POSITIVE POLYNOMIALS LECTURE NOTES (08: 06/05/10 - BEARBEITET 28/01/19)

SALMA KUHLMANN

Contents

1. Proof of Hilbert's theorem 1

1. PROOF OF HILBERT'S THEOREM (Continued)

Theorem 1.1. (Recall Theorem 2.8 of lecture 6) (Hilbert) $\sum_{n,m} = \mathcal{P}_{n,m}$ iff

(i) $n = 2$ or

(ii) $m = 2$ or

(iii) $(n, m) = (3, 4)$.

In lecture 7 (Theorem 3.2) we showed the proof of (Hilbert's) Theorem 1.1 part (iii), i.e. for ternary quartic forms: $\mathcal{P}_{3,4} = \sum_{3,4}$ using generalization of Krein-Milman theorem (applied to our context), plus the following lemma:

Lemma 1.2. (3.1 of lecture 7) Let $T(x, y, z) \in \mathcal{P}_{3,4}$. Then $\exists$ a quadratic form $q(x, y, z) \neq 0$ s.t. $T \geq q^2$, i.e. $T - q^2$ is psd.

Proof. Consider three cases concerning the zero set of T.

Case 1. $T > 0$, i.e. T has no non trivial zeros.
Let

$$\phi(x, y, z) := \frac{T(x, y, z)}{(x^2 + y^2 + z^2)^2}, \forall\ (x, y, z) \neq 0.$$

Let $\mu := \inf_{\mathbb{S}^2} \phi \geq 0$, where $\mathbb{S}^2$ is the unit sphere.

Since $\mathbb{S}^2$ is compact and ϕ is continous, $\exists\ (a, b, c) \in \mathbb{S}^2$ s.t. $\mu = \phi(a, b, c) > 0$

Therefore $\forall\ (x, y, z) \in \mathbb{S}^2 :\ T(x, y, z) \geq \mu(x^2 + y^2 + z^2)^2$.

Claim: $T(x,y,z) \geq \mu(x^2+y^2+z^2)^2$ for all $(x,y,z) \in \mathbb{R}^3$.

Indeed, it is trivially true at the point $(0,0,0)$, and

for $(x,y,z) \in \mathbb{R}^3 \setminus \{0\}$ denote $N := \sqrt{x^2+y^2+z^2}$, then $\left(\frac{x}{N}, \frac{y}{N}, \frac{z}{N}\right) \in \mathbb{S}^2$, which implies that

$$T\left(\frac{x}{N}, \frac{y}{N}, \frac{z}{N}\right) \geq \mu\left(\left(\frac{x}{N}\right)^2 + \left(\frac{y}{N}\right)^2 + \left(\frac{z}{N}\right)^2\right)^2.$$

So, by homogeneity we get

$T(x,y,z) \geq \mu(x^2+y^2+z^2)^2 = \left(\sqrt{\mu}(x^2+y^2+z^2)\right)^2$, as claimed.

$\square$(**Case1**)

<u>**Case 2.**</u> T has exactly one (nontrivial) zero.

By changing coordinates, we may assume w.l.o.g. that zero to be $(1,0,0)$, i.e. $T(1,0,0) = 0$.

Writing T as a polynomial in x one gets

$$T(x,y,z) = ax^4 + (b_1y + b_2z)x^3 + f(y,z)x^2 + 2g(y,z)x + h(y,z),$$

where f, g and h are binary quadratic, cubic and quartic forms respectively.

<u>Reducing T</u>: Since $T(1,0,0) = 0$ we get $a = 0$.

Further, suppose $(b_1,b_2) \neq (0,0)$, it $\Rightarrow \exists\ (y_0,z_0) \in \mathbb{R}^2$ s.t $b_1y_0 + b_2z_0 < 0$, then

taking x big enough $\Rightarrow T(x_0,y_0,z_0) < 0$, a contradiction to $T \geq 0$. Thus $b_1 = b_2 = 0$ and therefore

$$T(x,y,z) = f(y,z)x^2 + 2g(y,z)x + h(y,z) \tag{1}$$

Next, clearly $h(y,z) \geq 0$ $\big[$since otherwise $T(0,y_0,z_0) = h(y_0,z_0) < 0$ for some $(y_0,z_0) \in \mathbb{R}^2$, a contradiction$\big]$.

Also $f(y,z) \geq 0$, if not, say $f(y_0,z_0) < 0$ for some (y_0,z_0), then taking x big enough we get $T(x,y_0,z_0) < 0$, a contradiction.

Thus $f,h \geq 0$.

From (1) we can write:

$$fT(x,y,z) = (xf+g)^2 + (fh - g^2) \tag{2}$$

Claim: $fh - g^2 \geq 0$

If not, say $(fh-g^2)(y_0,z_0) < 0$ for some (y_0,z_0). Then there are two cases to be considered here:

Case (i): $f(y_0,z_0) = 0$. In this case we claim $g(y_0,z_0) = 0$ because if not then $T(x,y_0,z_0) = 2g(y_0,z_0)x + h(y_0,z_0) < 0$ and we take $|x_0|$ large enough

so that $2g(y_0,z_0)x_0 + h(y_0,z_0) < 0$, a contradiction.
Case (ii): $f(y_0,z_0) > 0$, we take $|x_0|$ such that $x_0 f(y_0,z_0) + g(y_0,z_0) = 0$, then $fT(x_0,y_0,z_0) = (fh - g^2)(y_0,z_0) < 0$, a contradiction.

So our claim is established and $fh - g^2 \geq 0$.

Now the polynomial f is a psd binary form, thus by Lemma 1.3 below f is sum of two squares. Let us consider the two subcases:

__**Case 2.1.**__ f is a perfect square. Then $f = f_1^2$, with $f_1 = by + cz$ for some $b, c \in \mathbb{R}$. Up to multiplication by a constant $(-c, b)$ is the unique zero of f_1 and so of f. Thus

$$(fh - g^2)(-c, b) = -(g(-c,b))^2 \leq 0$$

which is a contradiction unless $g(-c,b) = 0$ which means [1] that $f_1 \mid g$, i.e. $g(y,z) = f_1(y,z)g_1(y,z)$. Then from (2) we get

$$\begin{aligned} fT &\geq (xf + g)^2 \\ &= (xf_1{}^2 + f_1 g_1)^2 \\ &= f_1{}^2(xf_1 + g_1)^2 \\ &= f(xf_1 + g_1)^2. \end{aligned}$$

Hence $T \geq (xf_1 + g_1)^2$ as required.

__**Case 2.2.**__ $f = f_1^2 + f_2^2$, with f_1, f_2 linear in y, z.
Now $f_1 \not\equiv \lambda f_2$ [otherwise we are in **Case 2.1**]
i.e. f_1, f_2 don't have same non-trivial zeroes, otherwise they would be multiples of each other and f would be a perfect square. Hence $f > 0$.

Claim 1: $fh - g^2 > 0$
If not, i.e. if $\exists\ (y_0,z_0) \neq (0,0)$ s.t. $(fh - g^2)(y_0,z_0) = 0$, then (y_0,z_0) could be completed to a zero $\Big(-\dfrac{g(y_0,z_0)}{f(y_0,z_0)}, y_0, z_0\Big)$ of T, which contradicts our hypothesis that T has only 1 zero $(1,0,0)$. Thus $fh - g^2 > 0$.

Claim 2: $\dfrac{fh - g^2}{f^3}$ has a minimum $\mu > 0$ on the unit circle $\mathbb{S}^1$. (clear)

So, just as in **Case 1**,

$fh - g^2 \geq \mu f^3,\ \forall\ (y,z) \in \mathbb{R}^2$.

$\Rightarrow fT \geq fh - g^2 \geq \mu f^3$, by (2)

[1] See (5) implies (2) of Theorem 4.5.1 in *Real Algebraic Geometry* by J. Bochnak, M. Coste, M.-F. Roy or (5) implies (2) of Theorem 12.7 in *Positive Polynomials and Sum of Squares* by M. Marshall.

$\Rightarrow T \geq \mu f^2 = \left(\sqrt{\mu} f\right)^2$, as claimed. □(**Case 2**)

Case 3. T has more than one zero.

Without loss of generality, assume $(1,0,0)$ and $(0,1,0)$ are two of the zeros of T.

As in case 2, reduction $\Rightarrow$ T is of degree at most 2 in x as well as in y and so we can write:

$$T(x,y,z) = f(y,z)x^2 + 2g(y,z)zx + z^2 h(y,z),$$

where f, g, h are quadratic forms and $f, h \geq 0$.

And so

$$fT = (xf + zg)^2 + z^2(fh - g^2), \tag{3}$$

with $fh - g^2 \geq 0$ [Indeed, if $(fh - g^2)(y_0, z_0) < 0$ for some (y_0, z_0), then we must have case distinction as on bottom of page 2 i.e. $f(y_0, z_0) = 0$ or $f(y_0, z_0) > 0$].

Using Lemma 1.3 if f or h is a perfect square, then we get the desired result as in the **Case 2.1**. Hence we suppose f and h to be sum of two squares and again as before (as in **Case 2.2**) $f, h > 0$. We consider the following two possible subcases on $fh - g^2$:

Case 3.1. Suppose $fh - g^2$ has a zero $(y_0, z_0) \neq (0,0)$.

Set $x_0 = -\dfrac{g(y_0, z_0)}{f(y_0, z_0)}$ and

$$T_1 := T(x + x_0 z, y, z) = x^2 f + 2xz(g + x_0 f) + z^2(h + 2x_0 g + x_0^2 f) \tag{4}$$

Evaluating (3) at $(x + x_0 z, y, z)$, we get

$$fT_1 = fT(x + x_0 z, y, z) = \Big((x + x_0)f + zg\Big)^2 + z^2(fh - g^2), \qquad (3)'$$

Multyplying (4) by f, we get

$$fT_1 = fT(x + x_0 z, y, z) = x^2 f^2 + 2xzf(g + x_0 f) + z^2 f(h + 2x_0 g + x_0^2 f) \qquad (4)'$$

Now compare the coefficients of z^2 in $(3)'$ and $(4)'$ to get

$$(x_0 f + g)^2 + (fh - g^2) = f(h + 2x_0 g + x_0^2 f),$$

i.e. $h + 2x_0 g + x_0^2 f = \dfrac{(fh - g^2) + (x_0 f + g)^2}{f} \quad \forall\ (y,z) \neq (0,0)$

In particular, $h + 2x_0 g + x_0^2 f$ is psd and has a zero, namely $(y_0, z_0) \neq (0,0)$.

Thus $(h + 2x_0 g + x_0^2 f)$, being a psd quadratic in y, z, which has a nontrivial zero (y_0, z_0), is a perfect square [since by the arguments similar to **Case 2.2**, it cannot be a sum of two (or more) squares].

Say $(h + 2x_0 g + x_0^2 f) = h_1^2$, with $h_1(y, z)$ linear and $h_1(y_0, z_0) = 0$

Now $(g + x_0 f)(y_0, z_0) = g(y_0, z_0) + x_0 f(y_0, z_0) = 0$. So, $g + x_0 f$ vanishes at every zero of the linear form h_1. Therefore, we have $g + x_0 f = g_1 h_1$ for some g_1.

So (from (4)), $T_1 = fx^2 + 2xzg_1h_1 + z^2h_1^2$

$$= (zh_1 + xg_1)^2 + x^2(f - g_1^2)$$

$$\Rightarrow h_1^2 T_1 = h_1^2(zh_1 + xg_1)^2 + x^2\big(h_1^2 f - (h_1 g_1)^2\big)$$

$$= h_1^2(zh_1 + xg_1)^2 + x^2 \underbrace{(hf - g^2)}_{\geq\, 0}$$

$$\Rightarrow h_1^2 T_1 \geq h_1^2(zh_1 + xg_1)^2$$

$$\Rightarrow T(x + x_0 z, y, z) =: T_1 \geq (zh_1 + xg_1)^2$$

By change of variables $(x \to x - x_0 z)$, we get $T \geq$ a square of a quadratic form, as desired.

Case 3.2. Suppose $fh - g^2 > 0$ (i.e. $fh - g^2$ has no zero).

Then (as in **Case 2.2**), $\exists\ \mu > 0$ s.t $\dfrac{fh - g^2}{(y^2 + z^2)f} \geq \mu$ on $\mathbb{S}^1$

and so $fh - g^2 \geq \mu(y^2 + z^2)f\ \forall\ (y, z) \in \mathbb{R}^2$.

Hence, by (3) we get

$$fT = (xf + zg)^2 + z^2 \underbrace{(fh - g^2)}_{>0}$$

$$\geq z^2(fh - g^2)$$

$$\geq \mu z^2(y^2 + z^2)f,$$

giving as required

$$T \geq (\sqrt{\mu} zy)^2 + (\sqrt{\mu} z^2)^2$$

$$\Rightarrow T \geq (\sqrt{\mu} z^2)^2 \qquad \square(\textbf{Case 3})$$

This completes the proof of the Lemma 1.2. □□

Next we prove Theorem 1.1 part (i), i.e. for binary forms. This was also used as a helping lemma in the proof of above lemma:

Lemma 1.3. If f is a binary psd form of degree m, then f is a sum of squares of binary forms of degree $m/2$, that is, $\mathcal{P}_{2,m} = \sum_{2,m}$. In fact, f is sum of two squares.

Proof. If f is a binary form of degree m, we can write

$$f(x, y) = \sum_{k=0}^{m} c_k x^k y^{m-k}; \; c_k \in \mathbb{R}$$
$$= y^m \sum_{k=0}^{m} c_k \left(\frac{x}{y}\right)^k,$$

where m is an even number and $c_m \neq 0$, since f is psd.

Without loss of generality let $c_m = 1$.

Put $g(t) = \sum_{k=0}^{m} c_k t^k$.

Over $\mathbb{C}$, $g(t) = \prod_{k=1}^{m/2} (t - z_k)(t - \overline{z}_k); \;\; z_k = a_k + ib_k, a_k, b_k \in \mathbb{R}$

$$= \prod_{k=1}^{m/2} \Big((t - a_k)^2 + b_k^2\Big)$$

$$\Rightarrow f(x, y) = y^m g\Big(\frac{x}{y}\Big) = \prod_{k=1}^{m/2} \Big((x - a_k y)^2 + b_k^2 y^2\Big).$$

Then, using iteratively the identity

$$(X^2 + Y^2)(Z^2 + W^2) = (XZ - YW)^2 + (YZ + XW)^2,$$

we obtain that $f(x, y)$ is a sum of two squares. □

Example 1.4. Using the ideas in the proof of above lemma, we write the binary form

$$f(x, y) = 2x^6 + y^6 - 3x^4 y^2$$

as a sum of two squares:

Consider f written in the form

$$f(x, y) = y^6 \left(2\Big(\frac{x}{y}\Big)^6 + 1 - 3\Big(\frac{x}{y}\Big)^4\right).$$

The polynomial $g(t) = 2t^6 - 3t^4 + 1$. This polynomial has double roots 1 and

-1 and complex roots $\pm\frac{1}{\sqrt{2}}i$.

Thus

$$g(t) = 2(t-1)^2(t+1)^2(t^2+\frac{1}{2}) = (t^2-1)^2(2t^2+1).$$

Therefore, we have

$$f(x,y) = y^6 g\Big(\frac{x}{y}\Big) = (x^2-y^2)^2(2x^2+y^2) = 2x^2(x^2-y^2)^2 + y^2(x^2-y^2)^2$$

written as a sum of two squares. □

Next we prove Theorem 1.1 part (ii), i.e. for quadratic forms:

Lemma 1.5. If $f(x_1,\ldots,,x_n)$ is a psd quadratic form, then $f(x_1,\ldots,,x_n)$ is sos of linear forms, that is, $\mathcal{P}_{n,2} = \sum_{n,2}$.

Proof. If $f(x_1,\ldots,x_n)$ is a quadratic form, then we can write

$f(x_1,\ldots,x_n) = \sum_{i,j=1}^{n} x_i a_{ij} x_j$, where $A = [a_{ij}]$ is a symmetric matrix with $a_{ij} \in \mathbb{R}$.

We have $f = X^T A X$, where $X^T = [x_1, \ldots x_n]$.

By the spectral theorem for Hermitian matrices, there exists a real orthogonal matrix S and a diagonal matrix $D = \mathrm{diag}(d_1,\ldots,d_n)$ such that $D = S^T A S$. Then

$$f = X^T S S^T A\ S S^T X = (S^T X)^T S^T A\ S\ (S^T X).$$

Putting $Y = [y_1,\ldots,y_n]^T = S^T X$, we get

$$f = Y^T S^T A\ SY = Y^T D\ Y = \sum_{i=1}^{n} d_i {y_i}^2, d_i \in \mathbb{R}\ .$$

Since f is psd, we have $d_i \geq 0\ \forall\ i$, and so

$$f = \sum_{i=1}^{n} \left(\sqrt{d_i} y_i\right)^2.$$

Thus,

$$f(x_1,\ldots,x_n) = \sum_{i=1}^{n} \left(\sqrt{d_i}(s_{1,i}x_1 + \ldots + s_{n,i}x_n)\right)^2,$$

that is, f is sos of linear forms. □