Course on LMI optimization with applications in control

Didier Henrion

Maximal positively invariant set approximation

Prague 1 Dec 2022

Recall the main steps of the moment-SOS aka Lasserre hierarchy.

Given a **nonlinear nonconvex** problem:

1. Reformulate it as a **linear** problem (at the price of enlarging or changing the space of solutions);

2. Solve approximately the linear problem with a hierarchy of tractable **convex** relaxations (of increasing size);

3. Ensure **convergence**: either the original problem is solved at a finite relaxation size, or its solution is approximated with increasing quality.

At each step, **conic duality** is an essential ingredient.

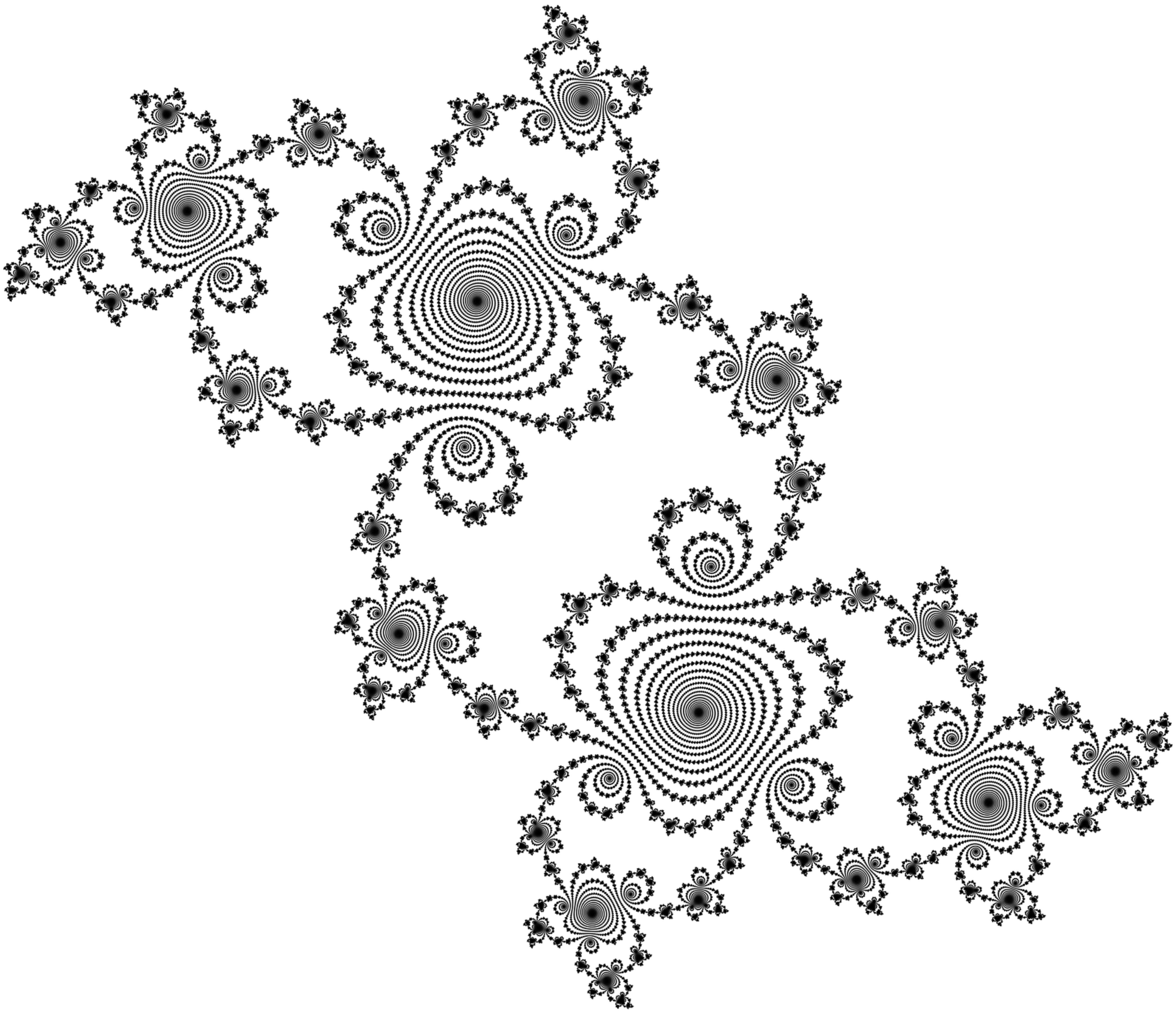
Maximal positively invariant set

We will follow this programme for the problem of estimating the maximal positively invariant (**MPI**) set for a discrete time dynamical system

Given $f : \mathbb{R}^n \to \mathbb{R}^n$ polynomial and a bounded basic semialgebraic set $X := \{x \in \mathbb{R}^n : p_k(x) \geq 0, k = 1, \ldots, m\}$, the MPI is

$$X_I := \{x_0 \in X : x_{t+1} = f(x_t) \in X, \forall t = 0, 1, \ldots\}$$

Even in the simplest cases (e.g. $n = 2$ and f quadratic) the MPI can be very complicated



Approximations

We content ourselves with approximations of X_I, whose quality improves at the price of more computation

During this course we will describe the following solution:

The Lasserre hierarchy can generate **outer approximations**

$$X_{I_r} := \{x \in X : v_r(x) \geq 0\} \supset X_I$$

with polynomials $v_r \in \mathbb{R}[x]_r$ **converging in volume**

$$\text{vol}\, X_{I_\infty} = \text{vol}\, X_I$$

Step 1 - Linear reformulation

For POP the key idea was to formulate an LP on probability measures: an optimal solution to the LP was then the Dirac measure at an optimal solution of the POP

Now for dynamical systems we proceed similarly

Given a trajectory $t \mapsto x_t$, define

$$\mu_t(dx) := \delta_{x_t}(dx)$$

or equivalently for all $A \in \mathscr{B}(X)$

$$\mu_t(A) := I_A(x_t) = \begin{cases} 1 & \text{if } x_t \in A \\ 0 & \text{otherwise} \end{cases}$$

Given an initial condition x_0 and a discount factor $\alpha \in (0, 1)$, define the discounted **occupation measure**

$$\mu(dx|x_0) := \sum_{t=0}^{\infty} \alpha^t \mu_t(dx|x_0)$$

and observe that its mass is finite

$$\mu(X|x_0) = \sum_{t=0}^{\infty} \alpha^t = \frac{1}{1-\alpha}$$

Now suppose that the initial condition x_0 in X is not a single point but a distribution of mass i.e. a probability measure μ_0 in X, and define the average discounted occupation measure

$$\mu(dx) := \int_X \mu(dx|x_0)\mu_0(dx_0)$$

Now we derive an equation linking μ and μ_0

Consider a trajectory $t \mapsto x_t$ staying in X, and observe that for any observable $v \in C(X)$ it holds

$$\begin{aligned}
\int_X v(x)\mu(dx|x_0) &= \sum_{t=0}^{\infty} \alpha^t v(x_t) = v(x_0) + \alpha \sum_{t=0}^{\infty} \alpha^t v(x_{t+1}) \\
&= v(x_0) + \alpha \sum_{t=0}^{\infty} \alpha^t v(f(x_t)) \\
&= v(x_0) + \alpha \int_X v(f(x)) d\mu(dx|x_0)
\end{aligned}$$

Integrating with respect to the initial distribution yields

$$\int_X v(x) d\mu(dx) = \int_X v(x) d\mu_0(dx) + \alpha \int_X v(f(x)) d\mu(dx)$$

We obtain a **linear equation**

$$\int_X v(x) d\mu(dx) = \int_X v(x) d\mu_0(dx) + \alpha \int_X v(f(x)) d\mu(dx)$$

that we can write by duality

$$\begin{aligned} \langle v, \mu \rangle &= \langle v, \mu_0 \rangle + \alpha \langle v \circ f, \mu \rangle \\ &= \langle v, \mu_0 \rangle + \alpha \langle v, f_{\#} \mu \rangle \end{aligned}$$

as a linear equation on measures

$$\mu = \mu_0 + \alpha f_{\#} \mu$$

called the **Liouville equation**

The **Koopman** or composition operator

$$v \mapsto v \circ f$$

is adjoint to the **Frobenius-Perron** or push-forward operator

$$\mu \mapsto f_{\#}\mu$$

i.e. $\langle v \circ f, \mu \rangle = \langle v, f_{\#}\mu \rangle$ for all $v \in C(X)$, $\mu \in C(X)'$

Information on the behavior of dynamical system $x_{t+1} = f(x_t)$ can be inferred from a spectral analysis of these linear operators

The push-forward or **image measure** is

$$f_{\#}\mu(A) := \mu(\{x \in X : f(x) \in A\})$$

for all $A \in \mathscr{B}(X)$

Measures μ satisfying $f_{\#}\mu = \mu$ are called **invariant**

The Krylov-Bogolyubov Theorem asserts that if f is continuous and X is compact there is always an invariant measure

Any invariant measure μ solves the Liouville equation

$$\mu = \mu_0 + \alpha f_{\#}\mu$$

for the choice $\mu_0 = (1 - \alpha)\mu$

For example, consider the logistic map

$$f(x) = 4x(1-x)$$

on

$$X := [0,1]$$

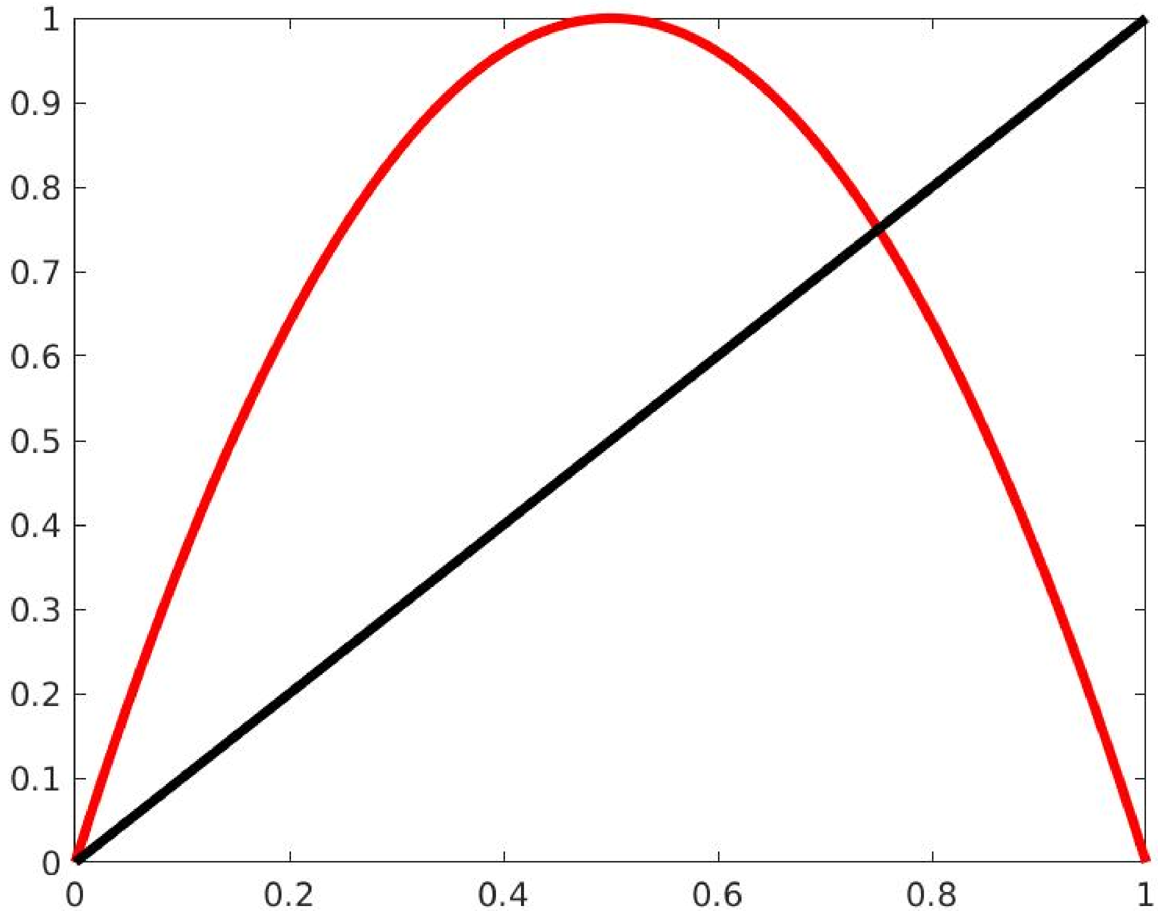
Exercise 2.1:

a. Given $\mu(dx) = m(x)dx$, derive analytically $f_{\#}\mu$.

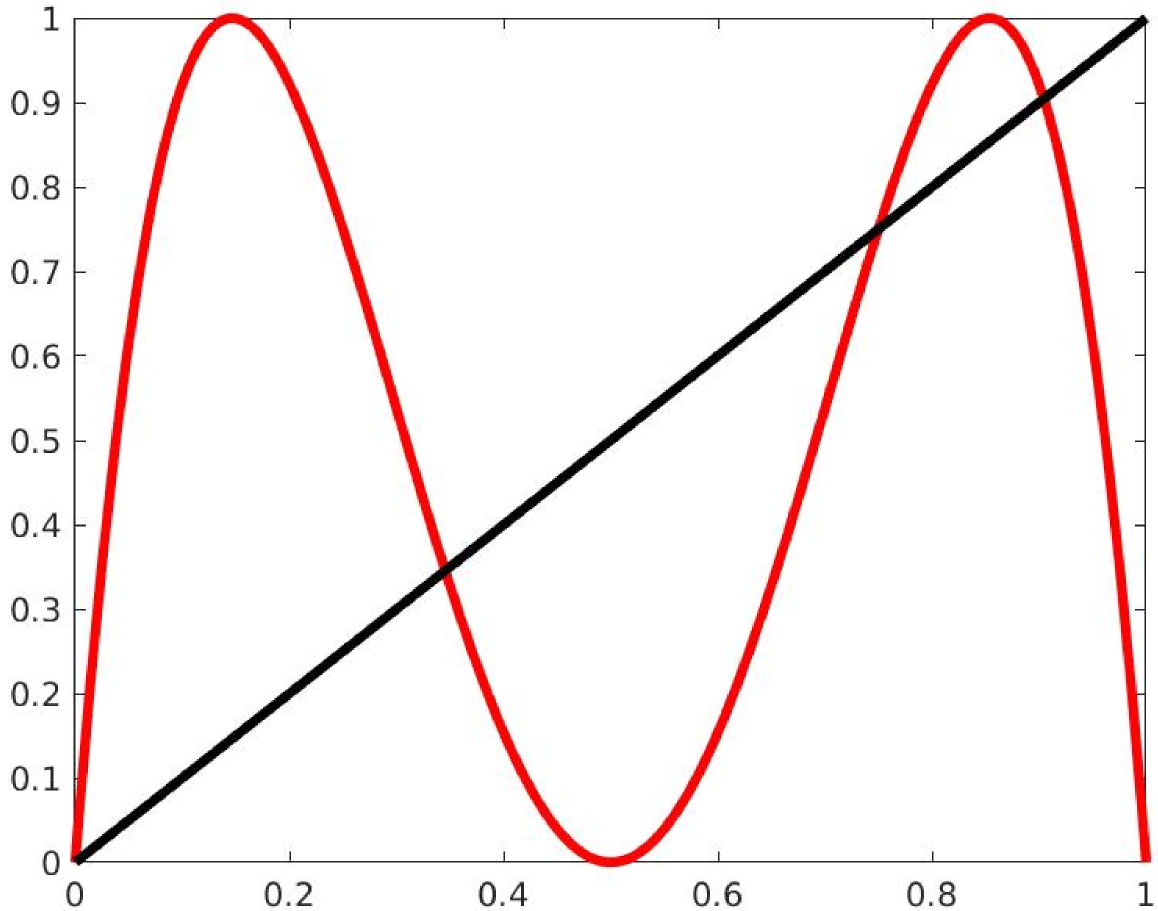
b. Given $\mu(dx) = I_{[0,1]}(x)dx$ compute $f_{\#}\mu$ and $f \circ f_{\#}\mu$.

c. Prove that $\mu(dx) = dx/(\pi\sqrt{x(1-x)})$ is invariant.

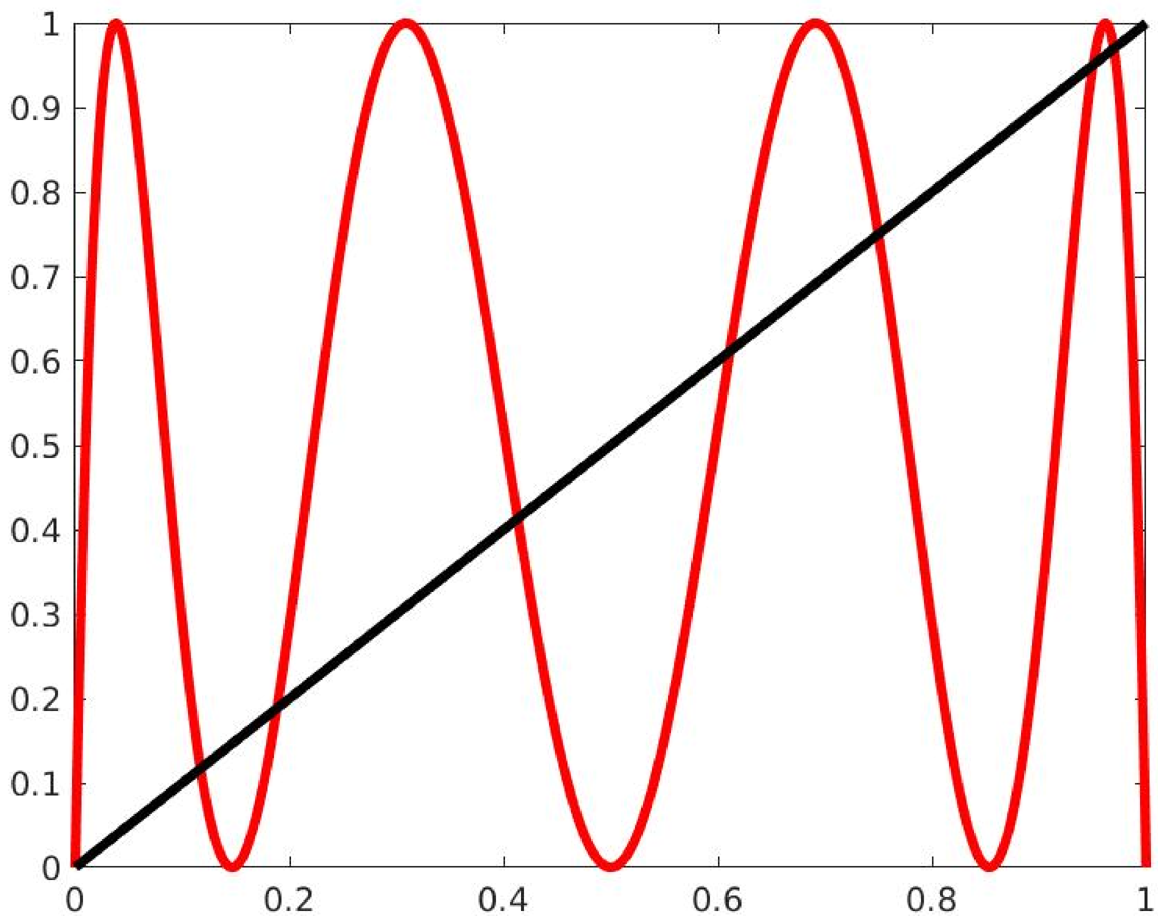
d. Prove that $\mu(dx) = \delta_{3/4}(dx)$ is invariant.



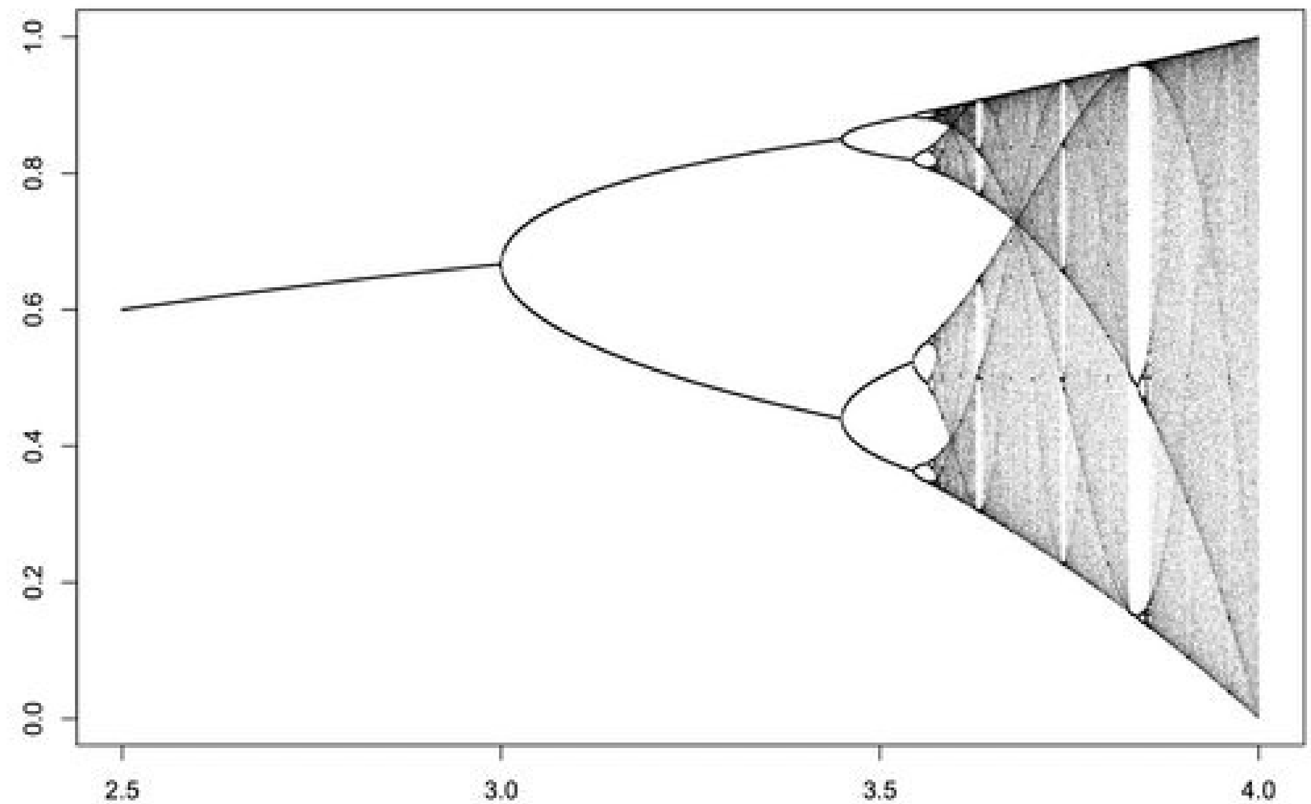
$$f(x) = x$$



$$f(f(x)) = x$$



$$f(f(f(x))) = x$$



$$f(x) = kx(1-x)$$

Instead of the **nonlinear** dynamical system

$$x_{t+1} = f(x_t)$$

defined on X we have now a **linear** Liouville equatiobn

$$\mu = \mu_0 + \alpha f_{\#}\mu$$

defined on occupation measures on X

For the MPI set $X_I := \{x_0 \in X : x_{t+1} = f(x_t) \in X, \forall t = 0, 1, \ldots\}$ we have the following result

Lemma: For any μ and μ_0 satisying the Liouville equation and $\operatorname{spt}\mu \subset X$ and $\operatorname{spt}\mu_0 \subset X$ it holds

$$\operatorname{spt}\mu_0 \subset X_I$$

where the support of a measure can be defined as

$$\operatorname{spt}\mu_0 := \{x_0 \in X : \mu_0(\{x : |x - x_0| \leq \varepsilon\}) > 0, \forall \varepsilon > 0\}$$

Since the support of the initial measure is contained in the MPI set we seek an initial measure with largest possible support

To achieve this, consider the LP

$$\begin{array}{rcl} p^* & = & \sup \ \langle 1, \mu_0 \rangle \\ & & \text{s.t.} \ \mu = \mu_0 + \alpha f_{\#}\mu \\ & & \qquad \mu_0 + \widehat{\mu}_0 = \lambda_X \end{array}$$

where λ_X is the Lebesgue measure on X and the optimization variables are μ, μ_0, $\widehat{\mu}_0$ all in $C(X)'_+$

Theorem: The supremum is attained by $\mu_0^* = \lambda_{X_I}$ and hence $p^* = \text{vol}\, X_I$

Exercise 2.2: Provide a graphical proof to the theorem.

The dual LP reads

$$\begin{aligned} d^* &= \inf \ \langle w, \lambda_X \rangle \\ & \text{s.t.} \ (v - \alpha v \circ f, \ w - v - 1, \ w) \in C(X)^3_+ \end{aligned}$$

or equivalently

$$\begin{aligned} d^* = \inf \ & \int_X w(x)dx \\ \text{s.t.} \ & \alpha v(f(x)) \leq v(x), \\ & w(x) \geq v(x) + 1, \\ & w(x) \geq 0, \quad \forall x \in X \end{aligned}$$

Exercise 2.3: Derive the dual using convex duality. Prove that there is no duality gap.

For the LP

$$\begin{aligned} d^* \;=\; \inf \;& \int_X w(x)dx \\ \text{s.t. } & \alpha v(f(x)) \le v(x), \\ & w(x) \ge v(x) + 1, \\ & w(x) \ge 0, \quad \forall x \in X \end{aligned}$$

any dual feasible pair (v, w) satisfies $v \ge 0$ and $w \ge 1$ on X_I

To prove this, consider a trajectory $(x_t)_{t=0,1,\dots} \subset X$ and note that the 1st inequality implies $v(x_0) \ge \alpha v(x_1) \ge \alpha^2 v(x_2) \ge \alpha^t v(x_t) \to 0$ as $t \to \infty$ since $\alpha \in (0,1)$, $x_t \in X$ and X is bounded

Therefore $v(x_0) \ge 0$ and from the 2nd inequality $w(x_0) \ge 1$

Step 2 - Convex hierarchy

To solve the primal LP

$$\begin{array}{rcll} p^* & = & \sup & \langle 1, \mu_0 \rangle \\ & & \text{s.t.} & \mu = \mu_0 + \alpha f_{\#}\mu \\ & & & \mu_0 + \hat{\mu}_0 = \lambda_X \\ & & & (\mu, \mu_0, \hat{\mu}_0) \in C(X)'^3_+ \end{array}$$

and dual LP

$$\begin{array}{rcll} d^* & = & \inf & \langle w, \lambda_X \rangle \\ & & \text{s.t.} & (v - \alpha v \circ f,\ w - v - 1,\ w) \in C(X)^3_+ \end{array}$$

with X bounded basic semialgebraic and f polynomial
we can readily use the moment-SOS hierarchy

We replace $C(X)_+$ with $Q(X)_r$ for increasing relaxation order r
and we get sequences p^*_r and d^*_r as well as pseudo-moments and
polynomials v_r, w_r in $\mathbb{R}[x]_r$

Step 3 - Convergence

Recall that for the primal LP

$$\begin{array}{rcll} p^* & = & \text{sup} & \langle 1, \mu_0 \rangle \\ & & \text{s.t.} & \mu = \mu_0 + \alpha f_{\#}\mu \\ & & & \mu_0 + \widehat{\mu}_0 = \lambda_X \\ & & & (\mu, \mu_0, \widehat{\mu}_0) \in C(X)'^3_+ \end{array}$$

the optimal value is $p^* = \text{vol}\, X_I$ **attained** by $\mu_0^* = \lambda_{X_I}$

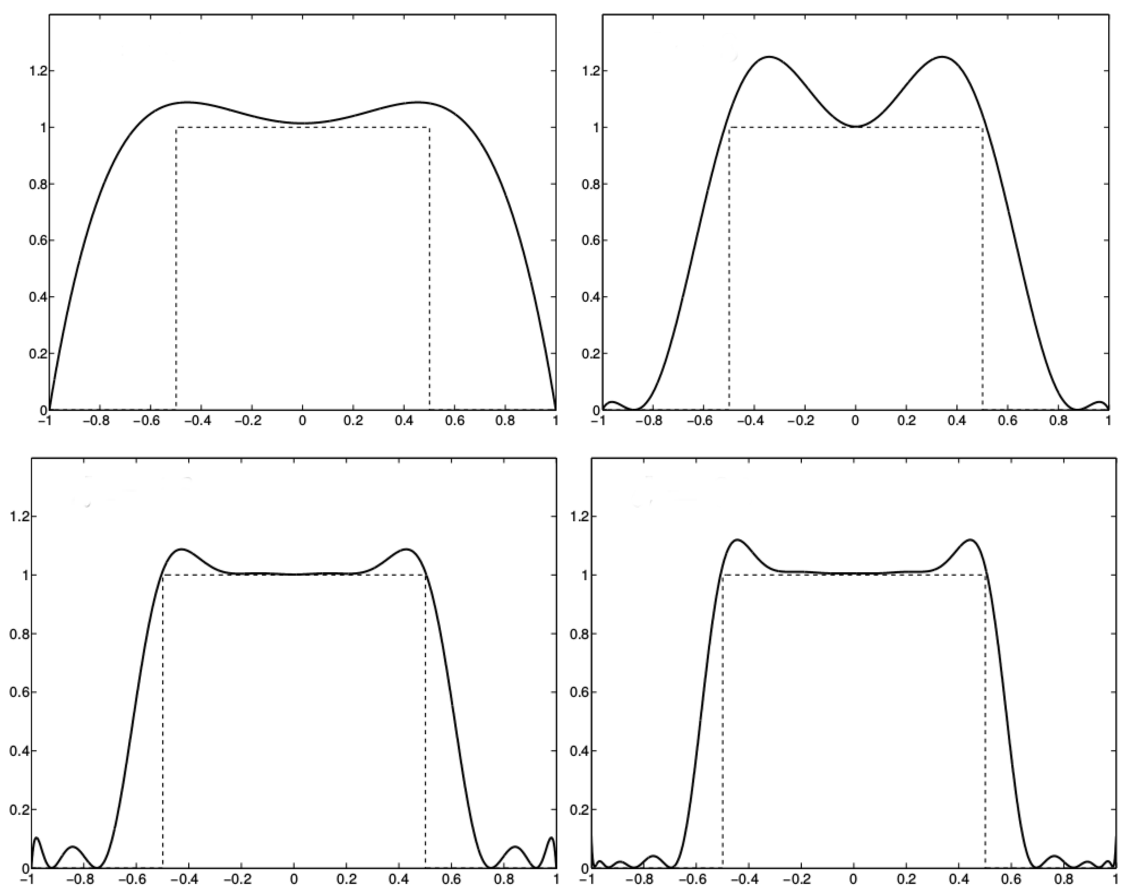
For the dual LP

$$\begin{array}{rcll} d^* & = & \text{inf} & \langle w, \lambda_X \rangle \\ & & \text{s.t.} & (v - \alpha v \circ f,\ w - v - 1,\ w) \in C(X)^3_+ \end{array}$$

since $w \geq I_{X_I}$ the objective function is $\int_X w(x)dx = \|w\|_{\mathscr{L}_1(X)}$

At optimality by strong duality it is equal to $\int_X I_{X_I} dx = \text{vol}\, X_I$ and hence it is **not attained** in $C(X)$

Here is an example of a minimizing sequence w_r for $X_I = [-\frac{1}{2}, \frac{1}{2}]$



Theorem: By replacing $C(X)_+$ with $Q(X)_r$ we get a monotone converging sequence of upper bounds

$$p_r^* = d_r^* \geq p_{r+1}^* = d_{r+1}^* \geq p_\infty^* = d_\infty^* = \text{vol } X_I$$

Exercise 2.4: Prove it with the Stone-Weierstrass Theorem.

Theorem: In the dual we obtain a sequence of polynomials v_r, w_r in $\mathbb{R}[x]_r$ such that

$$X_{I_r} := \{x \in X : v_r(x) \geq 0\} \supset X_I$$

and

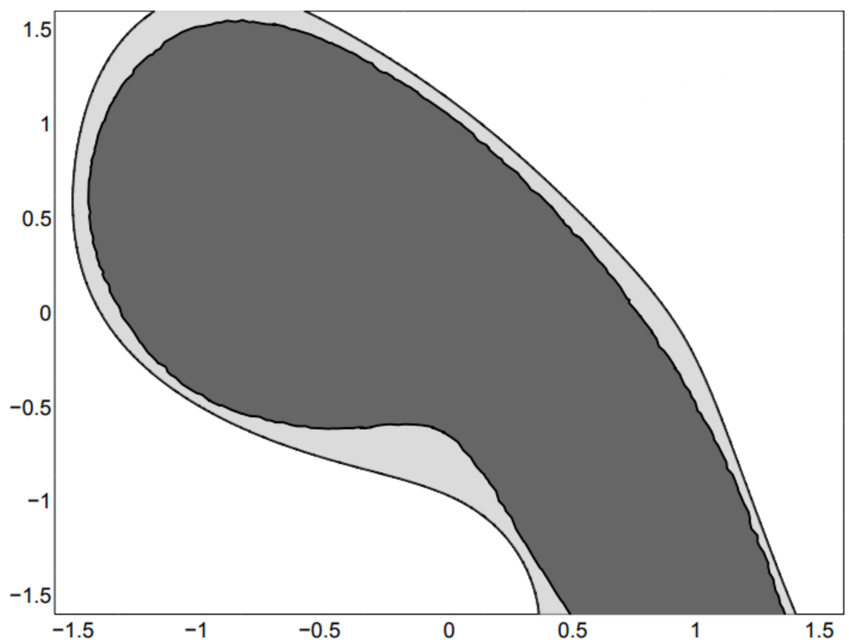
$$\lim_{r \to \infty} \text{vol}(X_{I_r} \setminus X_I) = 0$$

Exercise 2.5: Prove it by showing that $w_r \to I_{X_I}$ in $\mathscr{L}_1(X)$.

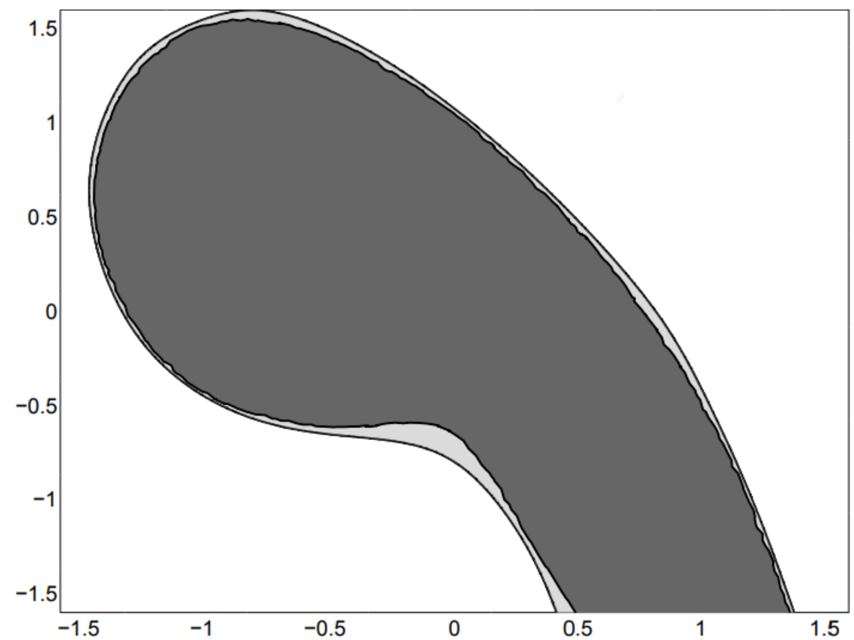
Examples

Cathala system

$$f(x) = \left(x_1 + x_2,\ -0.5952 + x_2 + x_1^2\right)$$



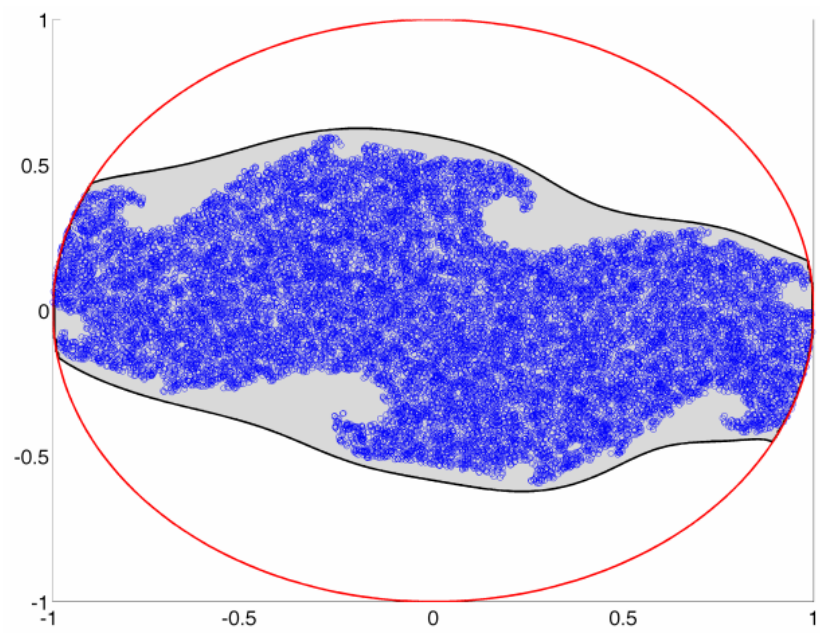
$r = 6$



$r = 10$

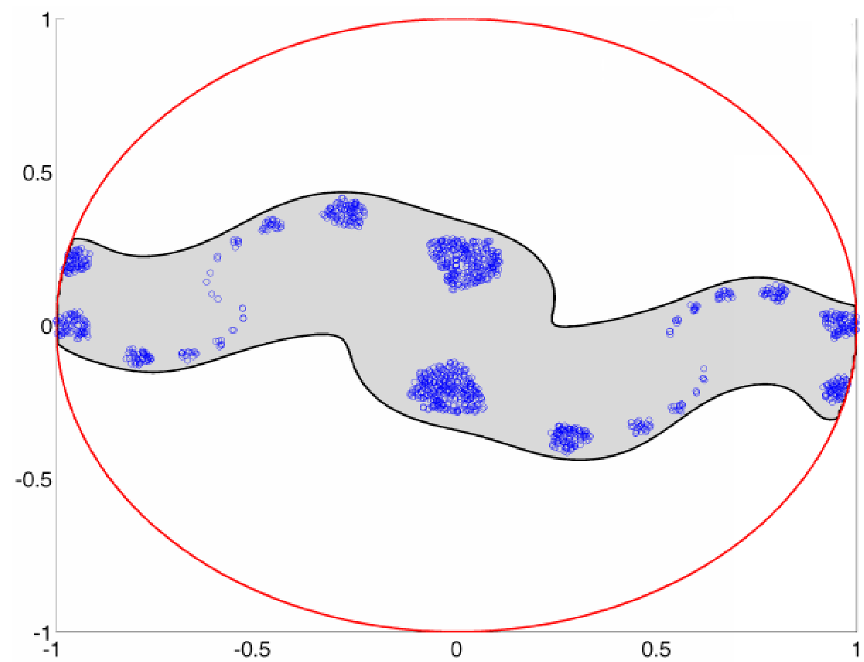
Julia sets

$$z_{t+1} = z_t^2 + c, \quad z_t \in \mathbb{C}$$



$c = -0.7 + i0.2$

$r = 12$



$c = -0.9 + i0.2$

$r = 12$