neq 1$, "Easy Fact" imples $\exists\ N \trianglelefteq G/J$ such that $[G/J : N] = 2$

If $\mathbb{M}$ is the field fixed by N over $\mathbb{F}(\sqrt{-1})$ then $[\mathbb{M} : \mathbb{F}(\sqrt{-1})] = 2$.

But $\mathbb{F}(\sqrt{-1})$ has no quadratic extensions. Hence $n = 1$ and $\mathbb{F}(\sqrt{-1})$ is

the algebraic closure of $\mathbb{F}$, as required. $\square$

Ordered Algebraic Extensions of Ordered Fields

Corollary 2: $\mathbb{F}$ is real closed $\Longrightarrow$ $\mathbb{F}$ has no ordered algebraic extensions.

Pf: The only algebraic extension of $\mathbb{F}$ is $\mathbb{F}(\sqrt{-1})$ which

cannot be ordered since -1 is a sum of squares. $\square$

Lemma 3: If $\mathbb{F}$ is an ordered field then $\mathbb{F}$ can be extended to an ordered

field $\mathbb{K}$ with every positive element of $\mathbb{F}$ being a square.

Pf: 'Our' field $\mathbb{K}$ is generated by $\{\sqrt{c} : c \in \mathbb{F}, c > 0\}$. Indeed, -1 is not

a ss in this field. If not, then -1 is a ss in $\mathbb{F}' := \mathbb{F}(\sqrt{c_0}, ..., \sqrt{c_n})$ for some

$c_0, ..., c_n \in \mathbb{F}$. All products of distinct $\sqrt{c_i}$ form an $\mathbb{F}$-basis for $\mathbb{F}'$.

Then $-1 = (\sum_{N \subset n} b_N(\prod_{i \in N} \sqrt{c_i}))^2 = \sum_{N \subset n} b_N^2(\prod_{i \in N} c_i) >_{\mathbb{F}} 0$,

but this is a contradiction. By Lemma 2, $\mathbb{K}$ can be ordered. $\square$

Lemma 4: If $\mathbb{F}$ is an ordered field and $f(x) \in \mathbb{F}[x]$ is irreducible of odd degree then $\mathbb{F}[x]/(f(x))$ can be ordered in a compatbile way with $\mathbb{F}$.

Pf: Extend $\mathbb{F}$ to $\mathbb{K}$ from Lemma 3. Induction on $n = \frac{deg(f)-1}{2}$ (n=0 clear).

If case "n-1" $\not\Rightarrow$ case "n", let $f(\alpha) = 0$ and $\mathbb{K}(\alpha)$ cannot be ordered

$\iff \exists\, g_0, ..., g_n$ such that $\sum_{i=0}^{n} g_i(\alpha)^2 = -1$ (Corollary 1).

Equivalently $\exists\ q$ s.th $f(x)q(x) + \sum_{i=0}^{n} g_i(x)^2 = -1$ in $\mathbb{K}[x]$. WLOG we may

assume $deg(g_i) < deg(f) \Rightarrow deg(q) < (deg(f) - 2)$ and is odd.

Say $\beta \in \overline{\mathbb{K}}$ s.th. $q(\beta) = 0 \Rightarrow f(\beta)q(\beta) + \sum_{i=0}^{n} g_i(\beta)^2 = \sum_{i=0}^{n} g_i(\beta)^2 = -1$,

which contradicts the inductive assumption. Hence, $\mathbb{K}[x]/(f(x))$ can be

ordered and $\mathbb{F}[x]/(f(x))$ can be ordered by restriction. This extends $<_{\mathbb{F}}$. $\square$

Summarizing Corollary 2 and Lemmas 3, 4 we have:

Theorem 2: If $\mathbb{F}$ is an ordered field then,

$\mathbb{F}$ is real closed $\iff$ $\mathbb{F}$ has no ordered algebraic extensions $\square$

Corollary 3: Ordered fields admit algebraic real closed extensions.

Pf: Extend $\mathbb{F}$ to $\mathbb{K}$ from Lemma 3. By Zorn's Lemma, $\mathbb{K}$ has a maximal ordered algebraic extension. By theorem 2, this extension is real closed. $\square$

Concepts from Model Theory with reminder (Alex's talk):

1^{st}-order language has quantifiers "for all" $\equiv \forall$ and "there exists" $\equiv \exists$.

Def 3: The 1^{st} order language $\mathbb{L}_{OR}$ contains the following symbols,

i) The binary functions $+$, $-$ and $\cdot$

ii) A binary relation $<$

iii) The constant symbols 0 and 1

Appendix contains explicit expressions for the axioms (called **RCF**) of

real closed fields in this language.

Def 4: A **theory** for a language $\mathbb{L}$ is a set of $\mathbb{L}$-sentences.

Def 5: An $\mathbb{L}$-structure $\mathbb{M}$ is called a model of a theory $\mathbf{T}$ if $\mathbb{M} \models \Phi$

for each $\Phi \in \mathbf{T}$. In this case we write $\mathbb{M} \models \mathbf{T}$.

Recall $|\mathbb{M}|$ stands for the underlying set of the model $\mathbb{M}$.

Def 6: $\mathbb{M}$ and $\mathbb{N}$ are $\mathbb{L}$-structures $\Rightarrow$ $\mathbb{M}$ is a submodel of $\mathbb{N}$ ($\mathbb{M} \subseteq \mathbb{N}$) if

i) $|\mathbb{M}| \subseteq |\mathbb{N}|$

ii) For each n-ary function symbol $f \in \mathbb{L}$, $f^{\mathbb{N}}|_{|\mathbb{M}|} = f^{\mathbb{M}}$

iii) For each n-ary relation symbol $R \in \mathbb{L}$, $R^{\mathbb{M}} = R^{\mathbb{N}} \bigcap |\mathbb{M}|^n$

Recall that $\mathbb{M} \models \Phi[a_0, ..., a_n]$ means Φ is true of $a_0, ..., a_n$ in model $\mathbb{M}$.

Def 7: If $\mathbb{M}$ and $\mathbb{N}$ are $\mathbb{L}$-structures then $\mathbb{M}$ is an elementary submodel of $\mathbb{N}$ (we write $\mathbb{M} \preceq \mathbb{N}$) provided:

i) $\mathbb{M} \subseteq \mathbb{N}$

ii) For each formula $\phi(v_0, ..., v_n)$ and each $(a_0, ..., a_n) \in |\mathbb{M}|^{n+1}$,

$\mathbb{M} \models \Phi[a_0, ..., a_n] \iff \mathbb{N} \models \Phi[a_0, ..., a_n]$.

Def 8: Theory $\mathbf{T}$ is model-complete when for all models $\mathbb{M}, \mathbb{N} \models \mathbf{T}$,

$\mathbb{M} \subseteq \mathbb{N} \Longrightarrow \mathbb{M} \preceq \mathbb{N}$ (we say that all submodels are elementary).

Def 9: Theory $\mathbf{T}$ has **quantifier elimination** if for a formula $\Phi(v_0, ..., v_n)$

$\mathbf{T} \models (\forall v_0 \cdots \forall v_n)(\Phi \leftrightarrow \Psi)$ with $\Psi(v_0, \ldots, v_n)$ quantifier-free.

Fact: The theory **RCF** admits quantifer elimination.

Lemma 5: If $\mathbf{T}$ has quantifier elimination, then $\mathbf{T}$ is model complete.

Pf: It suffices to show that if $\Psi(v_0, \ldots, v_n)$ is quantifier free and $\mathbb{M} \subset \mathbb{N}$

then $\mathbb{M} \models \Psi[a_0, ..., a_n] \iff \mathbb{N} \models \Psi[a_0, ..., a_n]$ for all $a_0, \ldots, a_n \in |\mathbb{M}|$.

This fact is proven by induction on complexity of Ψ (details attached). $\square$

Corollary 4: **RCF** is model complete.

Return to the Main Theorem (Hilbert's 17th Problem):

$f \in \mathbb{R}[x]$ and $f(x) \geq 0$, $\forall x \in \mathbb{R}^{n+1} \Rightarrow \exists\ f_1, \ldots, f_m \in \mathbb{R}(x)$ s.th $f = \sum_{i=1}^{m} f_i^2$.

Pf: If not true, say $f(x) \geq 0$ and f is not a ss in $\mathbb{R}(x)$. Since -1 is not a ss in the field $\mathbb{R}(x)$, there is a field ordering $<_{\mathbb{R}(x)}$ (shortly $<$) s.th $f < 0$ by Lemma 1. Every positive element of $\mathbb{R}$ is a square in $\mathbb{R}(x) \Rightarrow$ ordering $<_{\mathbb{R}(x)}$ extends $<_{\mathbb{R}}$. Therefore $\mathbb{R}$ and $\mathbb{R}(x)$ are $\mathbb{L}_{OR}$-models, by interpreting the $+, \cdot$ and $<$ symbols in the obvious way.

We can extend $\mathbb{R}(x)$ to a real closed field $\mathbb{F}$ (see page 12, Cor. 3).

We now have $\mathbb{R} \subset \mathbb{F}$ so by model completeness, we have $\mathbb{R} \preceq \mathbb{F}$.

Let $m = deg(f)$. Since the coefficients of f also lie in $\mathbb{F}$, we can view it

as an element of $\mathbb{F}$ or as a degree m polynomial in $\mathbb{F}[t]$, $t = (t_0, \ldots, t_n)$.

There is a formula $\Phi(v_0, ..., v_k)$ (see appendix) s.th for a model $\mathbb{K}$ of **RCF**,

$\mathbb{K} \models \Phi(v_0, ..., v_k)[a_0, ..., a_k]$ means that polynomial $g \in \mathbb{K}[t]$ of deg m

with coefficients $a_0, ..., a_k$ takes a negative value.

Then letting $[a_0, ..., a_k]$ be the coefficients of f we have $\mathbb{F} \models \Phi[a_0, ..., a_k]$

since the elements $x_0, ..., x_n \in \mathbb{F}$ make f negative by construction.

By model completeness we can infer that

$\mathbb{R} \models \Phi[a_0, ...a_k]$ which is to say that f takes a negative value

at a point $(p_0, ..., p_n) \in \mathbb{R}^{n+1}$, contradicting our assumption.

Then it must be the case that f is in fact a sum of squares in $\mathbb{R}(x)$

so we are done. $\square$

Appendix

1) A formula $\Phi(v_0, ..., v_k)$ stating "The polynomial of degree m with coefficients $v_0, ..., v_k$ is negative for some value" we write as:

$$\exists x_0 \cdots \exists x_n (v_0 + v_1 x_o + v_2 x_1 + \cdots + v_{n+1} x_n + \cdots + v_{k-n} x_0^m + \cdots + v_k x_n^m < 0)$$

2) Real Closed Field Axioms in $\mathbb{L}_{ORF}$

Total order: i)$(\forall x)\neg(x < x)$ ii)$(\forall x)(\forall y)\neg(x < y \wedge y < x)$

iii)$(\forall x)(\forall y)(\forall z)((x < y \wedge y < z) \rightarrow (x < z)$

iv)$(\forall x)(\forall y)((x < y \vee y < x \vee x = y)$

Field axioms:

v) $(\forall x)(\forall y)(\forall z)((x+y)+z=x+(y+z))$

vi) $(\forall x)(x+0=x)$ vii)$(\forall x)(\exists y)(x+y=0)$

viii) $(\forall x)(\forall y)(x+y=y+x)$

ix) $(\forall x)(\forall y)(\forall z)((x\cdot y)\cdot z=x\cdot(y\cdot z))$

x) $(\forall x)(x\cdot 1=0)$

xi) $(\forall x)(x=0 \vee (\exists y)(x\cdot y=1))$

xii) $(\forall x)(\forall y)(x\cdot y=y\cdot x)$

Ordered Field:

xiii) $(\forall x)(\forall y)(\forall z)(x < y \to x + z < y + z)$

xiv) $(\forall x)(\forall y)(\forall z)(0 < z \to (x < y \to x \cdot z < y \cdot z))$

Real Closed Axioms:

For each odd $n \in \mathbb{N}$, 'polynomials of degree n have a root'

we write $(\forall x_0) \cdots (\forall x_n)(\exists v)(x_0 + x_1 \cdot v + \cdots + x_n \cdot v^n = 0)$

And, 'positive elements have a square root'

we write $(\forall x)(\exists y)(0 < x \to (y \cdot y = x))$

Quantifier-free formulas preserved under submodels Pf:

Case 1: Ψ is of the form $t_1 = t_2$ for terms t_1, t_2. Then,

$$\mathbb{M} \models \Psi[\bar{a}] \iff t_1^{\mathbb{M}}[\bar{a}] = t_2^{\mathbb{M}}[\bar{a}] \iff t_1^{\mathbb{N}}[\bar{a}] = t_2^{\mathbb{N}}[\bar{a}] \iff \mathbb{N} \models \Psi[\bar{a}]$$

Case 2: Ψ is of the form $t_1 < t_2$ for terms t_1, t_2. Then,

$$\mathbb{M} \models \Psi[\bar{a}] \iff t_1^{\mathbb{M}}[\bar{a}] <_{\mathbb{M}} t_2^{\mathbb{M}}[\bar{a}] \iff t_1^{\mathbb{N}}[\bar{a}] <_{\mathbb{N}} t_2^{\mathbb{N}}[\bar{a}] \iff \mathbb{N} \models \Psi[\bar{a}]$$

Case 3: Ψ is of the form $\neg\Phi$ where the result holds for Φ. Then,

$$\mathbb{M} \models \Psi[\bar{a}] \iff \text{not } \mathbb{M} \models \Phi[\bar{a}] \iff \text{not } \mathbb{N} \models \Phi[\bar{a}] \iff \mathbb{N} \models \Psi[\bar{a}]$$

Case 4: Ψ is of the form $\Phi \wedge \Theta$ where the result holds for Φ and Θ. Then,

$$\mathbb{M} \models \Psi[\bar{a}] \iff \mathbb{M} \models \Phi[\bar{a}] \text{ and } \mathbb{M} \models \Theta[\bar{a}] \iff$$

$$\mathbb{N} \models \Phi[\bar{a}] \text{ and } \mathbb{N} \models \Theta[\bar{a}] \iff \mathbb{N} \models \Psi[\bar{a}]$$

The cases of formulas built from $\vee$ and $\rightarrow$ follow by the equivalences,

i) $A \vee B \iff \neg(\neg A \wedge \neg B)$

ii) $A \rightarrow B \iff \neg(\neg B \wedge A)$