

## 8. Sum of Squares

- Polynomial nonnegativity
- Sum of squares (SOS) decomposition
- Example of SOS decomposition
- Computing SOS using semidefinite programming
- Necessary conditions
- Newton Polytopes and Sparsity
- Positivity in one variable
- Background
- Global optimization
- Optimizing in parameter space
- Lyapunov functions

# Polynomial Programming

So far

- Polynomial equations over the complex field

Objectives

- General quantified formulae
- Boolean connectives
- Polynomial equations, inequalities, and inequations over the reals

e.g., does there exist  $x$  such that for all  $y$

$$(f(x, y) \geq 0) \wedge (g(x, y) = 0) \vee (h(x, y) \neq 0)$$

## Polynomial Nonnegativity

First, consider the case of one inequality; given  $f \in \mathbb{R}[x_1, \dots, x_n]$

does there exist  $x \in \mathbb{R}^n$  such that  $f(x) < 0$

- If not, then  $f$  is globally non-negative

$$f(x) \geq 0 \quad \text{for all } x \in \mathbb{R}^n$$

and  $f$  is called *positive semidefinite* or *PSD*

- The problem is *NP-hard*, but decidable
- Many applications

# Certificates

## Primal decision problem

does there exist  $x \in \mathbb{R}^n$  such that  $f(x) < 0$

- Answer *yes* is easy to verify; exhibit  $x$  such that  $f(x) < 0$
- Answer *no* is hard; we need a *certificate* or a *witness*  
i.e, a proof that there is no feasible point

## Sum of Squares Decomposition

If there are polynomials  $g_1, \dots, g_s \in \mathbb{R}[x_1, \dots, x_n]$  such that

$$f(x) = \sum_{i=1}^s g_i^2(x)$$

then  $f$  is nonnegative

A purely syntactic, easily checkable certificate, called a *sum-of-squares (SOS)* decomposition

- How do we find the  $g_i$
- When does such a certificate exist?

## Example

We can write any polynomial as a *quadratic form* on monomials

$$\begin{aligned}
 f &= 4x^4 + 4x^3y - 7x^2y^2 - 2xy^3 + 10y^4 \\
 &= \begin{bmatrix} x^2 \\ xy \\ y^2 \end{bmatrix}^T \begin{bmatrix} 4 & 2 & -\lambda \\ 2 & -7 + 2\lambda & -1 \\ -\lambda & -1 & 10 \end{bmatrix} \begin{bmatrix} x^2 \\ xy \\ y^2 \end{bmatrix} \\
 &= z^T Q(\lambda) z
 \end{aligned}$$

which holds for all  $\lambda \in \mathbb{R}$

If for some  $\lambda$  we have  $Q(\lambda) \succeq 0$ , then we can factorize  $Q(\lambda)$

## Example, continued

e.g., with  $\lambda = 6$ , we have

$$Q(\lambda) = \begin{bmatrix} 4 & 2 & -6 \\ 2 & 5 & -1 \\ -6 & -1 & 10 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ 2 & 1 \\ 1 & -3 \end{bmatrix} \begin{bmatrix} 0 & 2 & 1 \\ 2 & 1 & -3 \end{bmatrix}$$

so

$$\begin{aligned} f &= \begin{bmatrix} x^2 \\ xy \\ y^2 \end{bmatrix}^T \begin{bmatrix} 0 & 2 \\ 2 & 1 \\ 1 & -3 \end{bmatrix} \begin{bmatrix} 0 & 2 & 1 \\ 2 & 1 & -3 \end{bmatrix} \begin{bmatrix} x^2 \\ xy \\ y^2 \end{bmatrix} \\ &= \left\| \begin{bmatrix} 2xy + y^2 \\ 2x^2 + xy - 3y^2 \end{bmatrix} \right\|^2 \\ &= (2xy + y^2)^2 + (2x^2 + xy - 3y^2)^2 \end{aligned}$$

which is an SOS decomposition

## Sum of Squares and Semidefinite Programming

Suppose  $f \in \mathbb{R}[x_1, \dots, x_n]$ , of degree  $2d$

Let  $z$  be a vector of all monomials of degree less than or equal to  $d$

$f$  is SOS if and only if there exists  $Q$  such that

$$\begin{aligned} Q &\succeq 0 \\ f &= z^T Q z \end{aligned}$$

- This is an SDP in standard primal form
- The number of components of  $z$  is  $\binom{n+d}{d}$
- Comparing terms gives affine constraints on the elements of  $Q$

## Sum of Squares and Semidefinite Programming

If  $Q$  is a feasible point of the SDP, then to construct the SOS representation factorize  $Q = VV^T$ , and write  $V = [v_1 \dots v_r]$ , so that

$$\begin{aligned} f &= z^T VV^T z \\ &= \|V^T z\|^2 \\ &= \sum_{i=1}^r (v_i^T z)^2 \end{aligned}$$

- One can factorize using e.g., Cholesky or eigenvalue decomposition
- The number of squares  $r$  equals the rank of  $Q$

## Example

$$\begin{aligned}
 f &= 2x^4 + 2x^3y - x^2y^2 + 5y^4 \\
 &= \begin{bmatrix} x^2 \\ xy \\ y^2 \end{bmatrix}^T \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{12} & q_{22} & q_{23} \\ q_{13} & q_{23} & q_{33} \end{bmatrix} \begin{bmatrix} x^2 \\ xy \\ y^2 \end{bmatrix} \\
 &= q_{11}x^4 + 2q_{12}x^3y + (q_{22} + 2q_{13})x^2y^2 + 2q_{23}xy^3 + q_{33}y^4
 \end{aligned}$$

So  $f$  is SOS if and only if there exists  $Q$  satisfying the SDP

$$\begin{aligned}
 Q \succeq 0 & & q_{11} &= 2 & & 2q_{12} &= 2 \\
 & & 2q_{12} + q_{22} &= -1 & & 2q_{23} &= 0 \\
 & & & & & q_{33} &= 5
 \end{aligned}$$

## Convexity

The sets of PSD and SOS polynomials are *convex cones*; i.e.,

$$f, g \text{ PSD} \quad \implies \quad \lambda f + \mu g \text{ is PSD for all } \lambda, \mu \geq 0$$

let  $P_{n,d}$  be the set of PSD polynomials of degree  $\leq d$

let  $\Sigma_{n,d}$  be the set of SOS polynomials of degree  $\leq d$

- Both  $P_{n,d}$  and  $\Sigma_{n,d}$  are *convex cones* in  $\mathbb{R}^N$  where  $N = \binom{n+d}{d}$
- We know  $\Sigma_{n,d} \subseteq P_{n,d}$ , and testing if  $f \in P_{n,d}$  is NP-hard
- But testing if  $f \in \Sigma_{n,d}$  is an SDP

## Polynomials in One Variable

If  $f \in \mathbb{R}[x]$ , then  $f$  is SOS if and only if  $f$  is PSD

### Example

All real roots must have even multiplicity, and highest coeff. is positive

$$\begin{aligned} f &= x^6 - 10x^5 + 51x^4 - 166x^3 + 342x^2 - 400x + 200 \\ &= (x - 2)^2 (x - (2 + i)) (x - (2 - i)) (x - (1 + 3i)) (x - (1 - 3i)) \end{aligned}$$

Now reorder complex conjugate roots

$$\begin{aligned} &= (x - 2)^2 (x - (2 + i)) (x - (1 + 3i)) (x - (2 - i)) (x - (1 - 3i)) \\ &= (x - 2)^2 ((x^2 - 3x - 1) - i(4x - 7)) ((x^2 - 3x - 1) + i(4x - 7)) \\ &= (x - 2)^2 ((x^2 - 3x - 1)^2 + (4x - 7)^2) \end{aligned}$$

So every PSD scalar polynomial is the sum of *two* squares

## Quadratic Polynomials

A quadratic polynomial in  $n$  variables is PSD if and only if it is SOS

Because it is PSD if and only if

$$f = x^T Q x$$

where  $Q \geq 0$

And it is SOS if and only if

$$\begin{aligned} f &= \sum_i (v_i^T x)^2 \\ &= x^T \left( \sum_i v_i v_i^T \right) x \end{aligned}$$

## Some Background

In 1888, Hilbert showed that PSD=SOS if and only if

- $d = 2$ , i.e., quadratic polynomials
- $n = 1$ , i.e., univariate polynomials
- $d = 4, n = 2$ , i.e., quartic polynomials in two variables

$n \backslash d$	2	4	6	8
1	yes	yes	yes	yes
2	yes	yes	no	no
3	yes	no	no	no
4	yes	no	no	no

- In general  $f$  is PSD does not imply  $f$  is SOS

## Some Background

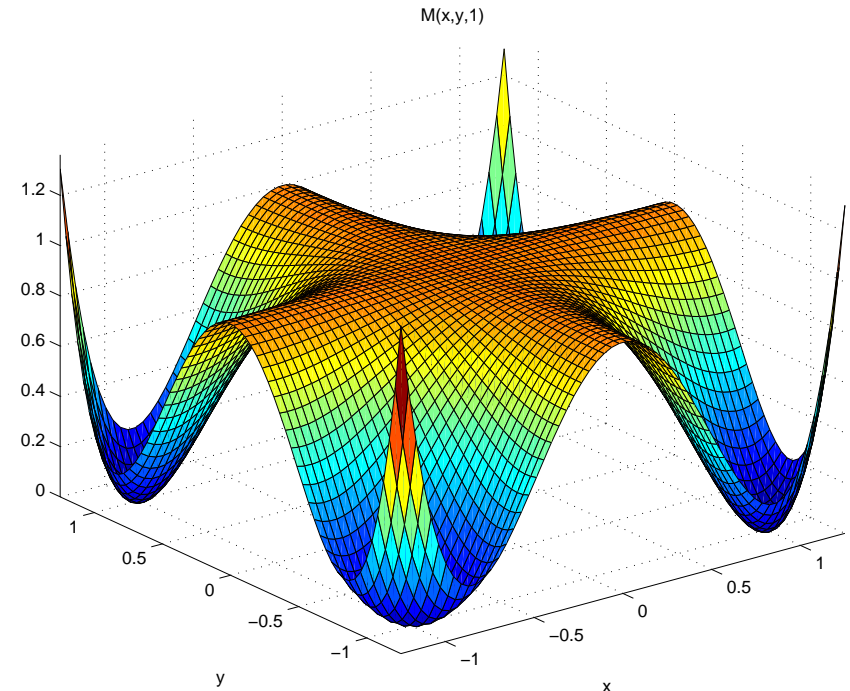
- Connections with Hilbert's 17th problem, solved by Artin: every PSD polynomial is a SOS of *rational functions*.
- If  $f$  is not SOS, then can try with  $gf$ , for some  $g$ .
  - For fixed  $f$ , can optimize over  $g$  too
  - Otherwise, can use a “universal” construction of Pólya-Reznick.

More about this later.

## The Motzkin Polynomial

A positive semidefinite polynomial, that is *not* a sum of squares.

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



- Nonnegativity follows from the arithmetic-geometric inequality applied to  $(x^2y^4, x^4y^2, 1)$
- Introduce a nonnegative factor  $x^2 + y^2 + 1$
- Solving the SDPs we obtain the decomposition:

$$\begin{aligned} (x^2 + y^2 + 1) M(x, y) &= (x^2y - y)^2 + (xy^2 - x)^2 + (x^2y^2 - 1)^2 + \\ &+ \frac{1}{4}(xy^3 - x^3y)^2 + \frac{3}{4}(xy^3 + x^3y - 2xy)^2 \end{aligned}$$

## The Univariate Case

$$\begin{aligned}
 f(x) &= a_0 + a_1x + a_2x^2 + a_3x^3 + \cdots + a_{2d}x^{2d} \\
 &= \begin{bmatrix} 1 \\ x \\ \vdots \\ x^d \end{bmatrix}^T \begin{bmatrix} q_{00} & q_{01} & \cdots & q_{0d} \\ q_{01} & q_{11} & \cdots & q_{1d} \\ \vdots & \vdots & \ddots & \vdots \\ q_{0d} & q_{1d} & \cdots & q_{dd} \end{bmatrix} \begin{bmatrix} 1 \\ x \\ \vdots \\ x^d \end{bmatrix} \\
 &= \sum_{i=0}^d \left( \sum_{j+k=i} q_{jk} \right) x^i
 \end{aligned}$$

- In the univariate case, the SOS condition is exactly equivalent to nonnegativity.
- The matrices  $A_i$  in the SDP have a Hankel structure. We will see how this can be exploited for efficient computation.

## Necessary Conditions

Suppose  $f = c_d x^d + c_{d-1} x^{d-1} + \dots + c_1 x + c_0$ ; then

$$f \text{ is PSD} \quad \implies \quad d \text{ is even, } c_d > 0 \text{ and } c_0 \geq 0$$

What is the analogue in  $n$  variables?

# The Newton Polytope

Suppose

$$f = \sum_{\alpha \in M} c_{\alpha} x^{\alpha}$$

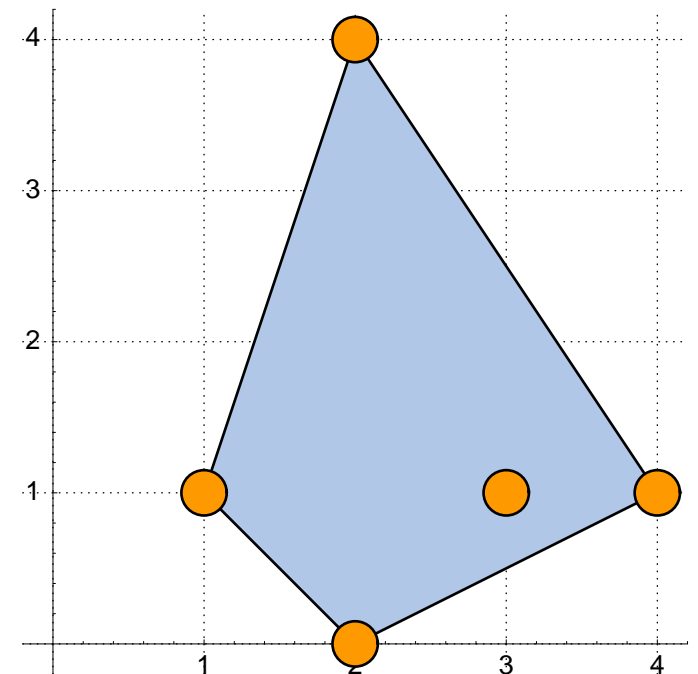
The set of monomials  $M \subset \mathbb{N}^n$  is called the *frame* of  $f$

The *Newton polytope* of  $f$  is its convex hull

$$\mathbf{new}(f) = \mathbf{co}(\mathbf{frame}(f))$$

The example shows

$$f = 7x^4y + x^3y + x^2y^4 + x^2 + 3xy$$

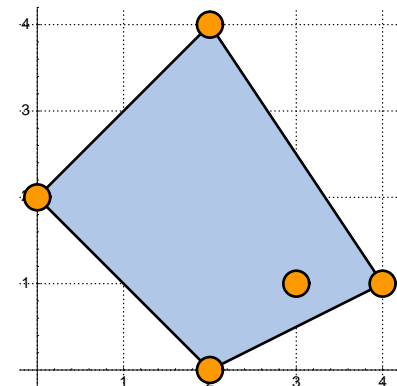
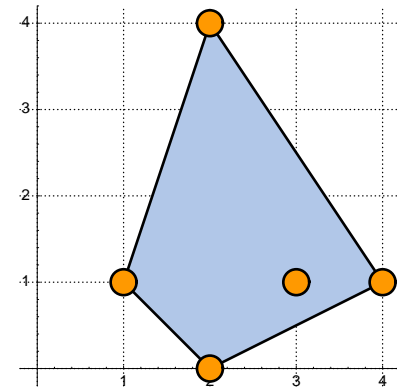


## Necessary Conditions for Nonnegativity

If  $f \in \mathbb{R}[x_1, \dots, x_n]$  is PSD, then

every vertex of  $\text{new}(f)$  has even coordinates, and a positive coefficient

- $f = 7x^4y + x^3y + x^2y^4 + x^2 + 3xy$   
 is not PSD, since term  $3xy$  has coords  $(1, 1)$
- $f = 7x^4y + x^3y - x^2y^4 + x^2 + 3y^2$   
 is not PSD, since term  $-x^2y^4$  has a negative coefficient



## Properties of Newton Polytopes

- Products:  $\mathbf{new}(fg) = \mathbf{new}(f) + \mathbf{new}(g)$
- Consequently  $\mathbf{new}(f^n) = n \mathbf{new}(f)$
- If  $f$  and  $g$  are PSD polynomials then

$$f(x) \leq g(x) \text{ for all } x \in \mathbb{R}^n \quad \Longrightarrow \quad \mathbf{new}(f) \subseteq \mathbf{new}(g)$$

- This tells us which monomials we have in an SOS decomposition

$$f = \sum_{i=1}^t g_i^2 \quad \Longrightarrow \quad \mathbf{new}(g_i) \subseteq \frac{1}{2} \mathbf{new}(f)$$

## Example of Sparse SOS Decomposition

Find an SOS representation for

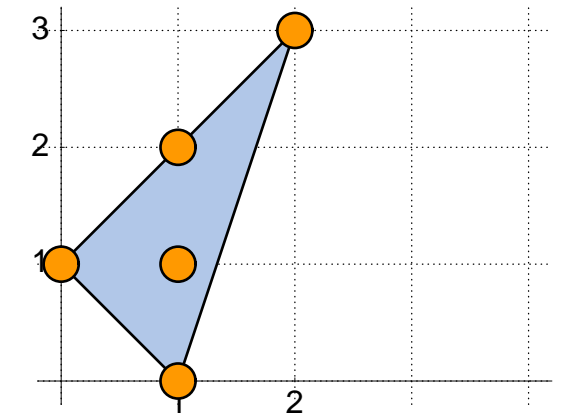
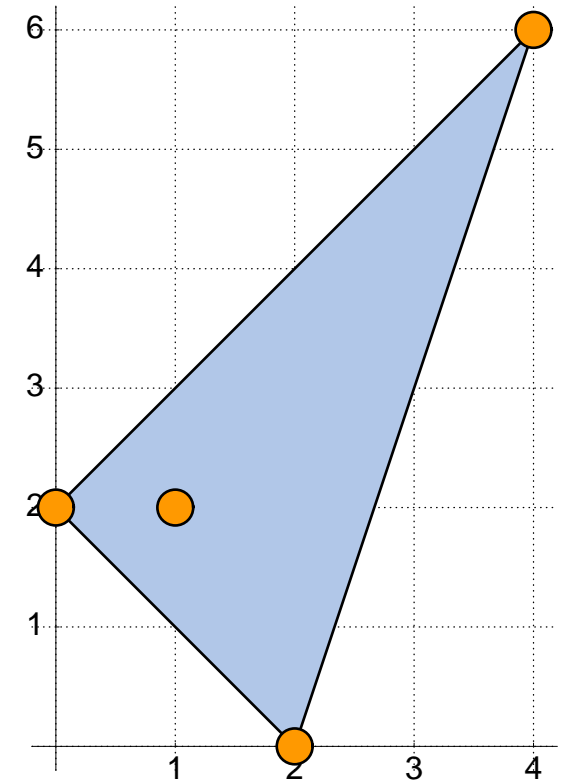
$$f = 4x^4y^6 + x^2 - xy^2 + y^2$$

The squares in an SOS decomposition can only contain the monomials

$$\text{new}\left(\frac{1}{2}f\right) \cap \mathbb{N}^n = \{x^2y^3, xy^2, xy, x, y\}$$

Without using sparsity, we would include all 21 monomials of degree  $< 5$  in the SDP

With sparsity, we only need 5 monomials



## About SOS/SDP

- The resulting SDP problem is polynomially sized (in  $n$ , for fixed  $d$ ).
- By properly choosing the monomials, we can exploit structure (sparsity, symmetries, ideal structure).
- An important feature: the problem is still a SDP *if the coefficients of  $F$  are variable*, and the dependence is affine.
- Can optimize over SOS polynomials in affinely described families.  
For instance, if we have  $p(x) = p_0(x) + \alpha p_1(x) + \beta p_2(x)$ , we can “easily” find values of  $\alpha, \beta$  for which  $p(x)$  is SOS.

This fact will be *crucial* in everything that follows...

# Global Optimization

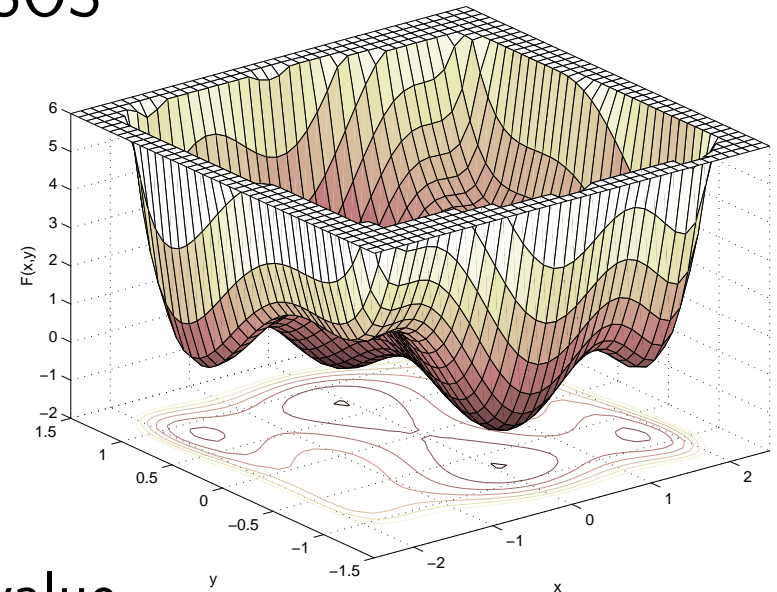
Consider the problem

$$\min_{x,y} f(x, y)$$

with

$$f(x, y) := 4x^2 - \frac{21}{10}x^4 + \frac{1}{3}x^6 + xy - 4y^2 + 4y^4$$

- Not convex. Many local minima. NP-hard.
- Find the largest  $\gamma$  s.t.  $f(x, y) - \gamma$  is SOS
- Essentially due to Shor (1987).
- A semidefinite program (convex!).
- If exact, can recover optimal solution.
- *Surprisingly* effective.



Solving, the maximum  $\gamma$  is -1.0316. Exact value.

## Why Does This Work?

Three *independent* facts, theoretical and experimental:

- The existence of efficient algorithms for SDP.
- The size of the SDPs grows much slower than the Bézout number  $\mu$ .
  - A bound on the number of (complex) critical points.
  - A reasonable estimate of complexity.
  - The bad news:  $\mu = (2d - 1)^n$  (for dense polynomials).
  - Almost all (exact) algebraic techniques scale as  $\mu$ .
- The lower bound  $f^{SOS}$  very often coincides with  $f^*$ .  
(Why? what does *often* mean?)

SOS provides *short proofs*, even though they're not guaranteed to exist.

## Coefficient Space

Let  $f_{\alpha\beta}(x) = x^4 + (\alpha + 3\beta)x^3 + 2\beta x^2 - \alpha x + 1$ .

What is the set of values of  $(\alpha, \beta) \in \mathbb{R}^2$  for which  $f_{\alpha\beta}$  is PSD? SOS?

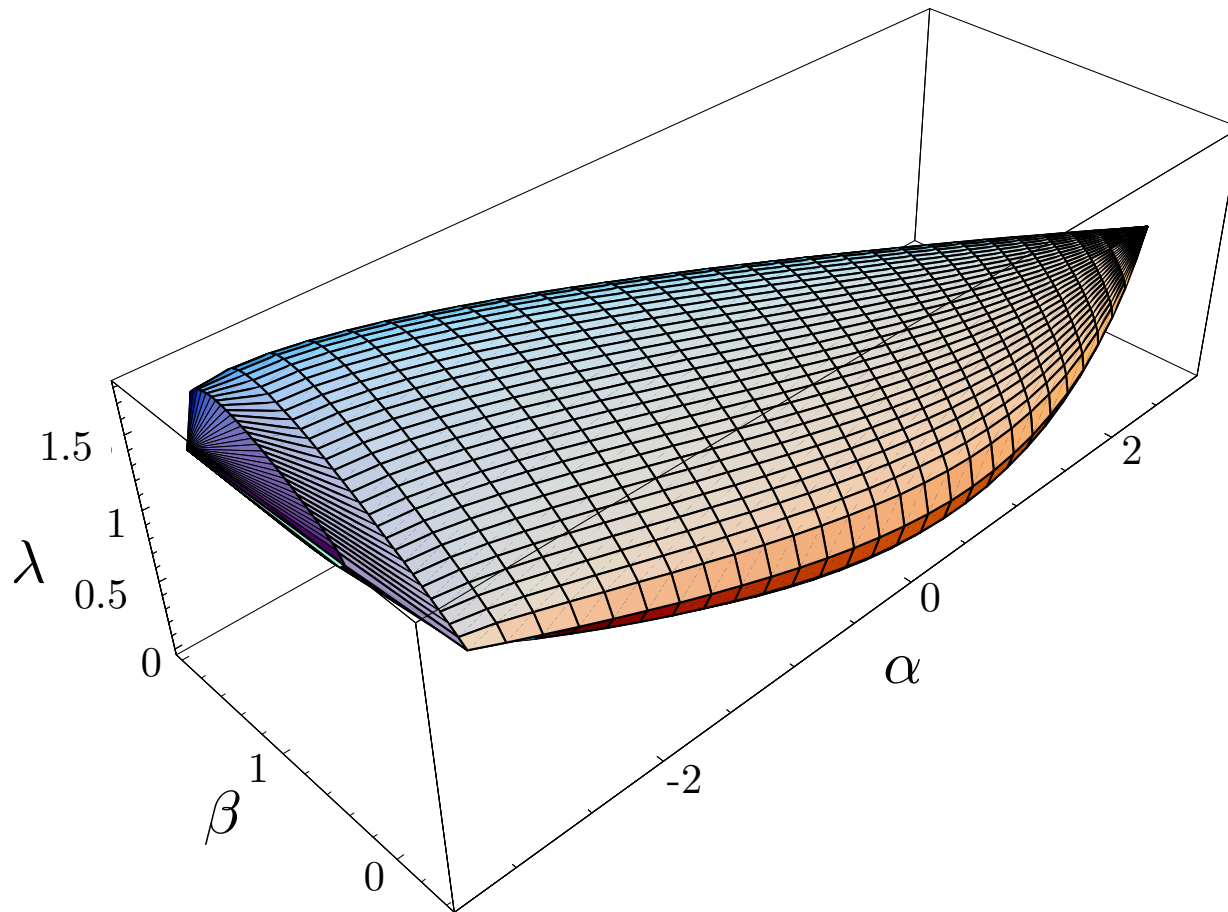
To find a SOS decomposition:

$$\begin{aligned} f_{\alpha,\beta}(x) &= 1 - \alpha x + 2\beta x^2 + (\alpha + 3\beta)x^3 + x^4 \\ &= \begin{bmatrix} 1 \\ x \\ x^2 \end{bmatrix}^T \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{12} & q_{22} & q_{23} \\ q_{13} & q_{23} & q_{33} \end{bmatrix} \begin{bmatrix} 1 \\ x \\ x^2 \end{bmatrix} \\ &= q_{11} + 2q_{12}x + (q_{22} + 2q_{13})x^2 + 2q_{23}x^3 + q_{33}x^4 \end{aligned}$$

The matrix  $Q$  should be PSD and satisfy the affine constraints.

The feasible set is given by:

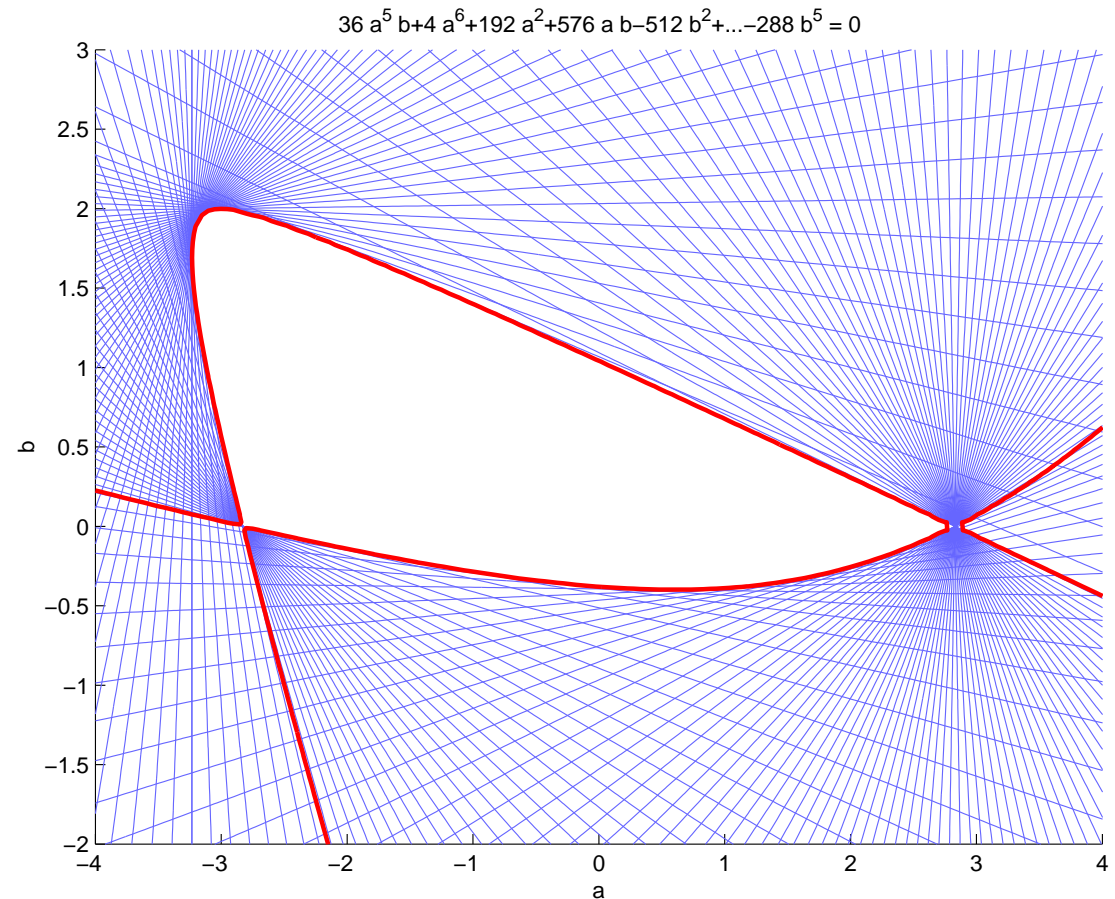
$$\left\{ (\alpha, \beta) \mid \exists \lambda \text{ s.t. } \begin{bmatrix} 1 & -\frac{1}{2}\alpha & \beta - \lambda \\ -\frac{1}{2}\alpha & 2\lambda & \frac{1}{2}(\alpha + 3\beta) \\ \beta - \lambda & \frac{1}{2}(\alpha + 3\beta) & 1 \end{bmatrix} \succeq 0 \right\}$$



What is the set of values of  $(\alpha, \beta) \in \mathbb{R}^2$  for which  $f_{\alpha\beta}$  PSD? SOS?

Recall: in the univariate case PSD=SOS, so here the sets are the same.

- Convex and semialgebraic.
- It is the projection of a spectrahedron in  $\mathbb{R}^3$ .
- We can easily test membership, or even optimize over it!

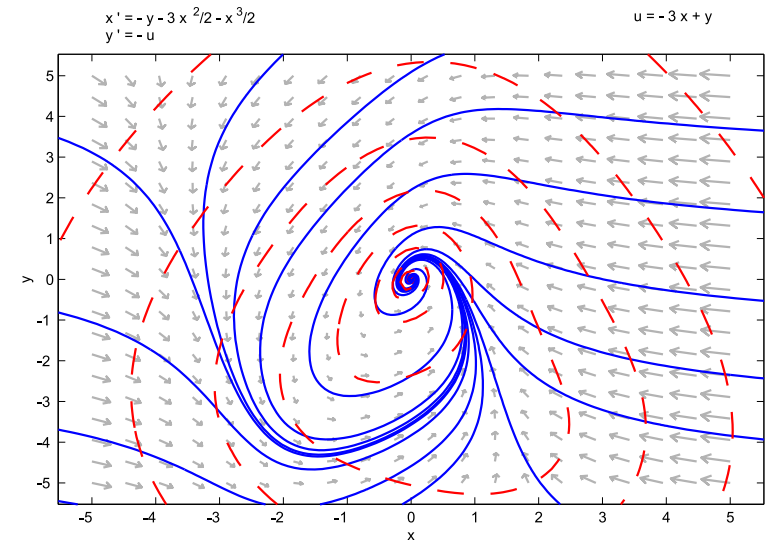


Defined by the curve:  $288\beta^5 - 36\alpha^2\beta^4 + 1164\alpha\beta^4 + 1931\beta^4 - 132\alpha^3\beta^3 + 1036\alpha^2\beta^3 + 1956\alpha\beta^3 - 2592\beta^3 - 112\alpha^4\beta^2 + 432\alpha^3\beta^2 + 1192\alpha^2\beta^2 - 1728\alpha\beta^2 + 512\beta^2 - 36\alpha^5\beta + 72\alpha^4\beta + 360\alpha^3\beta - 576\alpha^2\beta - 576\alpha\beta - 4\alpha^6 + 60\alpha^4 - 192\alpha^2 - 256 = 0$

# Lyapunov Stability Analysis

To prove asymptotic stability of  $\dot{x} = f(x)$ ,

$$\begin{aligned} V(x) &> 0 \quad x \neq 0 \\ \dot{V}(x) &= \left(\frac{\partial V}{\partial x}\right)^T f(x) < 0, \quad x \neq 0 \end{aligned}$$



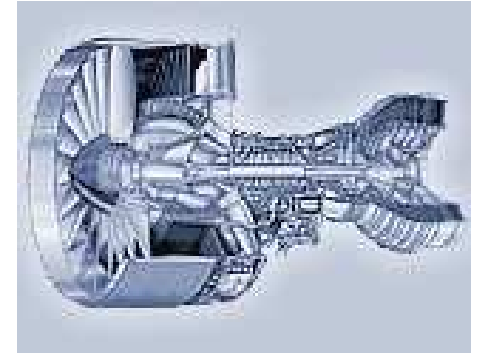
- For linear systems  $\dot{x} = Ax$ , quadratic Lyapunov functions  $V(x) = x^T P x$ 

$$P > 0, \quad A^T P + P A < 0.$$
- With an affine family of candidate polynomial  $V$ ,  $\dot{V}$  is also affine.
- Instead of *checking nonnegativity*, use a *SOS condition*.
- Therefore, for polynomial vector fields and Lyapunov functions, we can check the conditions using the theory described before.

## Lyapunov Example

A jet engine model (derived from Moore-Greitzer), with controller:

$$\begin{aligned}\dot{x} &= -y - \frac{3}{2}x^2 - \frac{1}{2}x^3 \\ \dot{y} &= 3x - y\end{aligned}$$



Try a generic 4th order polynomial Lyapunov function.

$$V(x, y) = \sum_{0 \leq j+k \leq 4} c_{jk} x^j y^k$$

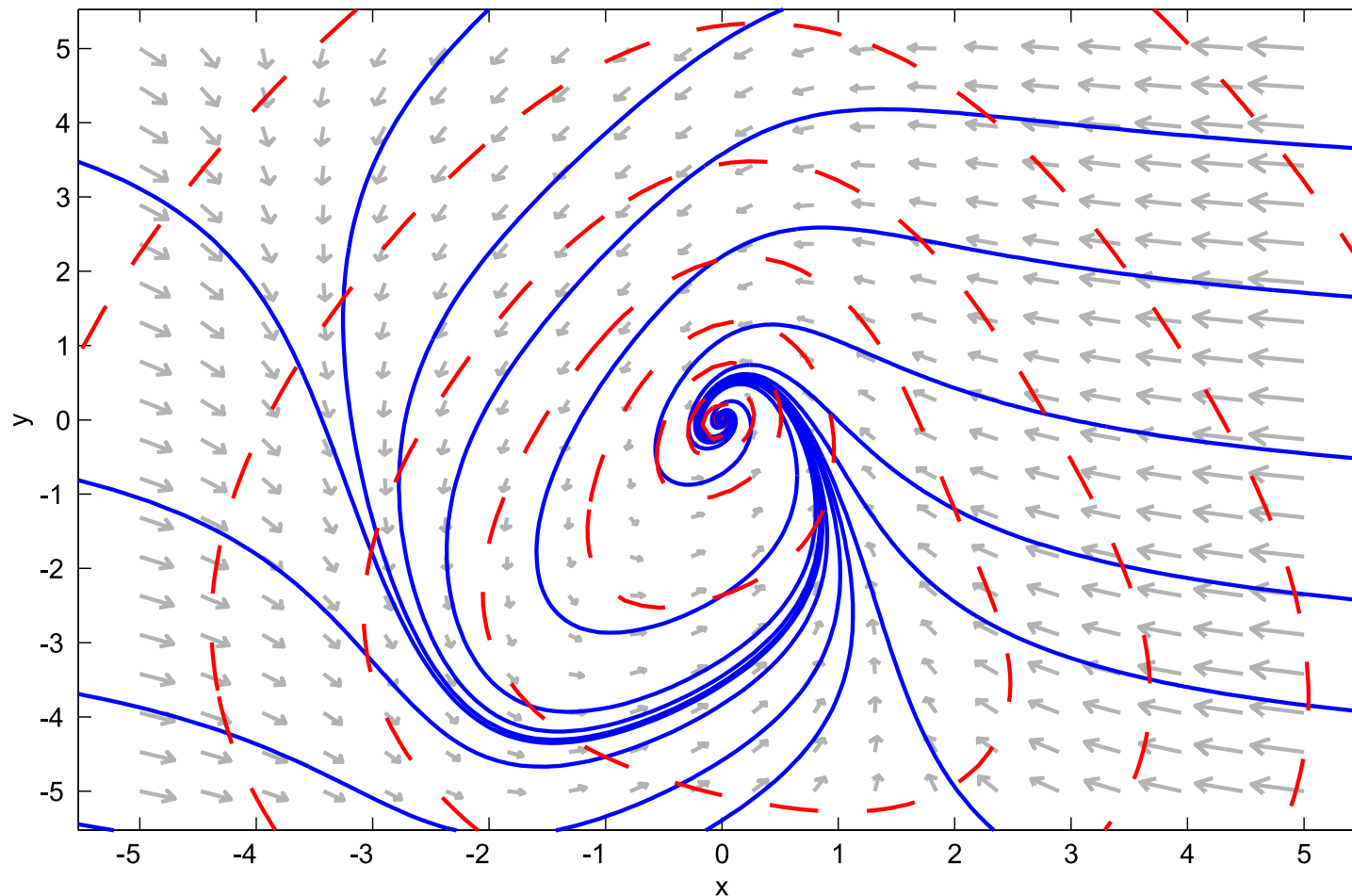
Find a  $V(x, y)$  that satisfies the conditions:

- $V(x, y)$  is SOS.
- $-\dot{V}(x, y)$  is SOS.

Both conditions are affine in the  $c_{jk}$ . Can do this directly using SOS/SDP!

# Lyapunov Example

After solving the SDPs, we obtain a Lyapunov function.



$$V = 4.5819x^2 - 1.5786xy + 1.7834y^2 - 0.12739x^3 + 2.5189x^2y - 0.34069xy^2 + 0.61188y^3 + 0.47537x^4 - 0.052424x^3y + 0.44289x^2y^2 + 0.0000018868xy^3 + 0.090723y^4$$

## Lyapunov Example

(M. Krstić) Find a Lyapunov function for

$$\dot{x} = -x + (1 + x)y$$

$$\dot{y} = -(1 + x)x.$$

we easily find a quartic polynomial

$$V(x, y) = 6x^2 - 2xy + 8y^2 - 2y^3 + 3x^4 + 6x^2y^2 + 3y^4.$$

Both  $V(x, y)$  and  $(-\dot{V}(x, y))$  are SOS:

$$V(x, y) = \begin{bmatrix} x \\ y \\ x^2 \\ xy \\ y^2 \end{bmatrix}^T \begin{bmatrix} 6 & -1 & 0 & 0 & 0 \\ -1 & 8 & 0 & 0 & -1 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 6 & 0 \\ 0 & -1 & 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \\ x^2 \\ xy \\ y^2 \end{bmatrix}, \quad -\dot{V}(x, y) = \begin{bmatrix} x \\ y \\ x^2 \\ xy \end{bmatrix}^T \begin{bmatrix} 10 & 1 & -1 & 1 \\ 1 & 2 & 1 & -2 \\ -1 & 1 & 12 & 0 \\ 1 & -2 & 0 & 6 \end{bmatrix} \begin{bmatrix} x \\ y \\ x^2 \\ xy \end{bmatrix}$$

The matrices are positive definite, so this proves asymptotic stability.

## Extensions

- Other linear differential inequalities (e.g. Hamilton-Jacobi).
- Many possible variations: nonlinear  $\mathcal{H}_\infty$  analysis, parameter dependent Lyapunov functions, etc.
- Can also do local results (for instance, on compact domains).
- Polynomial and rational vector fields, or functions with an underlying algebraic structure.
- Natural extension of the LMIs for the linear case.
- Only for *analysis*. Proper *synthesis* is trickier...

# Nonlinear Control Synthesis

Recently, Rantzer provided an alternative stability criterion, in some sense “dual” to the standard Lyapunov one.

$$\nabla \cdot (\rho f) > 0$$

- The *synthesis* problem is now *convex* in  $(\rho, u\rho)$ .

$$\nabla \cdot [\rho(f + gu)] > 0$$

- Parametrizing  $(\rho, u\rho)$ , can apply SOS methods.

Example:

$$\dot{x} = y - x^3 + x^2$$

$$\dot{y} = u$$

A stabilizing controller is:

$$u(x, y) = -1.22x - 0.57y - 0.129y^3$$

