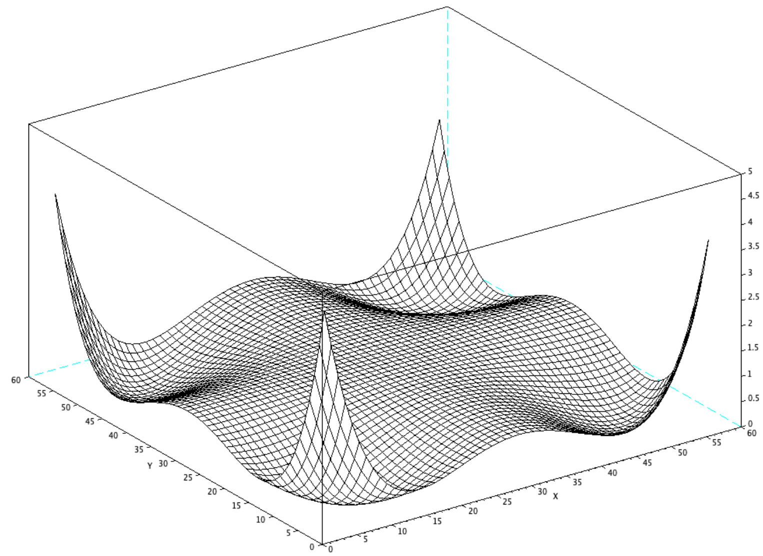
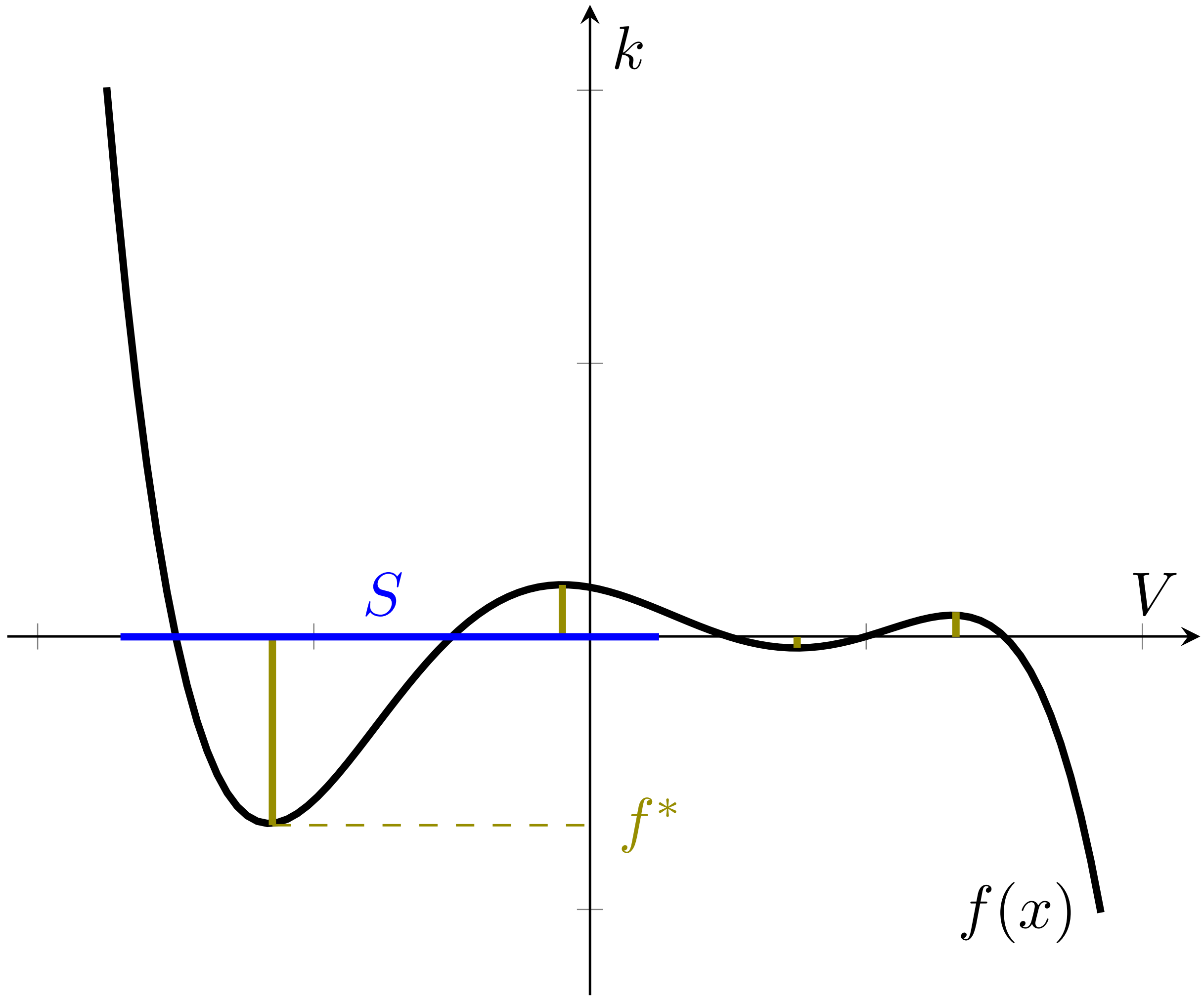
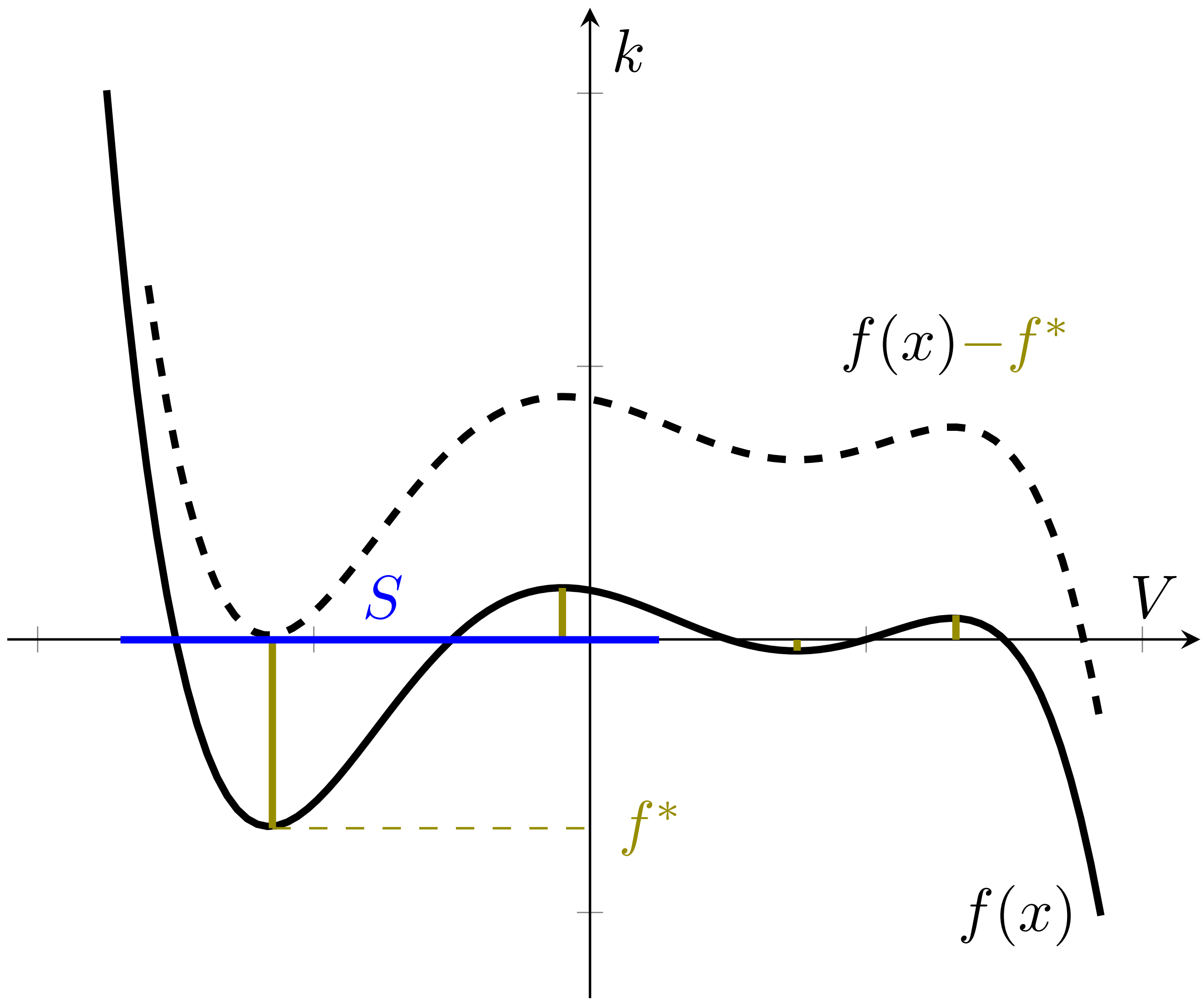
What is polynomial optimization?

Simone Naldi

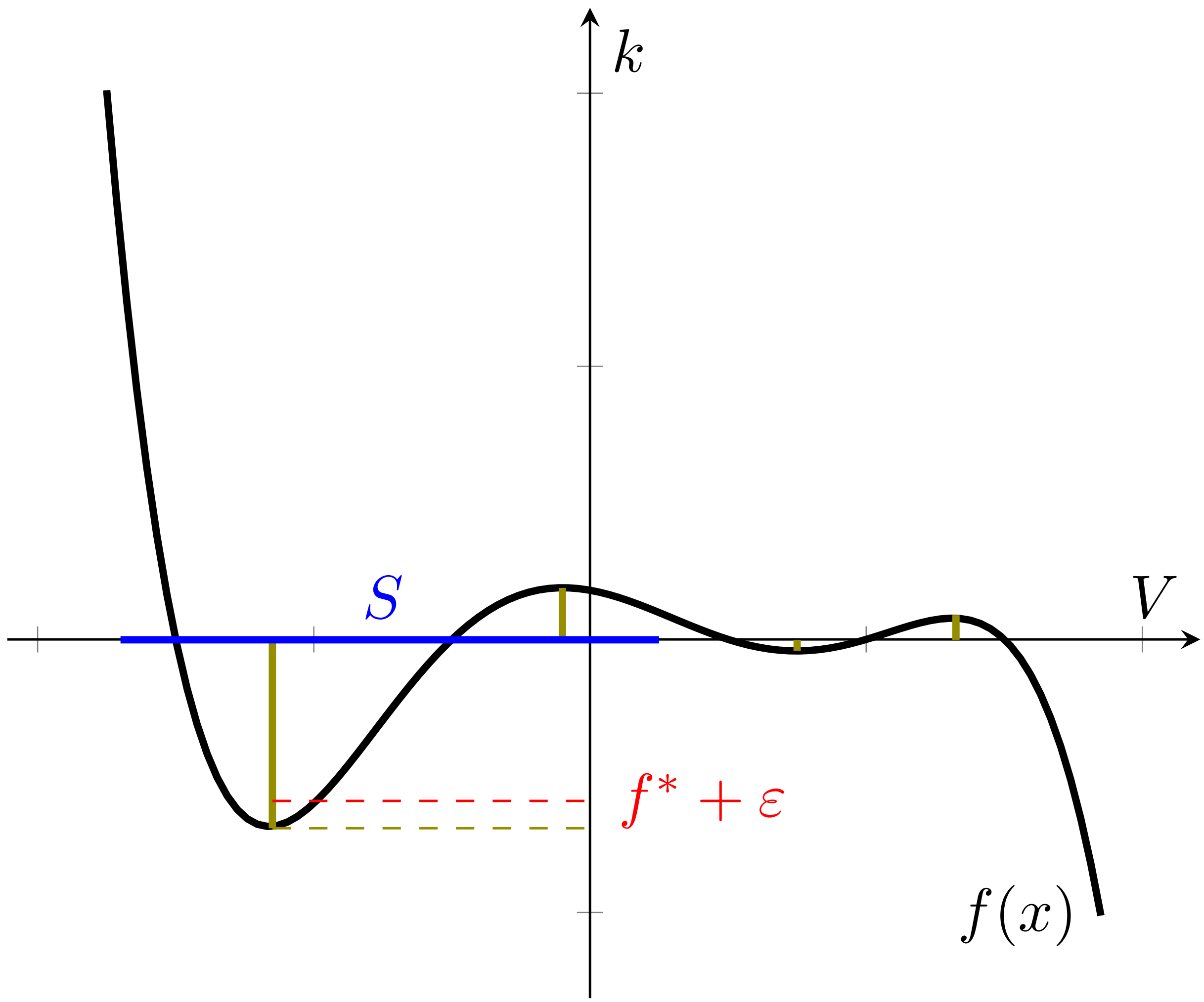
POEMA Meeting, Florence, 15-17/01/2020

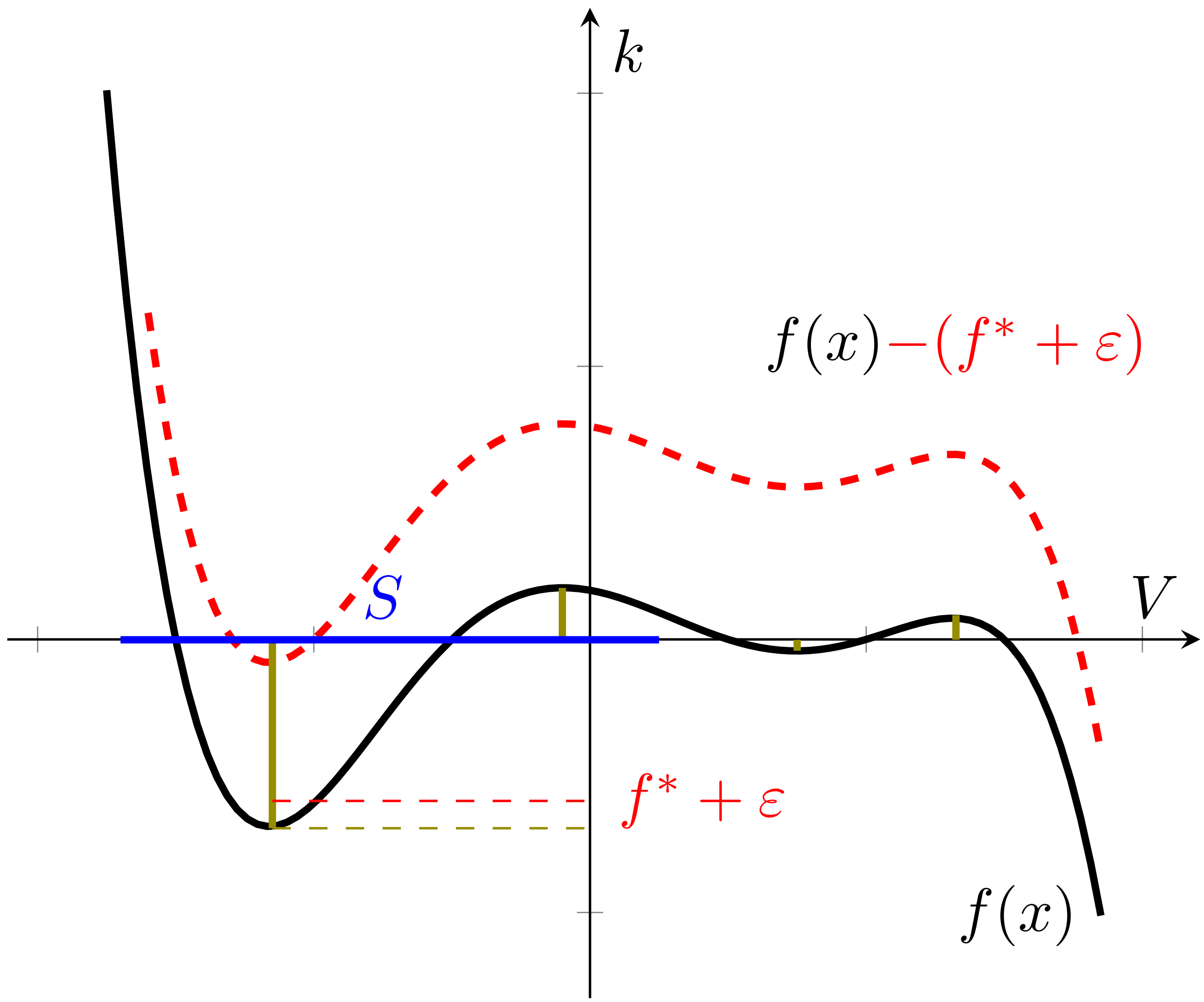



k
V
S
f*
f(x)



k
$f(x) - f^*$
V
S
f^*
$f(x)$


k
S
V
$f^* + \varepsilon$
$f(x)$



k
$f(x) - (f^* + \varepsilon)$
S
V
$f^* + \varepsilon$
$f(x)$

Optimization = Positivity

For $f : V \to k$, k with an order, $S \subset V$, then

$$\boxed{\begin{array}{rl} f^* := & \text{inf} \;\; f(x) \\ & \text{s.t.} \;\; x \in S \end{array}} \quad = \quad \boxed{\begin{array}{rl} \lambda^* := & \text{sup} \;\; \lambda \\ & \text{s.t.} \;\; f - \lambda \in \mathscr{P}(S) \end{array}}$$

where $\mathscr{P}(S) = \{g : V \to k \mid g(x) \geq 0 \; \forall \, x \in S\}$.

In other words, minimizing a function f over $S \subset V$ means characterizing those functions that are positive over S.

Problem : $\mathscr{P}(S)$ is hard to characterize since it is in general not "tractable", hence one strategy is to approximate it.

Polynomial optimization and SOS

Let $k = \mathbb{R}$, $V = \mathbb{R}^n$, $f, g_1, \ldots, g_m \in \mathbb{R}[x]$:

$$\begin{array}{rcll} f^* & = & \text{inf} & f(x) \\ & & \text{s.t.} & x \in S := \{a \in \mathbb{R}^n : g_i(a) \geq 0,\ i = 1, \ldots, m\} \end{array}$$

For $S = \mathbb{R}^n$ (global case) lower bounds are given by SOS :

$$\begin{array}{rcllcll} f^* & = & \text{sup} & \lambda & \geq & \text{sup} & \lambda \\ & & \text{s.t.} & f - \lambda \in \mathcal{P}(\mathbb{R}^n) & & \text{s.t.} & f - \lambda \in \boldsymbol{\Sigma} \subset \mathcal{P}(\mathbb{R}^n) \end{array}$$

SOS relaxations give lower bounds for polynomial optimization, and are “explicit” and “easier” to “solve”.

Semidefinite programming

In contrast to $\mathscr{P}(\mathbb{R}^n)$, the set Σ is tractable (in degree d): it is the projection of a section of the cone $\mathcal{S}_D^+$ of psd matrices.

Indeed, a $2d-$degree form $f \in \Sigma$ iff there is $X \in \mathcal{S}_D^+$ s.t.

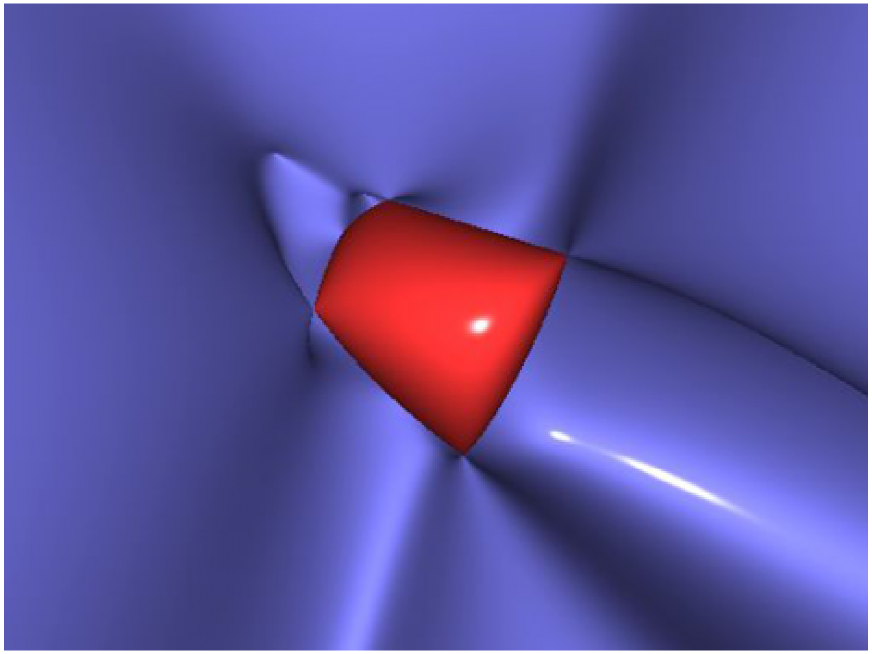
$$f = v^T X v \quad \text{where } \langle v \rangle = \mathbb{R}[x]_d \text{ and } D = \binom{n+d}{d}$$

Hence deciding membership in Σ is a feasibility problem in the convex cone $\mathcal{S}_D^+$ of psd matrices :

$$\begin{array}{ll} \text{inf} & \langle C, X \rangle := \text{Trace}(CX) \\ \text{s.t.} & \langle A_i, X \rangle = b_i \\ \text{and} & X \succeq 0 \end{array}$$

Despite the easy definition, there are open questions related to its complexity status, and the problem of solving it in practice.

The POEMA spectrahedron



An example of application

A method for sampling the real solutions of the poly. system

$$f_1(x) = 0,\ f_2(x) = 0,\ \ldots,\ f_m(x) = 0$$

is based on polynomial optimization. For a sufficiently generic polynomial function f, the minimizers of

$$\begin{array}{rcll} f^* & = & \text{inf} & f(x) \\ & & \text{s.t.} & x \in V_{\mathbb{R}}(f_1, \ldots, f_m) \end{array}$$

meet every connected component of $V_{\mathbb{R}}(f_1, \ldots, f_m)$.

This "critical point method", related to the construction of polar varieties, has led to better complexity estimates (more on Mohab's course).

Univariate polynomials – 1

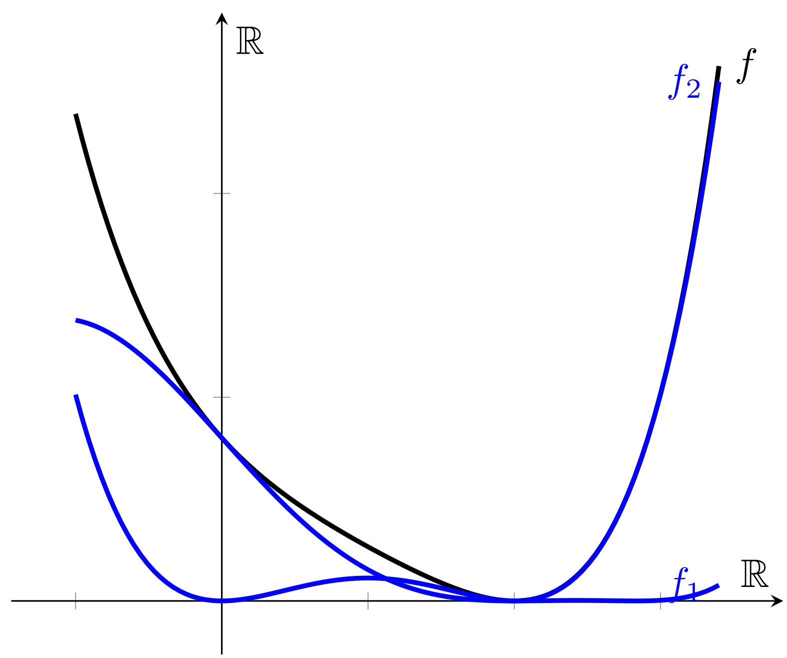
Fix $S = V = \mathbb{R}$, $k = \mathbb{R}$, and let $g \in \mathbb{R}[x]$

A polynomial is positive on $\mathbb{R}$ iff it is a sum of two squares:

$$g(x) \geq 0 \;\forall\, x \quad \Longleftrightarrow \quad g = r \cdot c \cdot \overline{c} \quad \Longleftrightarrow \quad g = g_1^2 + g_2^2$$

For instance

$$x^6 - 2x^5 + 6x^4 - 10x^3 + 9x^2 - 8x + 4$$
$$= (x^3 - x^2 - 2x + 2)^2 + (3x^2 - 3x)^2$$



SOS proof of the Cauchy-Scwharz inequality

Sums of squares can be used to prove inequalities, for instance:

For real vectors $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ one has:

$$\left|\sum x_i y_i\right| \leq \sqrt{\left(\sum x_i^2\right)\left(\sum y_i^2\right)}$$

Indeed, the polynomial

$$\left(\sum x_i^2\right)\left(\sum y_i^2\right) - \left(\sum x_i y_i\right)^2 = \sum_{i<j} \left(x_i y_j - x_j y_i\right)^2$$

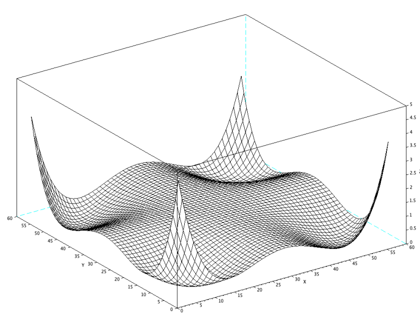
More generally, techniques based on sums of squares are used in formal proofs of semialgebraic inequalities.

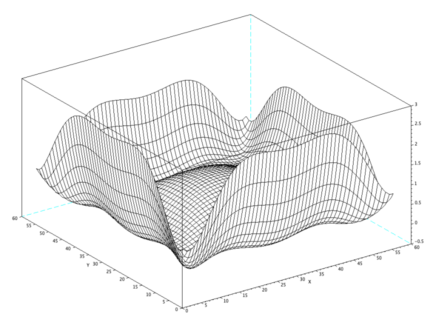
Explicit non-SOS positive polynomials

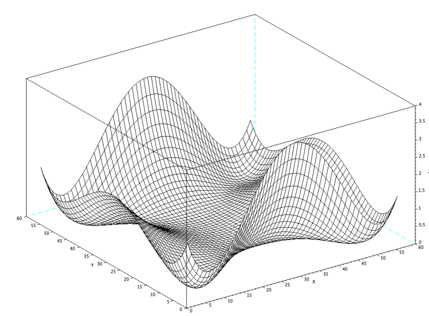
Motzkin : $M(x, y) = x^4y^2 + x^2y^4 + 1 - 3x^2y^2$

Robinson : $R(x, y) = x^6 + y^6 + 1 - (x^4y^2 + x^2y^4 + x^4 + x^2 + y^4 + y^2) + 3x^2y^2$

Choi-Lam : $CL(x, y) = x^4y^2 + y^4 + x^2 - 3x^2y^2$







Global certificates

The problem whether, and when, every polynomial is SOS (of polynomials) was fully solved by Hilbert in 1888

Every positive polynomial in n variables of degree $2d$ is a sum of squares if and only if $n = 1$ or $2d = 2$ or $(n, 2d) = (2, 4)$

Artin (1927) proved for the positive the 17th Hilbert problem :

Every positive polynomial is a SOS of rational functions

$$g(x) \geq 0, \ \forall x \in \mathbb{R} \quad \Longleftrightarrow \exists h, g_1, \ldots, g_m \in \mathbb{R}[x], \quad h^2 g = \sum_i g_i^2$$

Motzkin polynomial SOS in $\mathbb{R}(x, y, z)$

The Motzkin polynomial M has a SOS-representation in $\mathbb{R}(x, y, z)$

$$(x^2 + y^2) \cdot M = (y(1 - x^2))^2 + (x(1 - y^2))^2 + (xy(x^2 + y^2 - 2))^2$$

But in general, the SOS representation in $\mathbb{R}(x)$ requires high degree for the SOS-multipliers. A general bound has been given only recently by Lombardi-Perrucci-Roy :

The degrees of the numerator and the denominator in the Artin representation are bounded from above by $2^{2^{2^{d^{4^k}}}}$

Approximations of $\mathscr{P}(S)$

In order to have local positivity certificates, and lower bounds for polynomial optimization, one needs to approximate $\mathscr{P}(S)$ for a general S, so in some sense to generalize Σ.

One can construct explicit subsets of $\mathscr{P}(S)$, such as the quadratic module

$$M(g_1,\ldots,g_m) = \Sigma + g_1\Sigma + \cdots + g_m\Sigma$$

and the preordering

$$P(g_1,\ldots,g_m) = \Sigma + g_1\Sigma + \cdots + g_m\Sigma + g_1g_2\Sigma + g_1g_3\Sigma + \cdots$$

Truncations of M and P in degree d approximate $\mathscr{P}(S) \cap \mathbb{R}[x]_d$, and can still be represented using semidefinite programming.

Local certificates : Positivstellensätze

Local certificates are known in some cases, e.g. for intervals:

A univariate polynomial $f \in \mathbb{R}[x]$ is nonnegative on $[-1,1]$ if and only if $f = \sigma_0 + (1-x^2)\sigma_1, \ \ \sigma_i \in \Sigma$

More generally, for semialgebraic sets S satisfying some assumptions, one can get representation theorems for $f \in \mathscr{P}(S)$:

Schmüdgen PSS. If S is compact, then $P(g_1, \ldots, g_m)$ contains all polynomials that are strictly positive on S.

Putinar PSS. If $M(g_1, \ldots, g_m)$ is archimedean, then it contains all polynomials that are strictly positive on S.

Limits of SOS and beyond

Results by Blekherman show quantitative relationship between positive polynomials and sums of squares :

> There are much more nonnegative polynomials
>
> than sums of squares, for $n \gg 1$.

Recently new types of certificates, that is, new approximations of the cone $\mathscr{P}(S)$ have been proposed, to cite a few :

- SONC - Sums Of Nonnegative Circuits (Iliman, de Wolff)
  based on geometric programming

- Hyperbolic certificates of positivity (Saunderson)
  based on hyperbolic programming

Algebraic degree of polynomial optimization

The minimizers in a polynomial optimization problem

$$\begin{array}{rcll} f^* & = & \text{inf} & f(x) \\ & & \text{s.t.} & g_i(x) \geq 0,\ i = 1, \ldots, m \end{array}$$

are vectors of *algebraic* numbers, and there are formulas for upper bounds for their degree ($d = \deg f, d_i = \deg g_i$):

Nie-Ranestad. The algebraic degree of polynomial optimization is bounded from above by $d_1 d_2 \cdots d_m D_{n-m}(d-1, d_1 - 1, \ldots, d_m - 1)$. if $g_1, \ldots, g_m$ form a complete intersection.

This is a measure of the "output complexity".