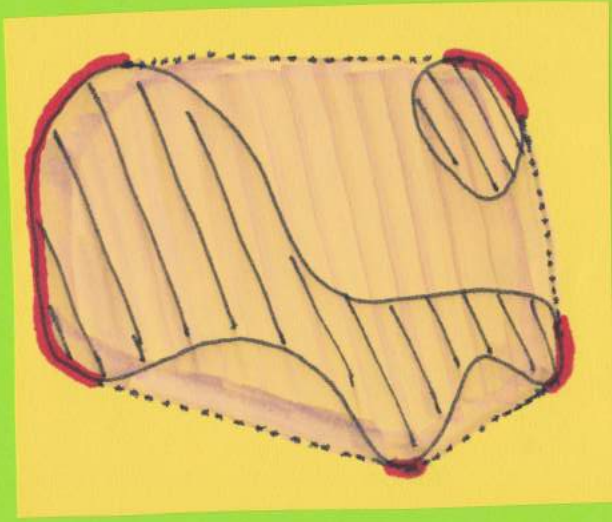
On the exactness of Lasserre relaxations of SPIs and POPs

June 22, 2019



Markus Schweighofer
Universität Konstanz

(joint work with Tom-Lukas Kriel)
now: TNG Technology Consulting, Munich

MEGAR 2019, Madrid, June 21-22, 2019

Métodos Efectivos en Geometría Algebraica

Reformulation and linearization technique

(RLT)

$$1 + 2x_1 + \ldots + x_1^3 x_2 - x_1^4 x_2 \geq 0$$

polynomial inequality of some degree

Reformulation and linearization technique

(RLT)

$$1+2x_1+\ldots+x_1^3x_2-x_1^4x_2\geq 0$$

polynomial inequality of some degree

$$\Longleftrightarrow \forall a_1,\ldots,a_6\in\mathbb{R}:(a_1+a_2x_1+\ldots+a_6x_2^2)^2(1+2x_1+\ldots+x_1^3x_2-x_1^4x_2)\geq 0$$

infinite system of polynamial inequalities of higher degree

Reformulation and linearization technique

(RLT)

$$1+2x_1+\dots+x_1^3x_2-x_1^4x_2\geq 0$$

polynomial inequality of some degree

$$\Longleftrightarrow \forall a_1,\dots,a_6\in\mathbb{R}: (a_1+a_2x_1+\dots+a_6x_2^2)^2(1+2x_1+\dots+x_1^3x_2-x_1^4x_2)\geq 0$$

infinite system of polynomial inequalities of higher degree

$$\Longleftrightarrow \forall a_1,\dots,a_6\in\mathbb{R}: \underset{\substack{\uparrow\\\text{row vector}}}{(a_1 \dots a_6)} \underbrace{\underset{\substack{\uparrow\\\text{scalar}}}{(1+2x_1+\dots+x_1^3x_2-x_1^4x_2)} \underbrace{\underset{\substack{\uparrow\\\text{column}\\\text{vector}}}{\begin{pmatrix}1\\x_1\\ \vdots\\ x_2^2\end{pmatrix}} \underset{\substack{\uparrow\\\text{row vector}}}{(1\ x_1 \dots x_2^2)}}_{\text{symmetric matrix}}}_{\text{symmetric matrix}} \underset{\substack{\uparrow\\\text{column}\\\text{vector}}}{\begin{pmatrix}a_1\\ \vdots\\ a_6\end{pmatrix}} \geq 0$$

Reformulation and linearization technique

(RLT)

$$1+2x_1+\ldots+x_1^3x_2-x_1^4x_2\geq 0$$

polynomial inequality of some degree

$$\iff \forall a_1,\ldots,a_6\in\mathbb{R}: (a_1+a_2x_1+\ldots+a_6x_2^2)^2(1+2x_1+\ldots+x_1^3x_2-x_1^4x_2)\geq 0$$

infinite system of polynanial inequalities of higher degree

$$\iff \forall a_1,\ldots,a_6\in\mathbb{R}: \underset{\substack{\uparrow\\\text{row vector}}}{(a_1\ldots a_6)}\underbrace{\underset{\substack{\uparrow\\\text{scalar}}}{(1+2x_1+\ldots+x_1^3x_2-x_1^4x_2)}\underbrace{\underset{\substack{\uparrow\\\text{column}\\\text{vector}}}{\begin{pmatrix}1\\x_1\\\vdots\\x_2^2\end{pmatrix}}\underset{\substack{\uparrow\\\text{row vector}}}{(1\ x_1\ldots x_2^2)}}_{\text{symmetric matrix}}}_{\text{symmetric matrix}}\underset{\substack{\uparrow\\\text{column}\\\text{vector}}}{\begin{pmatrix}a_1\\\vdots\\a_6\end{pmatrix}}\geq 0$$

$$\iff \begin{pmatrix}1+2x_1+\ldots-x_1^4x_2 & & \cdot\\ & \cdot & \\ \cdot & & x_2^4+\ldots\end{pmatrix}\underset{\substack{\uparrow\\\text{psd}}}{\succeq} 0$$

polynomial matrix Inequality of the higher degree

Reformulation and linearization technique

(RLT)

$1+2x_1+\dots+$ [redacted] y_{16} $-$ [redacted] y_{17} ≥ 0

[redacted] of
[redacted]

$\iff \forall a_1,\dots,a_6\in\mathbb{R}:$ [redacted] ≥ 0

[redacted]

$\iff \forall a_1,\dots,a_6\in\mathbb{R}:$ [redacted] $\begin{pmatrix}1\\ \vdots\\ x_2^2\end{pmatrix}$ [redacted] $\begin{pmatrix}a_1\\ \vdots\\ a_6\end{pmatrix}\geq 0$

row vector

scalar

column vector

row vector

column vector

symmetric matrix

symmetric matrix

$\iff \begin{pmatrix} 1+2x_1+\dots-y_{17} & & \cdot \\ & \cdot & \\ \cdot & & y_5+\dots \end{pmatrix} \succeq 0$

psd

[redacted] matrix Inequality
[redacted]

Reformulation and linearization technique

(RLT)

LINEARIZE ↓

$$1 + 2x_1 + \dots + y_{16} - y_{17} \geq 0$$

[redacted]

$$\Longleftrightarrow \forall a_1, \dots, a_6 \in \mathbb{R} : \text{[redacted]} \geq 0$$

[redacted]

EXPAND AND LINEARIZE

$$\Longleftrightarrow \forall a_1, \dots, a_6 \in \mathbb{R} : \text{[redacted]} \begin{pmatrix} 1 \\ x \\ \vdots \\ x_2^2 \end{pmatrix} \text{[redacted]} \begin{pmatrix} a_1 \\ \vdots \\ a_6 \end{pmatrix} \geq 0$$

↑ row vector

↑ scalar

↑ column vector

↑ row vector

↑ column vector

symmetric matrix

symmetric matrix

LINEARIZE ↓

$$\Longleftrightarrow \begin{pmatrix} 1+2x_1+\dots-y_{17} & & \cdot \\ & \cdot & \\ \cdot & & y_5+\dots \end{pmatrix} \succeq 0$$

↑ psd

linear

[redacted] matrix Inequality

[redacted]

Lasserre relaxation

$\mathbb{R}[\underline{X}] := \mathbb{R}[X_1, \dots, X_n]$

$g = (g_1, \dots, g_m) \in \mathbb{R}[\underline{X}]^m$

$$\left.\begin{array}{c} g_1(x) \geq 0 \\ \vdots \\ g_m(x) \geq 0 \end{array}\right] \text{SPI}$$

Lasserre relaxation

$$\mathbb{R}[\underline{X}] := \mathbb{R}[X_1, \dots, X_n]$$

$$g = (g_1, \dots, g_m) \in \mathbb{R}[\underline{X}]^m$$

$$\left.\begin{array}{r} 1 \geq 0 \\ g_1(x) \geq 0 \\ \vdots \\ g_m(x) \geq 0 \end{array}\right] \text{SPI} \quad \overset{\text{reformulation}}{\rightsquigarrow} \quad \left.\begin{array}{r} G_0(x) \succeq 0 \\ G_1(x) \succeq 0 \\ \vdots \\ G_m(x) \succeq 0 \end{array}\right] \text{PMI}$$

Lasserre relaxation

$\mathbb{R}[\underline{X}] := \mathbb{R}[X_1, \dots, X_n]$

$g = (g_1, \dots, g_m) \in \mathbb{R}[\underline{X}]^m$

$d \in \mathbb{N}_0$

„relaxation degree"

of degree d or $d-1$

$$\left.\begin{array}{r} 1 \geq 0 \\ g_1(x) \geq 0 \\ \vdots \\ g_m(x) \geq 0 \end{array}\right] \text{SPI} \xrightarrow{\text{reformulation}} \left.\begin{array}{r} G_0(x) \succeq 0 \\ G_1(x) \succeq 0 \\ \vdots \\ G_m(x) \succeq 0 \end{array}\right] \text{PMI} \xrightarrow{\text{linearization}}$$

Lasserre relaxation

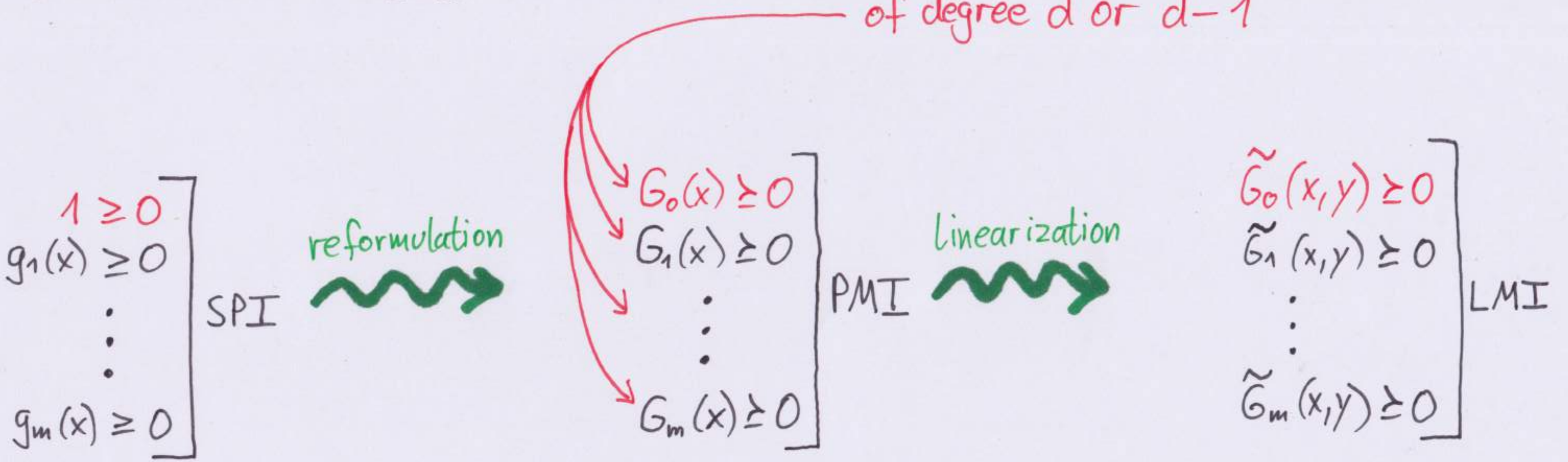
$\mathbb{R}[\underline{X}] := \mathbb{R}[X_1, \dots, X_n]$

$g = (g_1, \dots, g_m) \in \mathbb{R}[\underline{X}]^m$

$d \in \mathbb{N}_0$

„relaxation degree"

of degree d or d−1



$$\left.\begin{array}{r} 1 \geq 0 \\ g_1(x) \geq 0 \\ \vdots \\ g_m(x) \geq 0 \end{array}\right] \text{SPI} \quad \overset{\text{reformulation}}{\rightsquigarrow} \quad \left.\begin{array}{r} G_0(x) \succeq 0 \\ G_1(x) \succeq 0 \\ \vdots \\ G_m(x) \succeq 0 \end{array}\right] \text{PMI} \quad \overset{\text{linearization}}{\rightsquigarrow} \quad \left.\begin{array}{r} \tilde{G}_0(x,y) \succeq 0 \\ \tilde{G}_1(x,y) \succeq 0 \\ \vdots \\ \tilde{G}_m(x,y) \succeq 0 \end{array}\right] \text{LMI}$$

Lasserre relaxation

$\mathbb{R}[\underline{X}] := \mathbb{R}[X_1,\dots,X_n]$

$g = (g_1,\dots,g_m) \in \mathbb{R}[\underline{X}]^m$, $f \in \mathbb{R}[\underline{X}]$

constraints objective

$d \in \mathbb{N}_0$

„relaxation degree"

of degree d or $d-1$

minimize $f(x)$
over all $x \in \mathbb{R}^n$
s.t. $1 \geq 0$
$g_1(x) \geq 0$
$\vdots$
$g_m(x) \geq 0$

SPI

POP

reformulation

minimize $f(x)$
over all $x \in \mathbb{R}^n$
s.t. $G_0(x) \succeq 0$
$G_1(x) \succeq 0$
$\vdots$
$G_m(x) \succeq 0$

PMI

linearization

minimize $\tilde{f}(x,y)$
over all $x \in \mathbb{R}^n, y \in \mathbb{R}^?$
s.t. $\tilde{G}_0(x,y) \succeq 0$
$\tilde{G}_1(x,y) \succeq 0$
$\vdots$
$\tilde{G}_m(x,y) \succeq 0$

LMI

SD

Lasserre relaxation

$\mathbb{R}[\underline{X}] := \mathbb{R}[X_1,\dots,X_n]$

$g = (g_1,\dots,g_m) \in \mathbb{R}[\underline{X}]^m$, $f \in \mathbb{R}[\underline{X}]$

constraints — objective

$d \in \mathbb{N}_0$

„relaxation degree"

of degree d or d−1

minimize $f(x)$
over all $x \in \mathbb{R}^n$
s.t. $1 \geq 0$
$g_1(x) \geq 0$
$\vdots$
$g_m(x) \geq 0$

SPI

POP

reformulation $\rightsquigarrow$

minimize $f(x)$
over all $x \in \mathbb{R}^n$
s.t. $G_0(x) \succeq 0$
$G_1(x) \succeq 0$
$\vdots$
$G_m(x) \succeq 0$

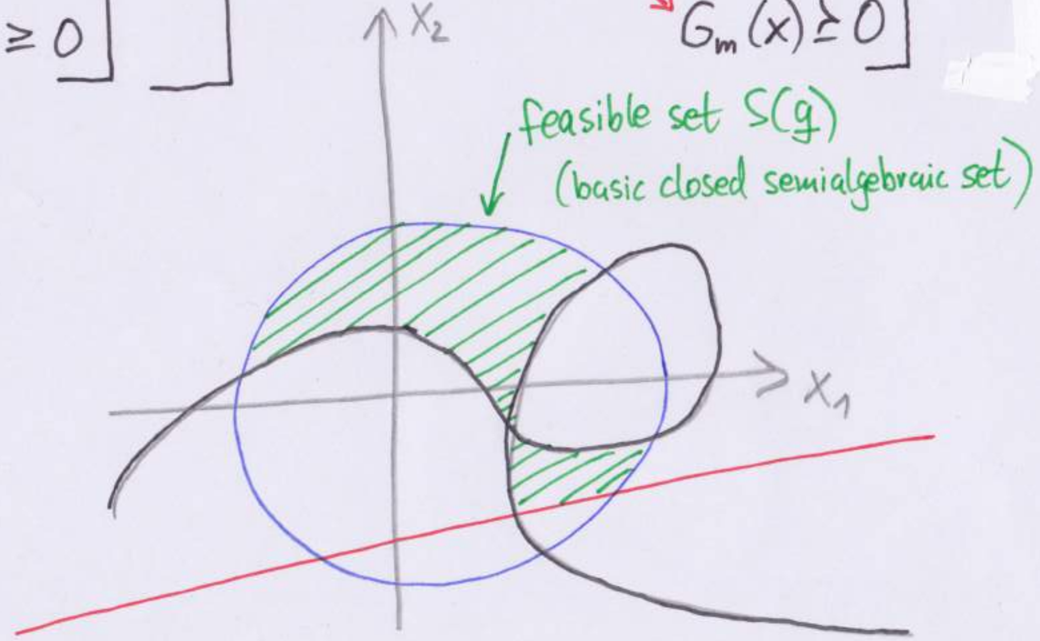
PMI

linearization $\rightsquigarrow$

minimize $\tilde{f}(x,y)$
over all $x \in \mathbb{R}^n, y \in \mathbb{R}^{?}$
s.t. $\tilde{G}_0(x,y) \succeq 0$
$\tilde{G}_1(x,y) \succeq 0$
$\vdots$
$\tilde{G}_m(x,y) \succeq 0$

LMI

SD



feasible set $S(g)$
(basic closed semialgebraic set)

Lasserre relaxation

$\mathbb{R}[\underline{X}] := \mathbb{R}[X_1,\dots,X_n]$

$g = (g_1,\dots,g_m) \in \mathbb{R}[\underline{X}]^m$, $f \in \mathbb{R}[\underline{X}]$

constraints, objective

$d \in \mathbb{N}_0$

„relaxation degree"

of degree d or $d-1$

minimize $f(x)$
over all $x \in \mathbb{R}^n$
s.t. $1 \geq 0$
$g_1(x) \geq 0$
$\vdots$
$g_m(x) \geq 0$

SPI

POP

reformulation

minimize $f(x)$
over all $x \in \mathbb{R}^n$
s.t. $G_0(x) \succeq 0$
$G_1(x) \succeq 0$
$\vdots$
$G_m(x) \succeq 0$

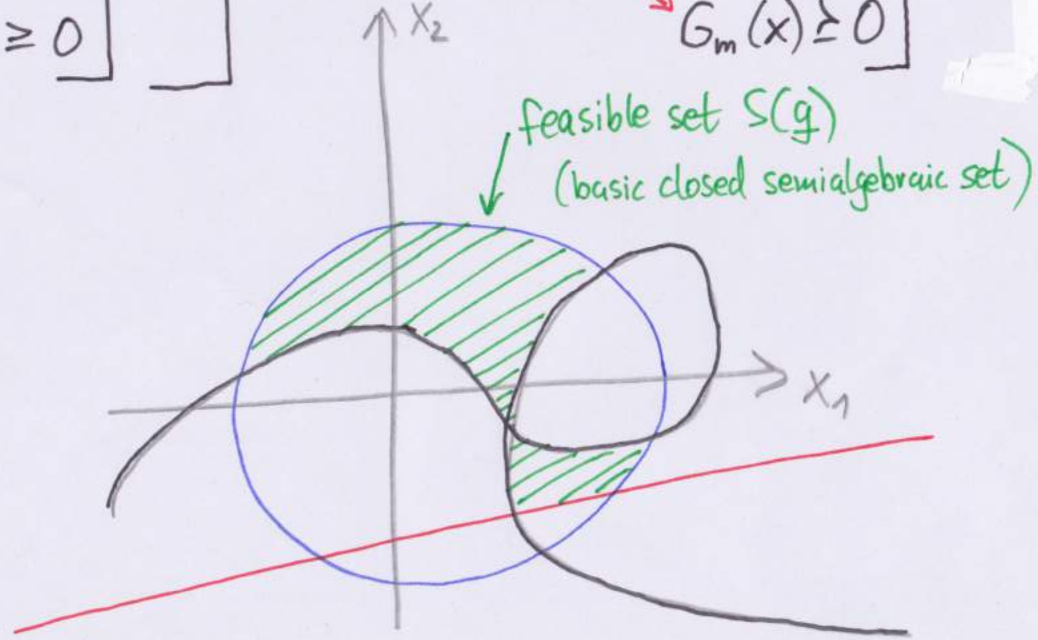
PMI

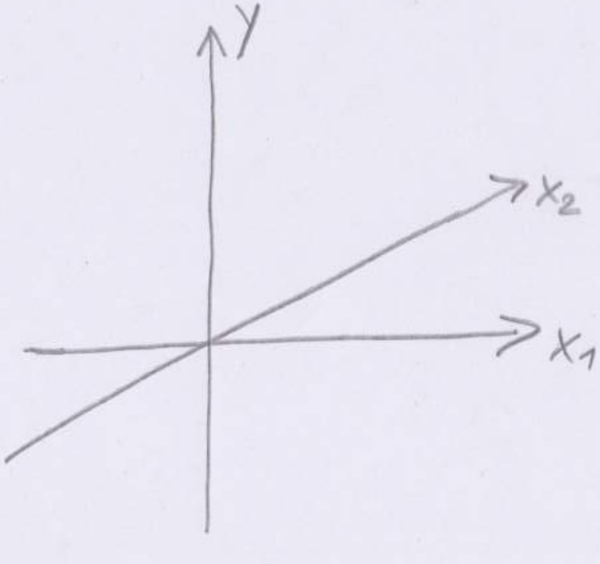
linearization

minimize $\tilde{f}(x,y)$
over all $x \in \mathbb{R}^n, y \in \mathbb{R}^?$
s.t. $\tilde{G}_0(x,y) \succeq 0$
$\tilde{G}_1(x,y) \succeq 0$
$\vdots$
$\tilde{G}_m(x,y) \succeq 0$

LMI

SD





Lasserre relaxation

$\mathbb{R}[\underline{X}] := \mathbb{R}[X_1, \dots, X_n]$

$g = (g_1, \dots, g_m) \in \mathbb{R}[\underline{X}]^m$, $f \in \mathbb{R}[\underline{X}]$

constraints, objective

$d \in \mathbb{N}_0$

„relaxation degree"

of degree d or $d-1$

minimize $f(x)$
over all $x \in \mathbb{R}^n$
s.t. $1 \geq 0$
$g_1(x) \geq 0$
$\vdots$
$g_m(x) \geq 0$

SPI

POP

reformulation $\rightsquigarrow$

minimize $f(x)$
over all $x \in \mathbb{R}^n$
s.t. $G_0(x) \succeq 0$
$G_1(x) \succeq 0$
$\vdots$
$G_m(x) \succeq 0$

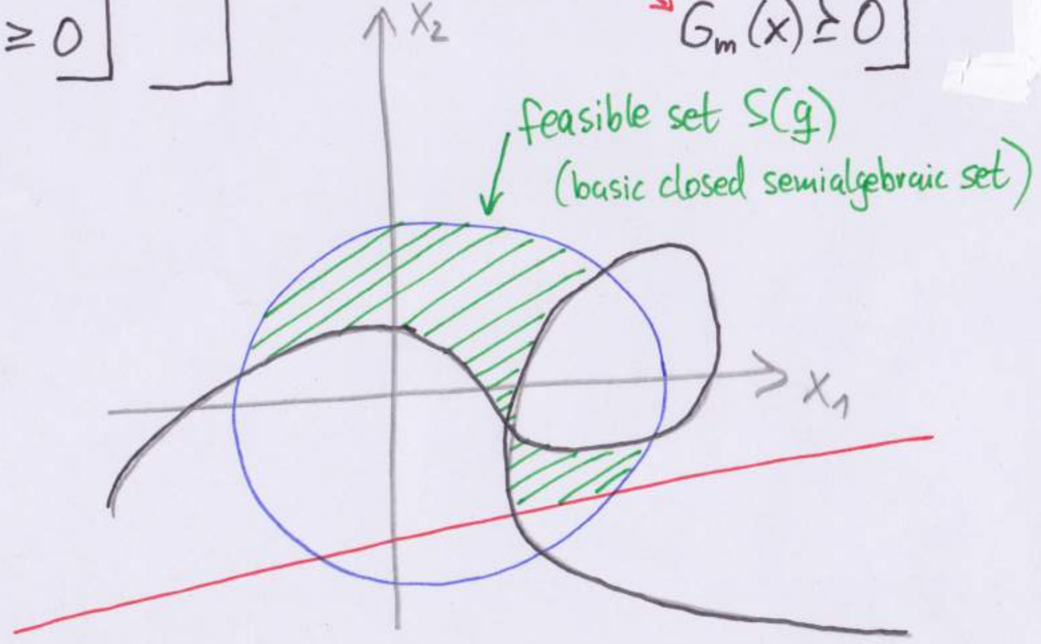
PMI

linearization $\rightsquigarrow$

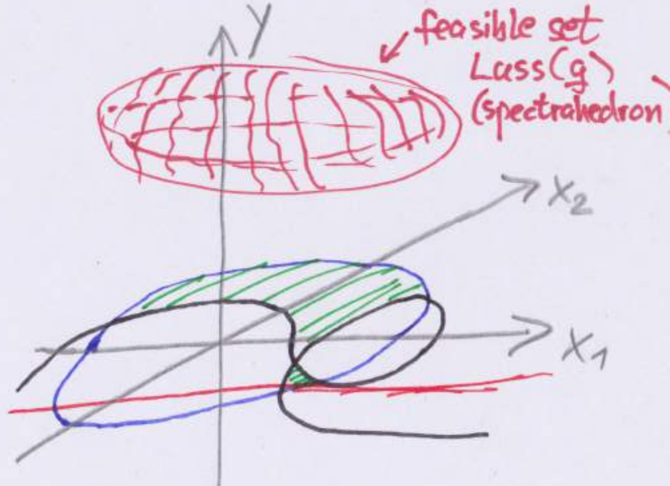
minimize $\tilde{f}(x,y)$
over all $x \in \mathbb{R}^n, y \in \mathbb{R}^{?}$
s.t. $\tilde{G}_0(x,y) \succeq 0$
$\tilde{G}_1(x,y) \succeq 0$
$\vdots$
$\tilde{G}_m(x,y) \succeq 0$

LMI

SD



feasible set $S(g)$
(basic closed semialgebraic set)



feasible set
$\mathrm{Lass}(g)$
(spectrahedron)

Lasserre relaxation

$\mathbb{R}[\underline{X}] := \mathbb{R}[X_1,\dots,X_n]$

$\underline{g} = (g_1,\dots,g_m) \in \mathbb{R}[\underline{X}]^m, \ f \in \mathbb{R}[\underline{X}]$

constraints objective

$d \in \mathbb{N}_0$

„relaxation degree"

of degree d or d−1

minimize $f(x)$
over all $x \in \mathbb{R}^n$
s.t. $1 \geq 0$
$g_1(x) \geq 0$
$\vdots$
$g_m(x) \geq 0$

SPI

POP

reformulation $\rightsquigarrow$

minimize $f(x)$
over all $x \in \mathbb{R}^n$
s.t. $G_0(x) \succeq 0$
$G_1(x) \succeq 0$
$\vdots$
$G_m(x) \succeq 0$

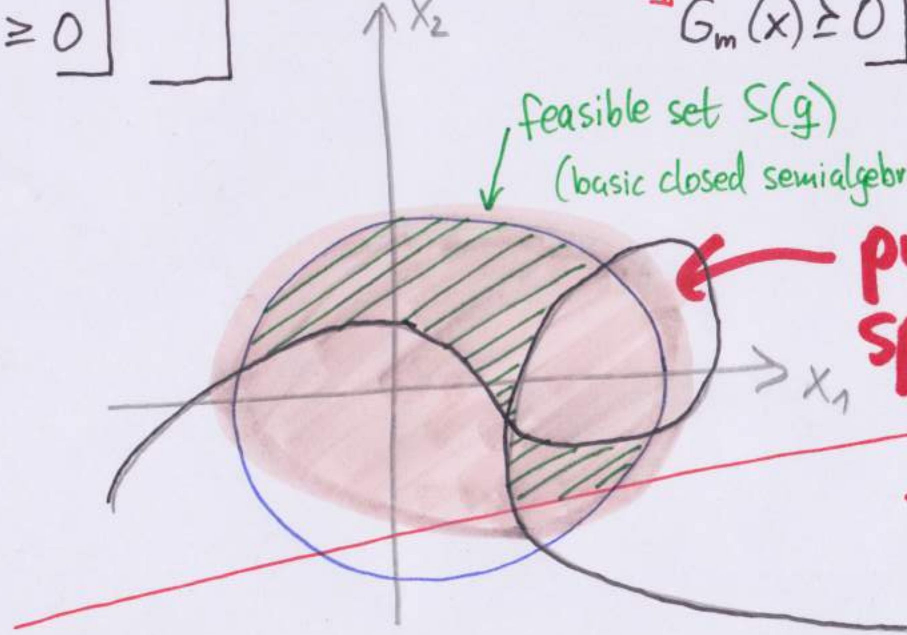
PMI

linearization $\rightsquigarrow$

minimize $\tilde{f}(x,y)$
over all $x \in \mathbb{R}^n, y \in \mathbb{R}^?$
s.t. $\tilde{G}_0(x,y) \succeq 0$
$\tilde{G}_1(x,y) \succeq 0$
$\vdots$
$\tilde{G}_m(x,y) \succeq 0$

LMI

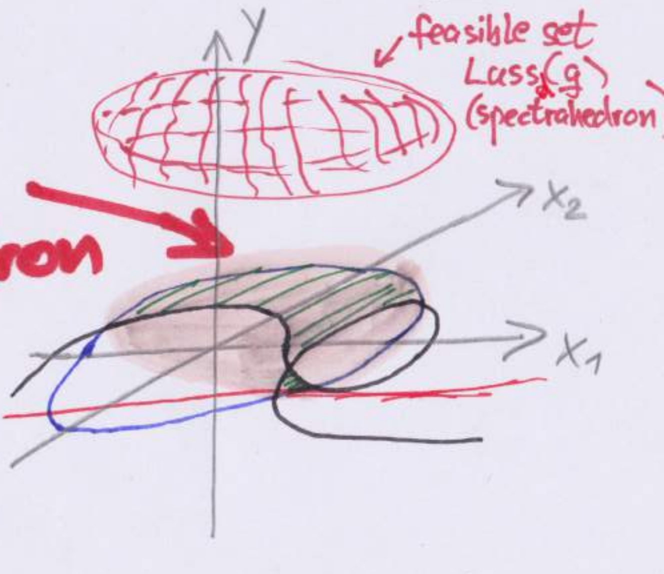
SD



feasible set $S(\underline{g})$
(basic closed semialgebraic set)

projected spectrahedron

$S_d(\underline{g})$



feasible set
$\mathrm{Lass}_d(\underline{g})$
(spectrahedron)

Lasserre in an abstract way

$g \in \mathbb{R}[\underline{X}]^m$, $f \in \mathbb{R}[\underline{X}]$, $d \in \mathbb{N}_0$

constraints objective relaxation degree

Lasserre in an abstract way

$g \in \mathbb{R}[\underline{X}]^m$, $f \in \mathbb{R}[\underline{X}]$, $d \in \mathbb{N}_0$

constraints objective relaxation degree

$S(g) := \{x \in \mathbb{R}^n \mid g_1(x) \geq 0, \dots, g_m(x) \geq 0\}$ „basic closed semialgebraic set"

$\mathrm{opt}(f, g) := \inf\{f(x) \mid x \in S(g)\} \in \{-\infty\} \cup \mathbb{R} \cup \{\infty\}$ „optimal value"

Lasserre in an abstract way

$\underline{g} \in \mathbb{R}[\underline{X}]^m$, $f \in \mathbb{R}[\underline{X}]$, $d \in \mathbb{N}_0$

constraints objective relaxation degree

$S(\underline{g}) := \{x \in \mathbb{R}^n \mid g_1(x) \geq 0, \dots, g_m(x) \geq 0\}$ „basic closed semialgebraic set"

$\operatorname{opt}(f, \underline{g}) := \inf\{f(x) \mid x \in S(\underline{g})\} \in \{-\infty\} \cup \mathbb{R} \cup \{\infty\}$ „optimal value"

$M_d(\underline{g}) := \left\{ \sum_{i=0}^{m} \sum_{j} p_{ij}^2 g_i \mid p_{ij} \in \mathbb{R}[\underline{X}], \deg(p_{ij}^2 g_i) \leq d \right\} \subseteq \mathbb{R}[\underline{X}]_d$

where $g_0 := 1 \in \mathbb{R}[\underline{X}]$ „truncated quadratic module"

$M(\underline{g}) := \bigcup_{d \in \mathbb{N}_0} M_d(\underline{g})$ „quadratic module"

Lasserre in an abstract way

$g \in \mathbb{R}[\underline{X}]^m$, $f \in \mathbb{R}[\underline{X}]$, $d \in \mathbb{N}_0$

constraints objective relaxation degree

$S(g) := \{x \in \mathbb{R}^n \mid g_1(x) \geq 0, \dots, g_m(x) \geq 0\}$ „basic closed semialgebraic set"

$\operatorname{opt}(f,g) := \inf\{f(x) \mid x \in S(g)\} \in \{-\infty\} \cup \mathbb{R} \cup \{\infty\}$ „optimal value"

$M_d(g) := \left\{ \sum_{i=0}^{m} \sum_{j} p_{ij}^2 g_i \;\middle|\; p_{ij} \in \mathbb{R}[\underline{X}], \deg(p_{ij}^2 g_i) \leq d \right\} \subseteq \mathbb{R}[\underline{X}]_d$

where $g_0 := 1 \in \mathbb{R}[\underline{X}]$ „truncated quadratic module"

$M(g) := \bigcup_{d \in \mathbb{N}_0} M_d(g)$ „quadratic module"

$\operatorname{Lass}_d(g) := \{L \mid L : \mathbb{R}[\underline{X}]_d \to \mathbb{R} \text{ linear}, L(M_d(g)) \subseteq \mathbb{R}_{\geq 0}, L(1) = 1\}$ „Lasserre spectrahedron"

$S_d(g) := \{(L(X_1), \dots, L(X_n)) \mid L \in \operatorname{Lass}_d(g)\}$ „projected Lasserre spectrahedron"

$\operatorname{opt}_d(f,g) := \inf\{L(f) \mid L \in \operatorname{Lass}_d(g)\}$ „relaxed optimal value"

Lasserre in an abstract way

$g \in \mathbb{R}[\underline{X}]^m$, $f \in \mathbb{R}[\underline{X}]$, $d \in \mathbb{N}_0$

constraints objective relaxation degree

$S(g) := \{x \in \mathbb{R}^n \mid g_1(x) \geq 0, \dots, g_m(x) \geq 0\}$ „basic closed semialgebraic set"

$\operatorname{opt}(f,g) := \inf\{f(x) \mid x \in S(g)\} \in \{-\infty\} \cup \mathbb{R} \cup \{\infty\}$ „optimal value"

$M_d(g) := \left\{ \sum_{i=0}^{m} \sum_j p_{ij}^2 g_i \mid p_{ij} \in \mathbb{R}[\underline{X}], \deg(p_{ij}^2 g_i) \leq d \right\} \subseteq \mathbb{R}[\underline{X}]_d$

where $g_0 := 1 \in \mathbb{R}[\underline{X}]$ „truncated quadratic module"

$M(g) := \bigcup_{d \in \mathbb{N}_0} M_d(g)$ „quadratic module"

$\operatorname{Lass}_d(g) := \{L \mid L : \mathbb{R}[\underline{X}]_d \to \mathbb{R} \text{ linear}, L(M_d(g)) \subseteq \mathbb{R}_{\geq 0}, L(1) = 1\}$ „Lasserre spectrahedron"

$S_d(g) := \{(L(X_1), \dots, L(X_n)) \mid L \in \operatorname{Lass}_d(g)\}$ „projected Lasserre spectrahedron"

$\operatorname{opt}_d(f,g) := \inf\{L(f) \mid L \in \operatorname{Lass}_d(g)\}$ „relaxed optimal value"

Obviously, $S(g) \subseteq \dots \subseteq S_{d+1}(g) \subseteq S_d(g) \subseteq \dots$ „SPIs through LMIs"

and $\operatorname{opt}(f,g) \geq \dots \geq \operatorname{opt}_{d+1}(f,g) \geq \operatorname{opt}_d(f,g)$ „POPs through SDPs"

History

Let $\underset{\text{constraints}}{g} \in \mathbb{R}[\underline{x}]^m$, $\underset{\text{objective}}{f} \in \mathbb{R}[\underline{x}]$.

Suppose $M(g)$ is Archimedean $\left(\underset{\substack{\text{slightly}\\ \text{change } g}}{\overset{\Longrightarrow}{\leftsquigarrow}} S(g) \text{ is compact} \right)$.

History

Let $g \in \mathbb{R}[\underline{X}]^m$, $f \in \mathbb{R}[\underline{X}]$.

constraints objective

Suppose $M(g)$ is Archimedean $\left(\underset{\text{slightly change } g}{\overset{\Longrightarrow}{\Longleftarrow}} S(g) \text{ is compact} \right)$.

Lasserre 2001 based on Putinar 1993: $\operatorname{opt}_d(f,g) \xrightarrow{d\to\infty} \operatorname{opt}(f,g)$

Lasserre 2009 based on Prestel 2001 : $S_d(f,g) \xrightarrow{d\to\infty}$ conv $S(f,g)$

Nie & S. 2007

History

Let $\overset{\text{constraints}}{g \in \mathbb{R}[\underline{x}]^m}$, $\overset{\text{objective}}{f \in \mathbb{R}[\underline{x}]}$.

Suppose $M(g)$ is Archimedean $\left(\underset{\text{slightly change } g}{\overset{\Longrightarrow}{\leftsquigarrow}} \right.$ $S(g)$ is compact$\left. \right)$.

Always:

Lasserre 2001 based on Putinar 1993: $\mathrm{opt}_d(f,g) \xrightarrow{d\to\infty} \mathrm{opt}(f,g)$

Lasserre 2009 based on Prestel 2001, Nie & S. 2007: $S_d(f,g) \xrightarrow{d\to\infty} \mathrm{conv}\ S(f,g)$

Under mild conditions:

Scheiderer 2005: For large d, $\mathrm{opt}_d(f,g) = \mathrm{opt}(f,g)$.

Helton & Nie 2009: For large d, conv $S(g)$ is a projected spectrahedron

UNDER MILD CONDITIONS

History

Let $\underset{\text{constraints}}{g} \in \mathbb{R}[\underline{x}]^m$, $\underset{\text{objective}}{f} \in \mathbb{R}[\underline{x}]$.

Suppose $M(g)$ is Archimedean $\left(\underset{\text{slightly change } g}{\overset{\Longrightarrow}{\leftsquigarrow}} S(g) \text{ is compact} \right)$.

Always:

Lasserre 2001 based on Putinar 1993: $\operatorname{opt}_d(f,g) \xrightarrow{d\to\infty} \operatorname{opt}(f,g)$

Lasserre 2009 based on Prestel 2001 : $S_d(f,g) \xrightarrow{d\to\infty} \operatorname{conv} S(f,g)$
Nie & S. 2007

Under mild conditions:

Scheiderer 2005 : For large d, $\operatorname{opt}_d(f,g) = \operatorname{opt}(f,g)$.

Helton & Nie 2009: For large d, $\operatorname{conv} S(g)$ is a projected spectrahedron.

Kriel & S. 2018 SIOPT : $S(g)$ convex $\Rightarrow$ For large d, $S(g) = S_d(g)$.

Kriel & S. FoCM : For large d, $\operatorname{conv} S(g) = S_d(g)$.

UNDER MILD CONDITIONS

History

Let $\underset{\text{constraints}}{g} \in \mathbb{R}[\underline{x}]^m$, $\underset{\text{objective}}{f} \in \mathbb{R}[\underline{x}]$.

Suppose $M(g)$ is Archimedean $\left(\underset{\text{slightly change } g}{\overset{\Longrightarrow}{\Longleftarrow}} S(g) \text{ is compact} \right)$.

Always:

Lasserre 2001 based on Putinar 1993: $\mathrm{opt}_d(f,g) \xrightarrow{d\to\infty} \mathrm{opt}(f,g)$

Lasserre 2009 based on Prestel 2001, Nie & S. 2007: $S_d(f,g) \xrightarrow{d\to\infty} \mathrm{conv}\ S(f,g)$

Under mild conditions:

Scheiderer 2005: For large d, $\mathrm{opt}_d(f,g) = \mathrm{opt}(f,g)$.

Helton & Nie 2009: For large d, conv $S(g)$ is a projected spectrahedron.

Kriel & S. 2018 SIOPT: $S(g)$ convex $\Rightarrow$ For large d, $S(g) = S_d(g)$.

Linear g_i are fine. / a problem.

Kriel & S. FoCM: For large d, conv $S(g) = S_d(g)$.

UNDER MILD CONDITIONS

History

Let $g \in \mathbb{R}[\underline{x}]^m$, $f \in \mathbb{R}[\underline{x}]$.

constraints objective

Suppose $M(g)$ is Archimedean $\Big(\overset{\Longrightarrow}{\underset{\text{slightly change } g}{\Longleftarrow}}$ $S(g)$ is compact$\Big)$.

Always:

Lasserre 2001 based on Putinar 1993: ● $\mathrm{opt}_d(f,g) \xrightarrow{d\to\infty} \mathrm{opt}(f,g)$

Lasserre 2009 based on Prestel 2001, Nie&S. 2007: ● $S_d(f,g) \xrightarrow{d\to\infty}$ conv $S(f,g)$

Under mild conditions:

Scheiderer 2005: ● For large d, $\mathrm{opt}_d(f,g) = \mathrm{opt}(f,g)$.

Helton & Nie 2009: ● For large d, conv $S(g)$ is a projected spectrahedron.

Kriel & S. 2018 SIOPT: ● $S(g)$ convex $\Rightarrow$ For large d, $S(g) = S_d(g)$.

Linear g_i are ~~a problem.~~ fine.

Kriel & S. FoCM: ● For large d, conv $S(g) = S_d(g)$.

proof uses real closed fields

● $\forall g$: $\forall$complexity bounds: $\exists d_0$: $\forall f$ of bounded complexity: $\forall d \geq d_0$: $\mathrm{opt}_d(f,g) = \mathrm{opt}(f,g)$.

The result for convex sets

For $h \in \mathbb{R}[\underline{X}]$ and $x \in \mathbb{R}^n$, we call h <u>strictly quasiconcave</u> at x if

$$\forall v \in \mathbb{R}^n \setminus \{0\} : \left(((\nabla h)(x))^T v = 0 \implies v^T ((\operatorname{Hess} h)(x)) v < 0 \right).$$

„Hessian negative definite on tangent space"

The result for convex sets

For $h \in \mathbb{R}[\underline{X}]$ and $x \in \mathbb{R}^n$, we call h strictly quasiconcave at x if

$$\forall v \in \mathbb{R}^n \setminus \{0\} : \left((\nabla h)(x)^T v = 0 \Longrightarrow v^T((\mathrm{Hess}\, h)(x))\, v < 0 \right).$$

„Hessian negative definite on tangent space"

Rough intuition: mountain hike

mountain: subgraph of h

$\{(x,y) \in \mathbb{R}^n \times \mathbb{R} \mid y \leq h(x)\}$

your ground position: x

Can a bird flying a straight line at constant altitude crash into you?

contour map

The result for convex sets

For $h \in \mathbb{R}[\underline{X}]$ and $x \in \mathbb{R}^n$, we call h strictly quasiconcave at x if

$$\forall v \in \mathbb{R}^n \setminus \{0\} : \left(((\nabla h)(x))^T v = 0 \Longrightarrow v^T ((\operatorname{Hess} h)(x)) v < 0 \right).$$

„Hessian negative definite on tangent space"

We call $h \in \mathbb{R}[\underline{X}]$ $\underline{g}$-sos-concave if there exists a certain sums-of-squares certificate for Hess $h \preceq 0$ on $S(\underline{g})$.

The result for convex sets

For $h \in \mathbb{R}[\underline{X}]$ and $x \in \mathbb{R}^n$, we call h strictly quasiconcave at x if

$$\forall v \in \mathbb{R}^n \setminus \{0\} : \left(((\nabla h)(x))^T v = 0 \Longrightarrow v^T ((\operatorname{Hess} h)(x)) v < 0 \right).$$

„Hessian negative definite on tangent space"

We call $h \in \mathbb{R}[\underline{X}]$ g-sos-concave if there exists a certain sums-of-squares certificate for Hess $h \preceq 0$ on $S(g)$.

Linear polynomials are trivially g-sos-concave!

The result for convex sets

For $h \in \mathbb{R}[\underline{X}]$ and $x \in \mathbb{R}^n$, we call h strictly quasiconcave at x if

$$\forall v \in \mathbb{R}^n \setminus \{0\} : \left(((\nabla h)(x))^T v = 0 \implies v^T ((\operatorname{Hess} h)(x)) v < 0 \right).$$

"Hessian negative definite on tangent space"

We call $h \in \mathbb{R}[\underline{X}]$ $\underline{g}$-sos-concave if there exists a certain sums-of-squares certificate for Hess $h \preceq 0$ on $S(\underline{g})$.

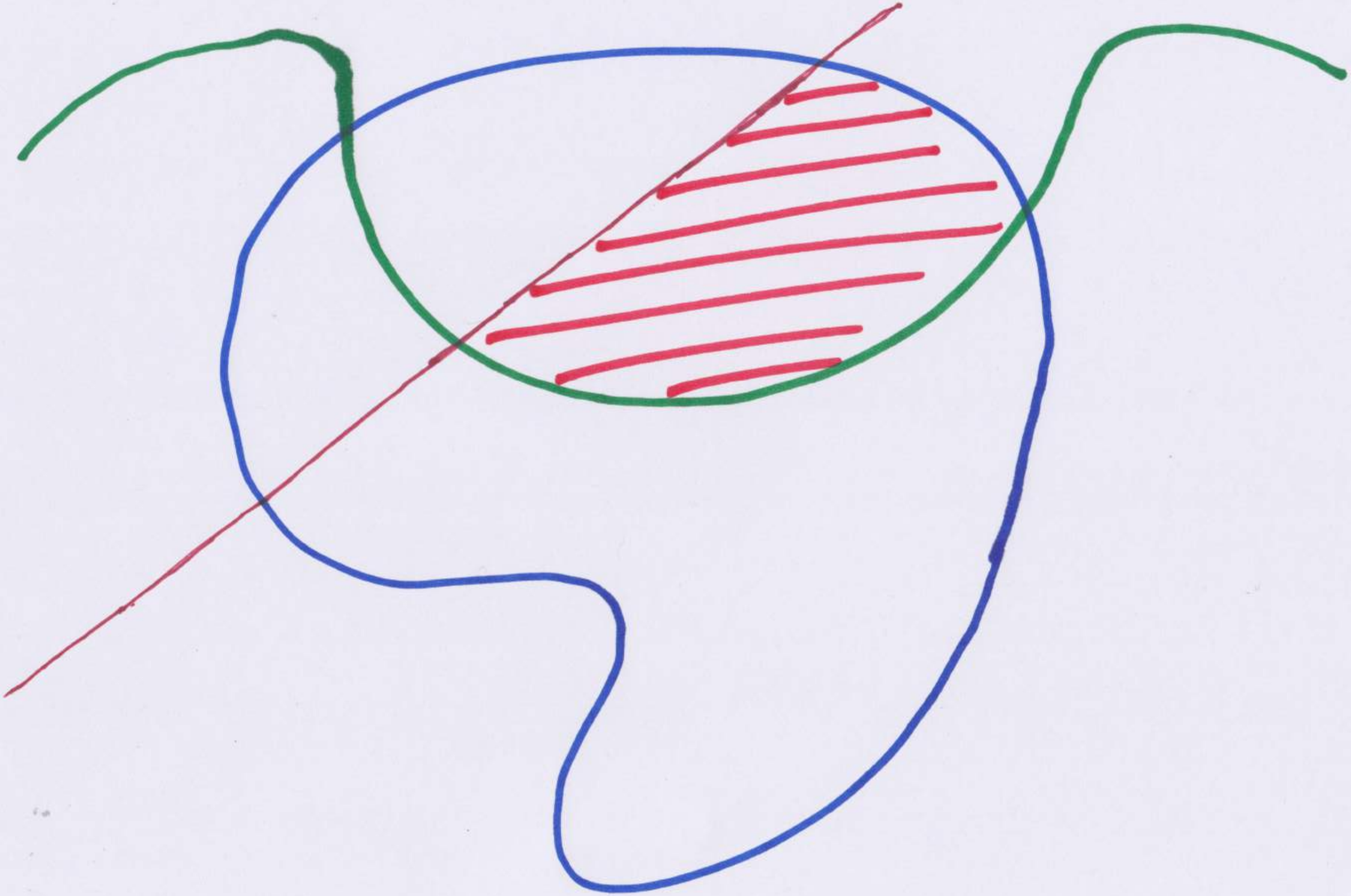
Linear polynomials are trivially $\underline{g}$-sos-concave!

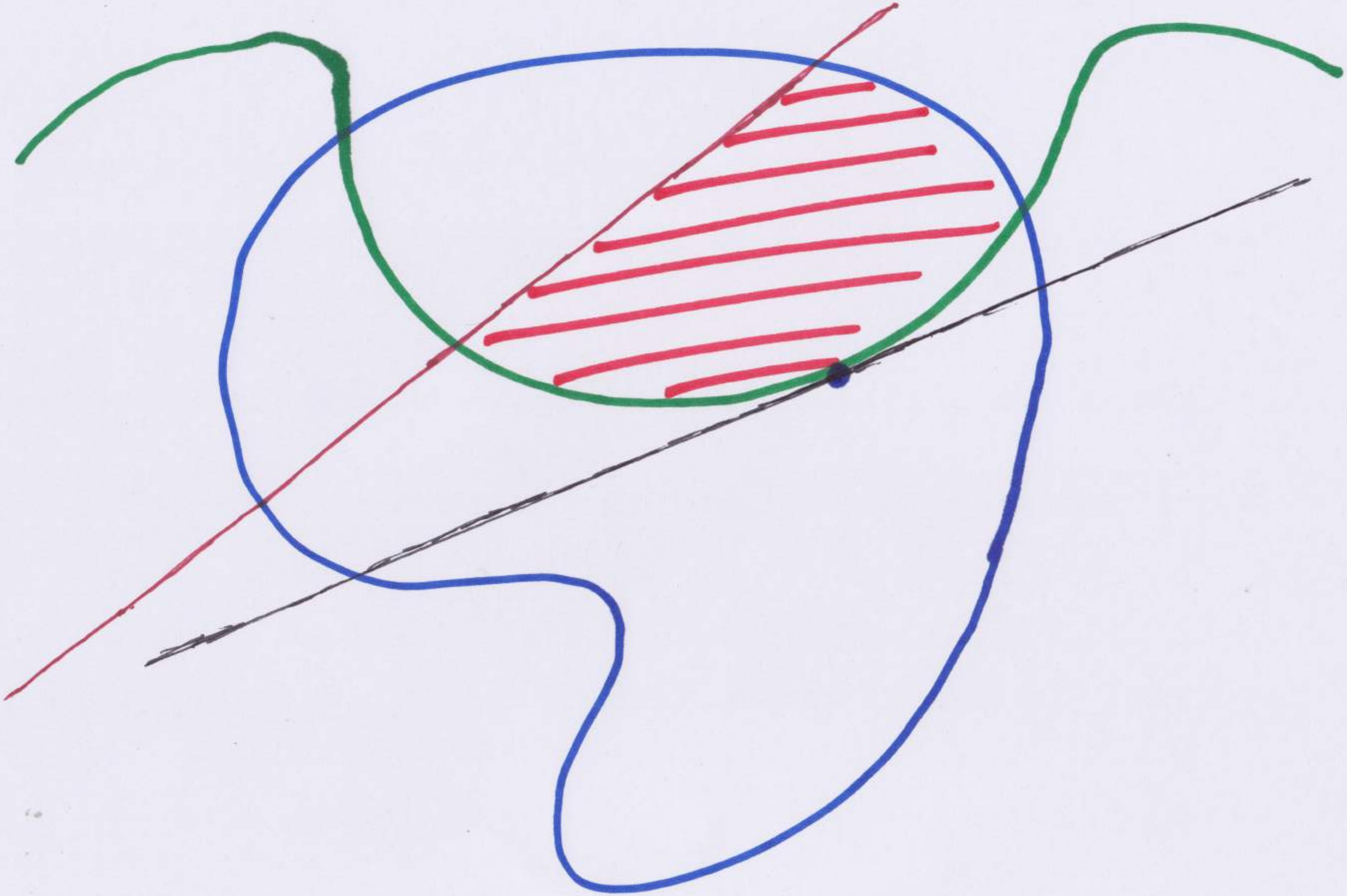
Thm (2018 SIOPT) Let $\underline{g} \in \mathbb{R}[\underline{X}]^m$ such that $M(\underline{g})$ is Archimedean.

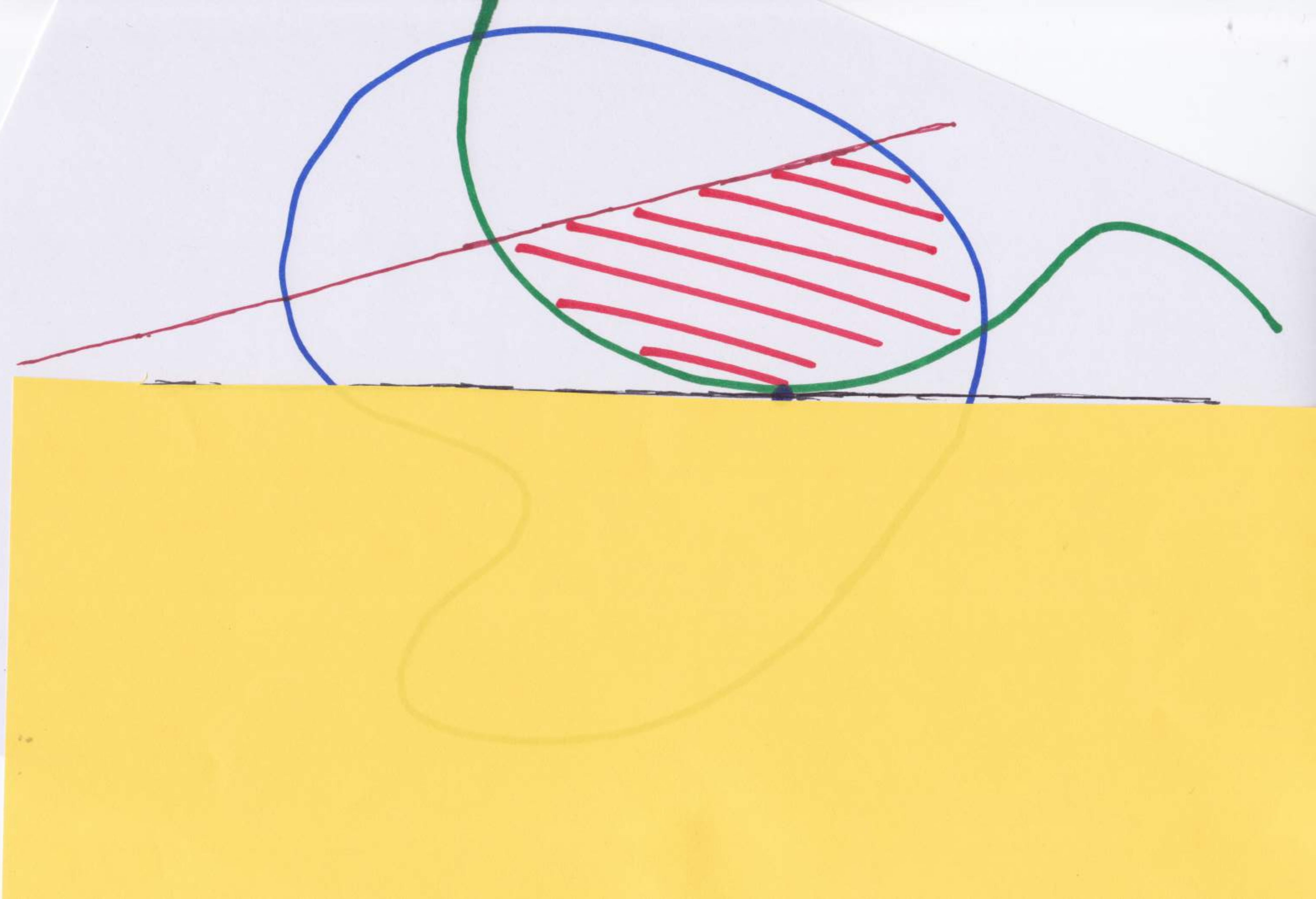
Suppose $S(\underline{g})$ is convex with nonempty interior.

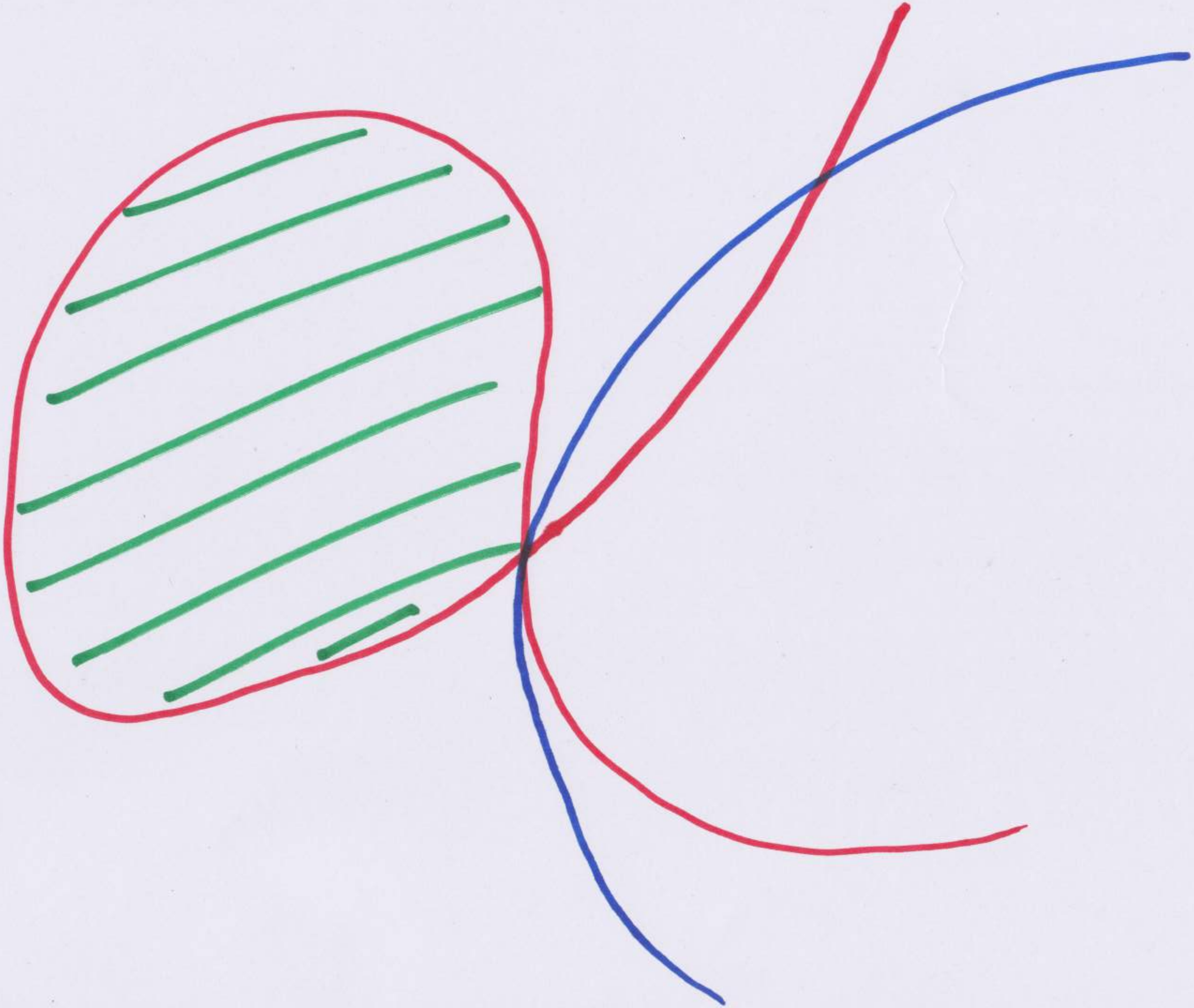
Suppose that each g_i is strictly quasiconcave on $S(\underline{g}) \cap Z(g_i)$ or $\underline{g}$-sos-concave (for example linear).

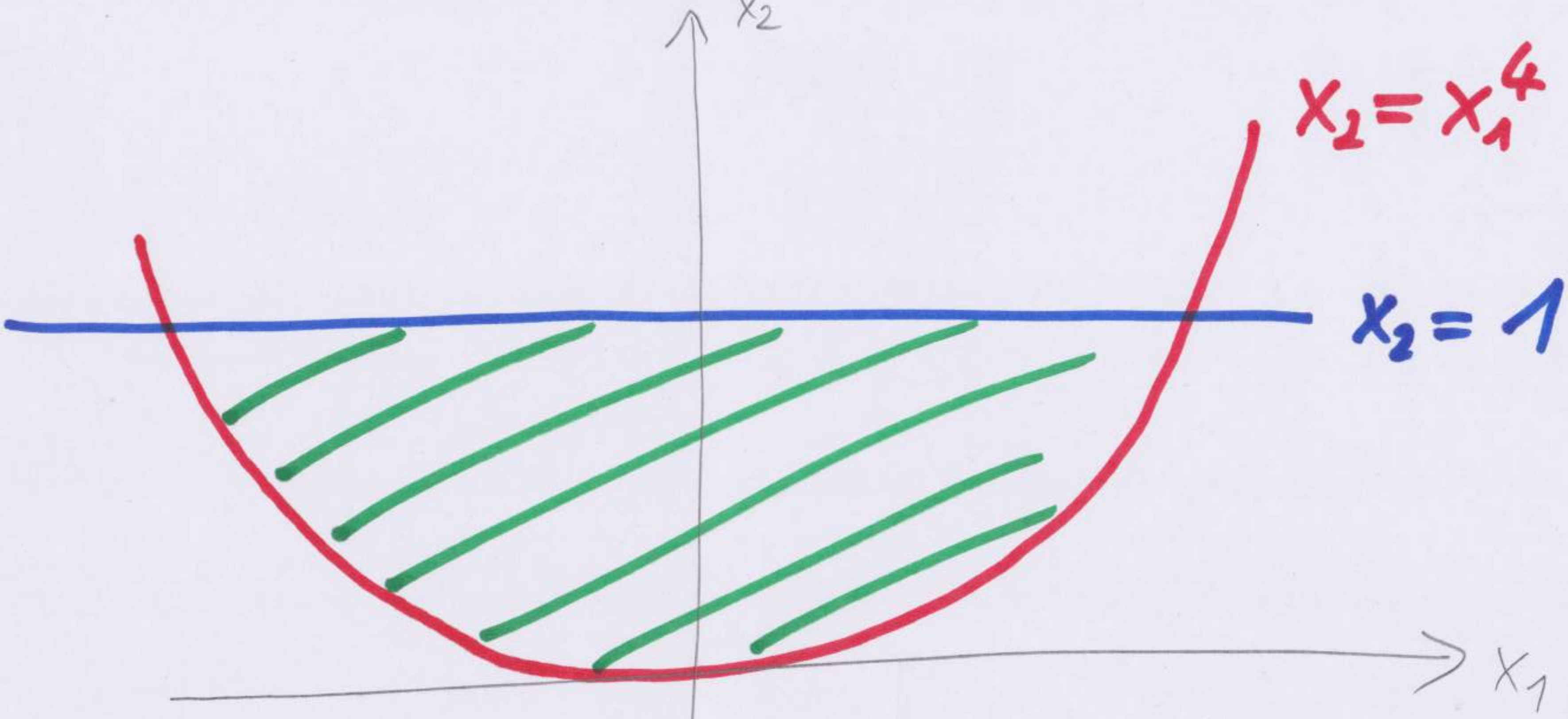
Then $S(\underline{g}) = S_d(\underline{g})$ for large d.



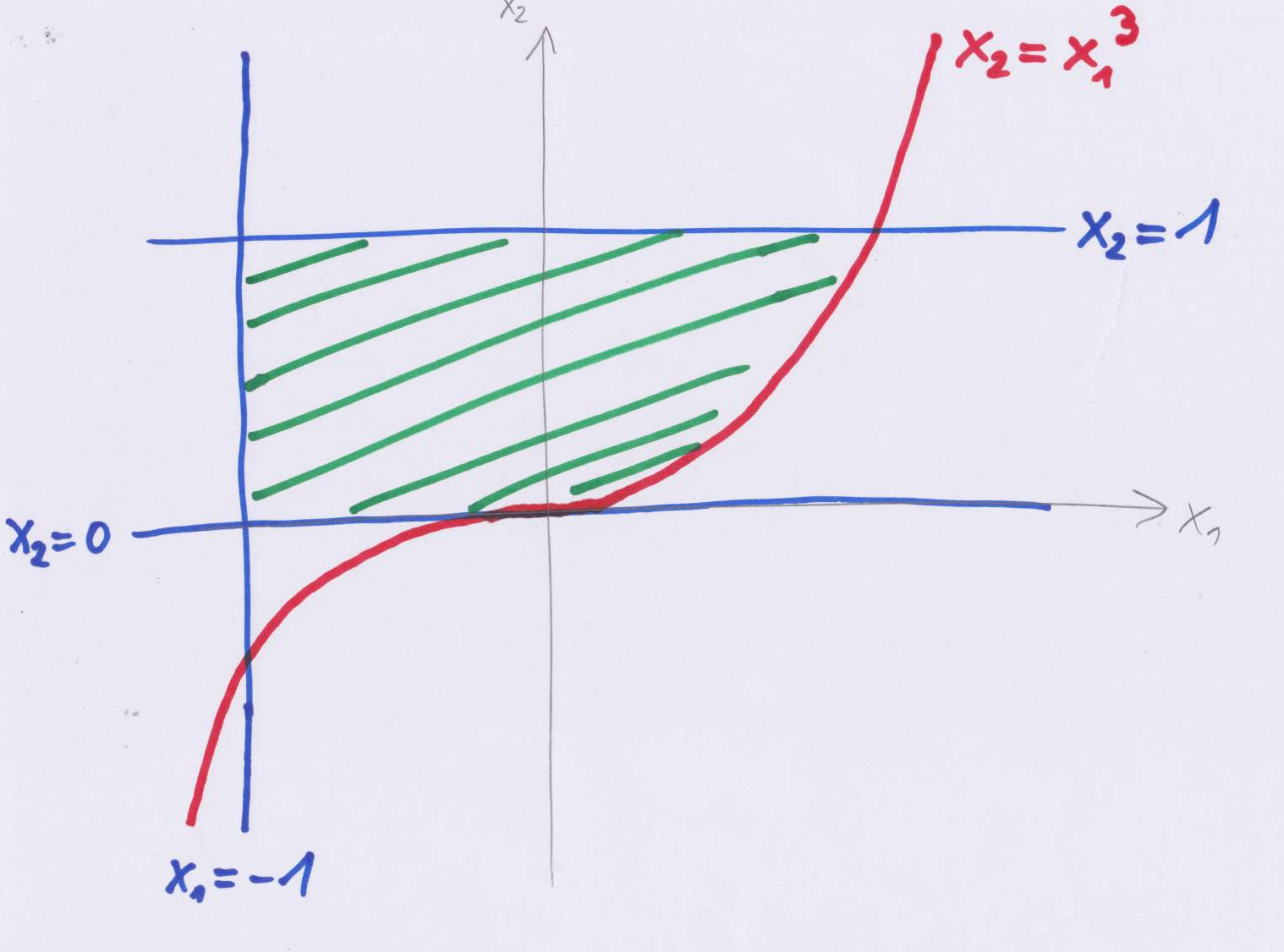


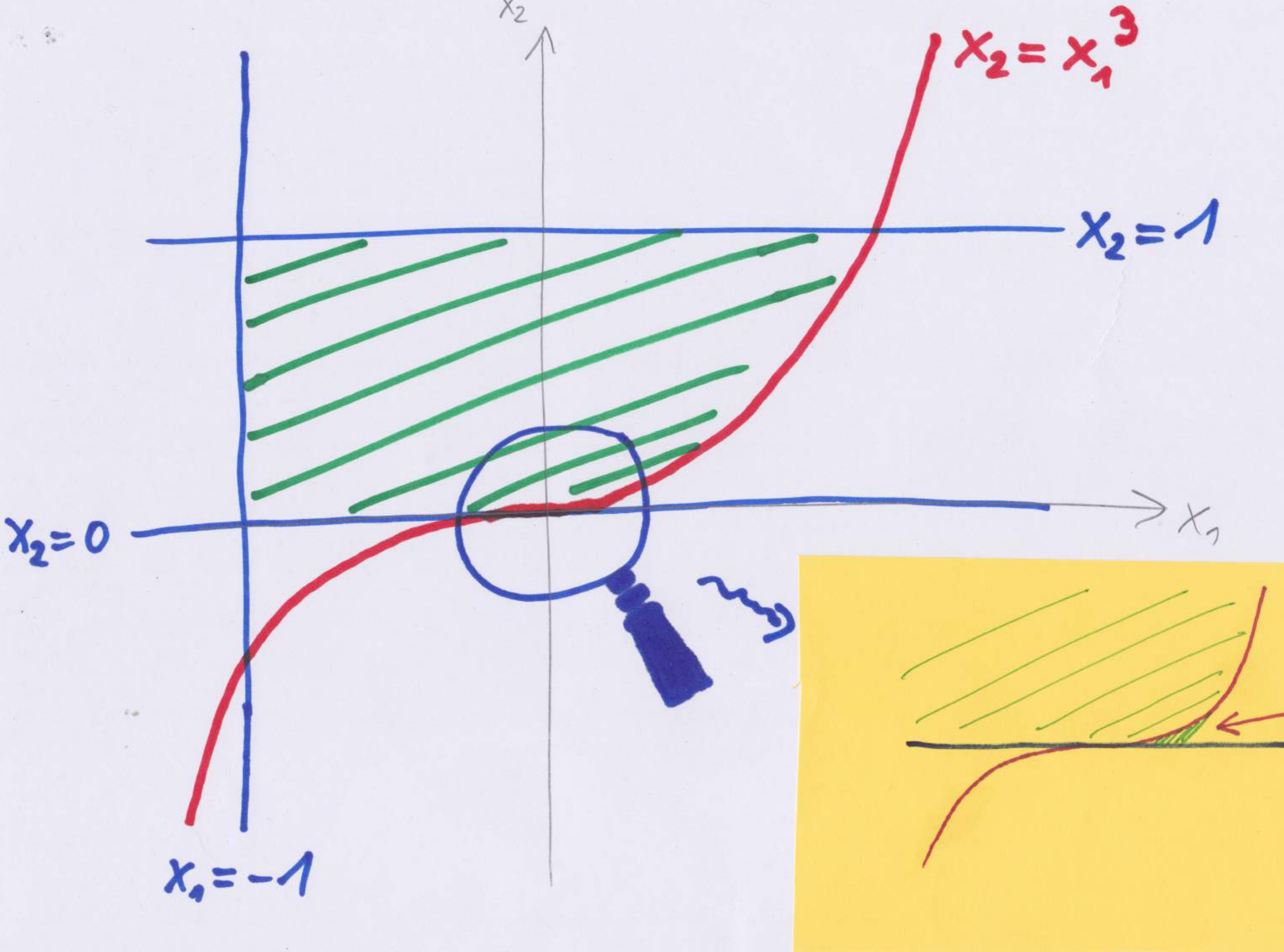






$$\mathrm{Hess}\left(x_2 - x_1^4\right) = \begin{pmatrix} -12x_1^2 & 0 \\ 0 & 0 \end{pmatrix}$$


x_2
$x_2 = x_1^3$
$x_2 = 1$
$x_2 = 0$
x_1
$x_1 = -1$


x_2
$x_2 = x_1^3$
$x_2 = 1$
$x_2 = 0$
x_1
$x_1 = -1$

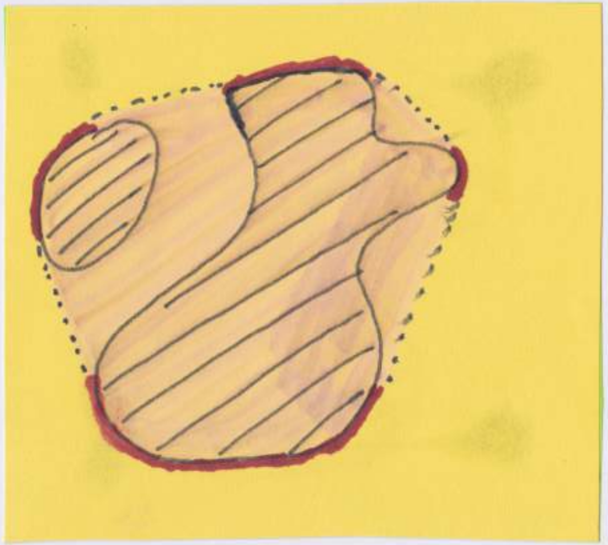
The result for not necessarily convex sets

For $S \subseteq \mathbb{R}^n$, we call

$$\mathrm{convbd}\, S := S \cap \partial \mathrm{conv}\, S$$

the convex boundary of S.

We say that S has nonempty interior near its convex boundary if $\mathrm{convbd}\, S \subseteq \overline{S^\circ}$.



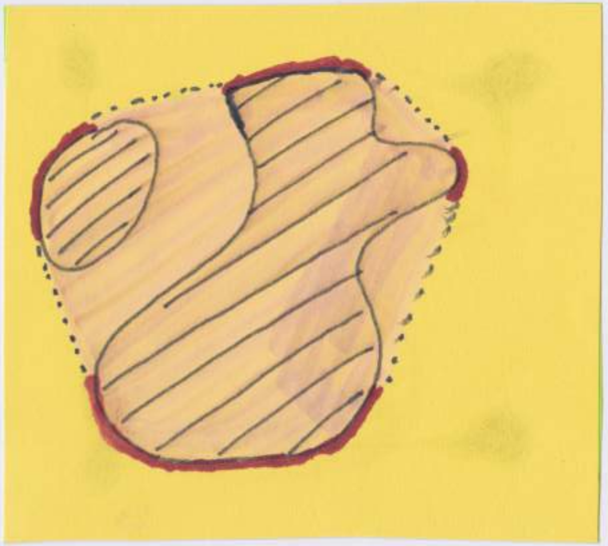
The result for not necessarily convex sets

For $S \subseteq \mathbb{R}^n$, we call

$$\operatorname{convbd} S := S \cap \partial \operatorname{conv} S$$

the *convex boundary* of S.

We say that *S has nonempty interior near its convex boundary* if $\operatorname{convbd} S \subseteq \overline{S^\circ}$.

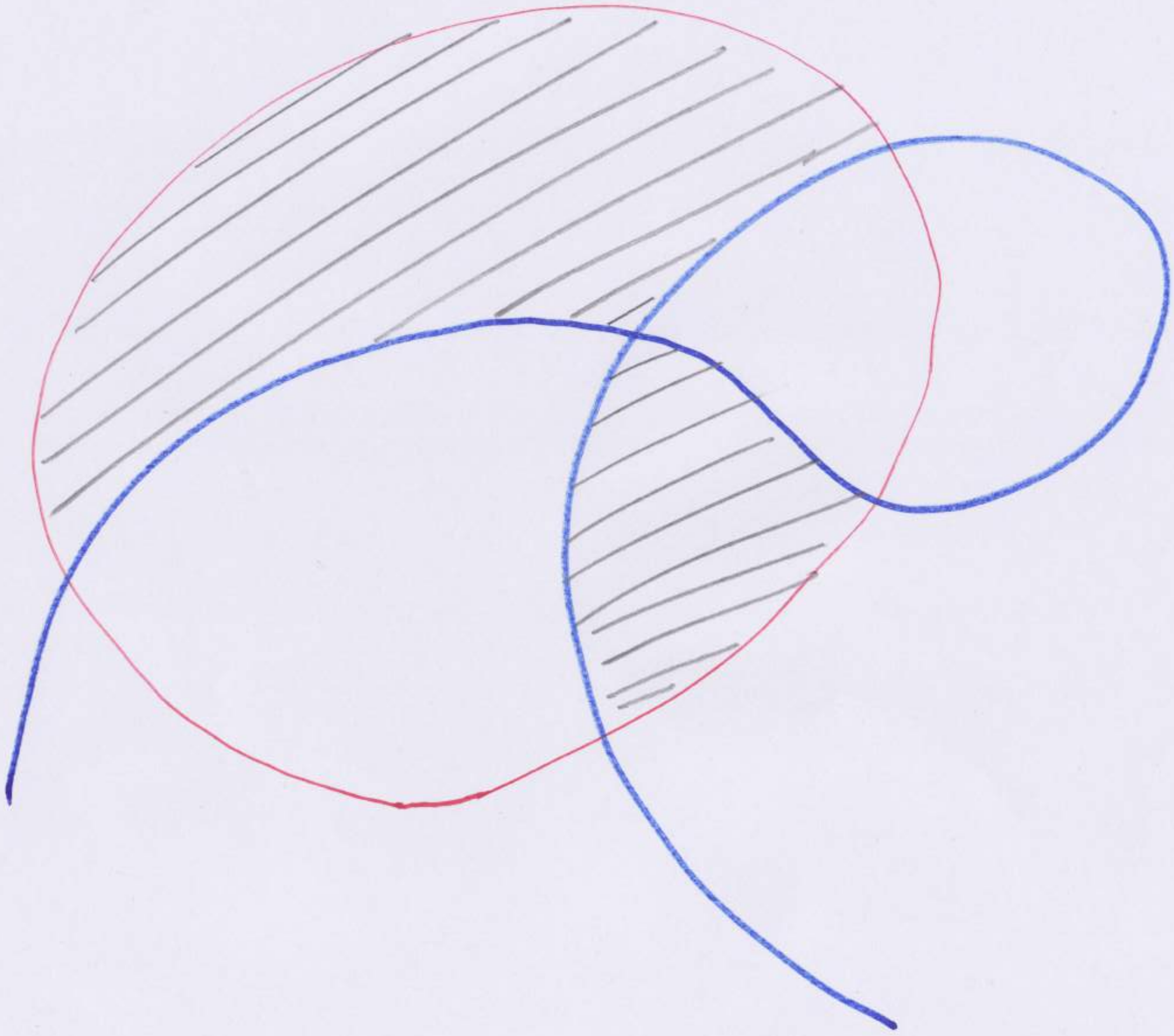


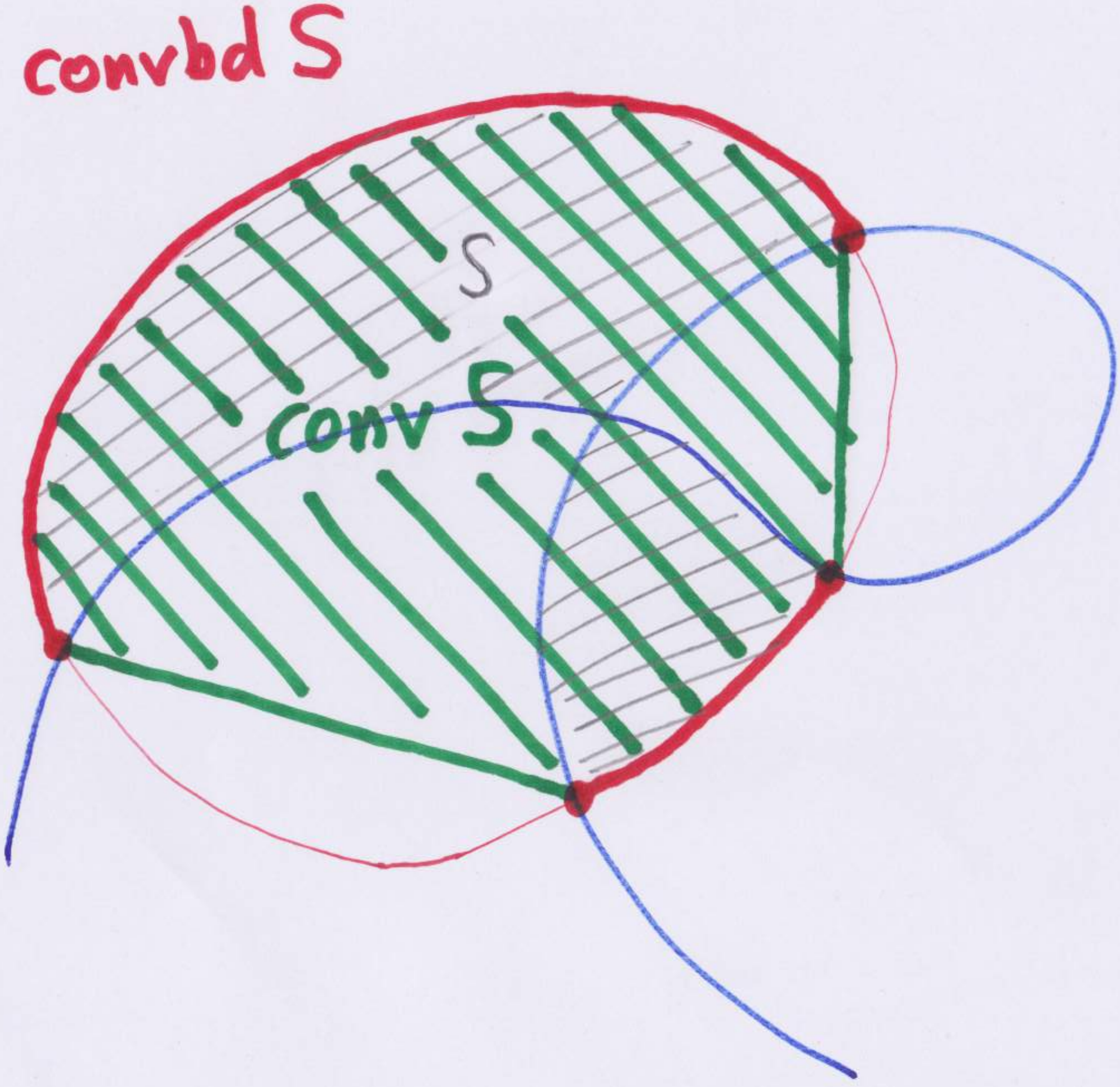
Thm (FoCM) Let $g \in \mathbb{R}[\underline{X}]^m$ such that $M(g)$ is Archimedean.

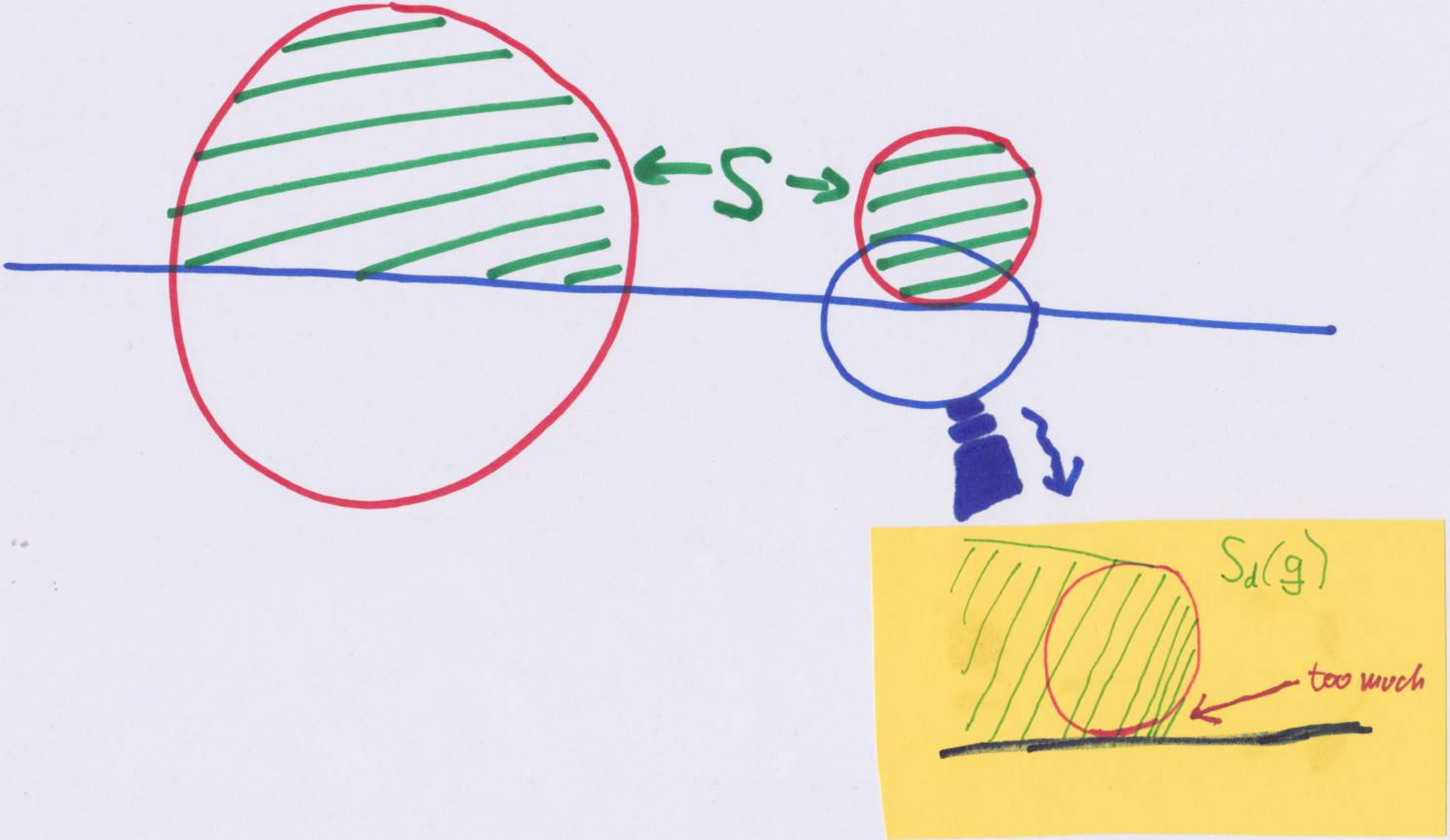
Suppose that $S(g)$ has nonempty interior near its convex boundary and each g_i is strictly quasiconcave on $(\operatorname{convbd} S) \cap Z(g_i)$.

Then $S(g) = \operatorname{conv} S(g) = S_d(\overline{g})$ for large d.

Proof uses real closed fields!




convbd S
S
conv S


S
$S_d(g)$
too much

The result for optimization

For $x \in \mathbb{R}^n$, $u_x := (X_1 - x_1)^2 + \ldots + (X_n - x_n)^2 \in \mathbb{R}[\underline{X}]$.

Thm (FoCM) Let $n, m \in \mathbb{N}_0$ and $\underline{g} \in \mathbb{R}[\underline{X}]^m$ such that $M(\underline{g})$ is Archimedean.

Suppose $S(\underline{g}) \neq \emptyset$. Moreover, let $k \in \mathbb{N}_0$, $N \in \mathbb{N}$ and $\varepsilon > 0$.

Then there exists $d \in \mathbb{N}_0$ such that for all $f \in \mathbb{R}[\underline{X}]_N$ with all coefficients in $[-N, N]$, we have :

The result for optimization

For $x \in \mathbb{R}^n$, $\quad U_x := (X_1 - x_1)^2 + \ldots + (X_n - x_n)^2 \in \mathbb{R}[\underline{X}]$.

Thm (FoCM) Let $n, m \in \mathbb{N}_0$ and $\underline{g} \in \mathbb{R}[\underline{X}]^m$ such that $M(\underline{g})$ is Archimedean. Suppose $S(\underline{g}) \neq \emptyset$. Moreover, let $k \in \mathbb{N}_0$, $N \in \mathbb{N}$ and $\varepsilon > 0$. Then there exists $d \in \mathbb{N}_0$ such that for all $f \in \mathbb{R}[\underline{X}]_N$ with all coefficients in $[-N, N]$, we have:

If $x_1, \ldots, x_k$ are the global minimizers of f on $S(\underline{g})$, if the balls of radius ε around the x_i are pairwise disjoint and contained in $S(\underline{g})$ and if we have

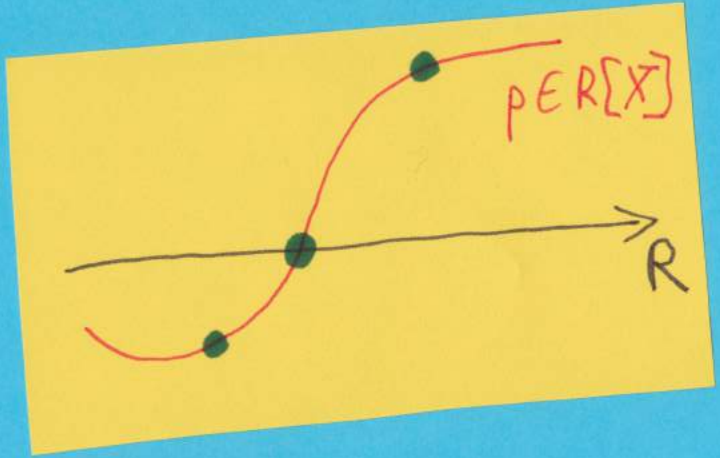
$$f \geq \operatorname{opt}(f, \underline{g}) + \varepsilon u \text{ on } S(\underline{g})$$

where $u := U_{x_1} \cdots U_{x_k} \in \mathbb{R}[\underline{X}]$, then $f - \operatorname{opt}(f, \underline{g}) \in M_d(\underline{g})$

and consequently $\operatorname{opt}(f, \underline{g}) = \operatorname{opt}_d(f, \underline{g})$.

Real closed fields

A field R is real closed if $a \leq b :\Longleftrightarrow \exists c \in R : a + c^2 = b$ $(a, b \in R)$ defines a linear order on the set R with respect to which the intermediate value theorem for polynomials holds.



Real closed fields

A field R is real closed if $a \leq b :\Longleftrightarrow \exists c \in R : a + c^2 = b$ $(a, b \in R)$ defines a linear order on the set R with respect to which the intermediate value theorem for polynomials holds.

From now on, let R be a real closed extension field of $\mathbb{R}$.

$$\mathcal{O}_R := \{a \in \mathbf{R} \mid \exists N \in \mathbb{N} : -N \leq a \leq N\}$$

„finite elements"

subring of R

$$\mathfrak{m}_R := \{a \in R \mid \forall N \in \mathbb{N} : -\frac{1}{N} \leq a \leq \frac{1}{N}\}$$

„infinitesimal elements"

unique maximal ideal of $\mathcal{O}_R$

Real closed fields

A field R is real closed if $a \leq b :\Longleftrightarrow \exists c \in R : a + c^2 = b$ $(a, b \in R)$ defines a linear order on the set R with respect to which the intermediate value theorem for polynomials holds.

From now on, let R be a real closed extension field of $\mathbb{R}$.

$$\mathcal{O}_R := \{a \in \mathbf{R} \mid \exists N \in \mathbb{N} : -N \leq a \leq N\}$$

„finite elements"

subring of R

$$\mathfrak{m}_R := \{a \in R \mid \forall N \in \mathbb{N} : -\frac{1}{N} \leq a \leq \frac{1}{N}\}$$

„infinitesimal elements"

unique maximal ideal of $\mathcal{O}_R$

„standard part"

For each $a \in \mathcal{O}_R$ there is exactly one $\mathrm{st}(a) \in \mathbb{R}$ such that $a - \mathrm{st}(a) \in \mathfrak{m}_R$.

Real closed fields

A field R is real closed if $a \leq b :\Longleftrightarrow \exists c \in R: a+c^2=b$ $(a,b \in R)$ defines a linear order on the set R with respect to which the intermediate value theorem for polynomials holds.

From now on, let R be a real closed extension field of $\mathbb{R}$.

$$\mathcal{O}_R := \{a \in \mathbf{R} \mid \exists N \in \mathbb{N}: -N \leq a \leq N\}$$

„finite elements"

subring of R

$$\mathfrak{m}_R := \{a \in R \mid \forall N \in \mathbb{N}: -\frac{1}{N} \leq a \leq \frac{1}{N}\}$$

„infinitesimal elements"

unique maximal ideal of $\mathcal{O}_R$

„standard part"

For each $a \in \mathcal{O}_R$ there is exactly one $\operatorname{st}(a) \in \mathbb{R}$ such that $a - \operatorname{st}(a) \in \mathfrak{m}_R$.

For $x \in \mathcal{O}^n$, we set $I_x := (X_1 - x_1, \dots, X_n - x_n) = \{f \in \mathcal{O}[X] \mid f(x) = 0\}$

so that $I_x^2 = \{f \in \mathcal{O}[X] \mid f(x) = 0, \nabla f(x) = 0\}$.

A key lemma

Let M be an Archimedean quadratic module of $\mathcal{O}[\bar{X}]$ and set

$$S := \{x \in \mathbb{R}^n \mid \forall p \in M:\ \operatorname{st}(p(x)) \geq 0\}.$$

Moreover, suppose $k \in \mathbb{N}_0$ and let $x_1, \dots, x_k \in \mathcal{O}^n$ have pairwise distinct standard parts. Let

$$f \in \bigcap_{i=1}^{k} I_{x_i}^2$$

such that

$$\underline{\operatorname{st}(f(x)) > 0}$$

for all $x \in S \setminus \{\operatorname{st}(x_1), \dots, \operatorname{st}(x_n)\}$ and

$$\underline{\operatorname{st}(v^T (\operatorname{Hess} f)(x_i)\, v) > 0}$$

for all $i \in \{1, \dots, k\}$ and $v \in \mathbb{R}^n \setminus \{0\}$. **Then $f \in M$.**

A key Lemma

Let M be an Archimedean quadratic module of $\mathcal{O}[\bar{X}]$ and set

$$S := \{x \in \mathbb{R}^n \mid \forall p \in M:\ \operatorname{st}(p(x)) \geq 0\}.$$

Moreover, suppose $k \in \mathbb{N}_0$ and let $x_1, \dots, x_k \in \mathcal{O}^n$ have pairwise distinct standard parts. Let

$$f \in \bigcap_{i=1}^{k} I_{x_i}^2$$

such that

$$\underline{\operatorname{st}(f(x)) > 0}$$

for all $x \in S \setminus \{\operatorname{st}(x_1), \dots, \operatorname{st}(x_n)\}$ and

$$\underline{\operatorname{st}(v^T (\operatorname{Hess} f)(x_i)\, v) > 0}$$

for all $i \in \{1, \dots, k\}$ and $v \in \mathbb{R}^n \setminus \{0\}$. **Then $f \in M$.**

Proof idea. Assume $f \notin M$. Separate f from the cone $M \cap I$ in the real vector space $I := I_{x_1}^2 \cdots I_{x_k}^2 \overset{\text{Chinese}}{\underset{\text{remainder}}{=}} I_{x_1}^2 \cap \dots \cap I_{x_k}^2$.

Choose extremal separating functional (real valued!) . . .

$\mathcal{O}_R^n$

line $\ell(x)=0$

$\ell \in \mathcal{O}_R[\underline{X}]$

Lagrange multipliers from $\mathcal{O}_{\geq 0} \subseteq \mathbb{R}_{\geq 0}$