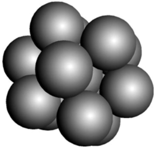
Flyspeck Inequalities and Semidefinite Programming

Victor Magron, RA Imperial College

Memory Optimization and Co-Design Meeting
29 June 2015



Errors and Proofs

- Mathematicians and Computer Scientists want to eliminate all the uncertainties on their results. Why?

Errors and Proofs

- Mathematicians and Computer Scientists want to eliminate all the uncertainties on their results. Why?

M. Lecat, Erreurs des Mathématiciens des origines à nos jours, 1935.

$\rightsquigarrow$ 130 pages of errors! (Euler, Fermat, Sylvester, . . .)

Errors and Proofs

- Mathematicians and Computer Scientists want to eliminate all the uncertainties on their results. Why?

M. Lecat, Erreurs des Mathématiciens des origines à nos jours, 1935.

$\leadsto$ 130 pages of errors! (Euler, Fermat, Sylvester, ...)

Ariane 5 launch failure, Pentium FDIV bug

Errors and Proofs

- Possible workaround: proof assistants

  Coq (Coquand, Huet 1984)

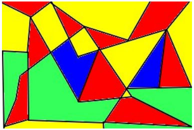
  HOL-LIGHT (Harrison, Gordon 1980)

  Built in top of OCAML

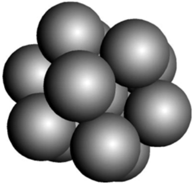
Complex Proofs

- Complex mathematical proofs / mandatory computation

K. Appel and W. Haken , Every Planar Map is Four-Colorable, 1989.



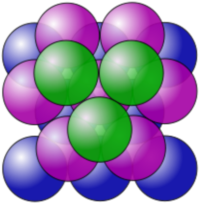
T. Hales, A Proof of the Kepler Conjecture, 1994.



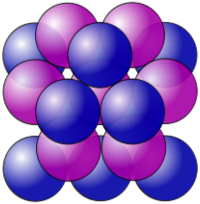
From Oranges Stack...

Kepler Conjecture (1611):

The maximal density of sphere packings in 3D-space is $\frac{\pi}{\sqrt{18}}$



Face-centered cubic Packing



Hexagonal Compact Packing

...to Flyspeck Nonlinear Inequalities

- The proof of T. Hales (1998) contains mathematical and computational parts
- Computation: check thousands of nonlinear inequalities
- Robert MacPherson, editor of The Annals of Mathematics: "[...] the mathematical community will have to get used to this state of affairs."
- **Flyspeck** [Hales 06]: **F**ormal **P**roof of **K**epler Conjecture

...to Flyspeck Nonlinear Inequalities

- The proof of T. Hales (1998) contains mathematical and computational parts
- Computation: check thousands of nonlinear inequalities
- Robert MacPherson, editor of The Annals of Mathematics: "[...] the mathematical community will have to get used to this state of affairs."
- **Flyspeck** [Hales 06]: **F**ormal **P**roof of **K**epler Conjecture
- **Project Completion on 10 August by the Flyspeck team!!**

...to Flyspeck Nonlinear Inequalities

- Nonlinear inequalities: quantified reasoning with "$\forall$"

$$\forall \mathbf{x} \in \mathbf{K}, f(\mathbf{x}) \geqslant 0$$

- NP-hard optimization problem

A "Simple" Example

In the computational part:

- Multivariate Polynomials:
  $\Delta\mathbf{x} := x_1x_4(-x_1 + x_2 + x_3 - x_4 + x_5 + x_6) + x_2x_5(x_1 - x_2 + x_3 + x_4 - x_5 + x_6) + x_3x_6(x_1 + x_2 - x_3 + x_4 + x_5 - x_6) - x_2(x_3x_4 + x_1x_6) - x_5(x_1x_3 + x_4x_6)$

A "Simple" Example

In the computational part:

- Semialgebraic functions: composition of polynomials with $|\cdot|, \sqrt{}, +, -, \times, /, \sup, \inf, \ldots$

$$p(\mathbf{x}) := \partial_4 \Delta \mathbf{x} \qquad q(\mathbf{x}) := 4 x_1 \Delta \mathbf{x}$$
$$r(\mathbf{x}) := p(\mathbf{x}) / \sqrt{q(\mathbf{x})}$$

$$l(\mathbf{x}) := -\frac{\pi}{2} + 1.6294 - 0.2213 \left(\sqrt{x_2} + \sqrt{x_3} + \sqrt{x_5} + \sqrt{x_6} - 8.0\right) + 0.913 \left(\sqrt{x_4} - 2.52\right) + 0.728 \left(\sqrt{x_1} - 2.0\right)$$

A “Simple” Example

In the computational part:

- Transcendental functions $\mathcal{T}$: composition of semialgebraic functions with arctan, exp, sin, $+, -, \times, \dots$

A “Simple” Example

In the computational part:

- Feasible set $\mathbf{K} := [4, 6.3504]^3 \times [6.3504, 8] \times [4, 6.3504]^2$

Lemma$_{9922699028}$ from Flyspeck:

$$\forall \mathbf{x} \in \mathbf{K}, \arctan\left(\frac{p(\mathbf{x})}{\sqrt{q(\mathbf{x})}}\right) + l(\mathbf{x}) \geqslant 0$$

Existing Formal Frameworks

Formal proofs for Global Optimization:

- Bernstein polynomial methods [Zumkeller's PhD 08]
- SMT methods [Gao et al. 12]
- Interval analysis and Sums of squares

Existing Formal Frameworks

Interval analysis

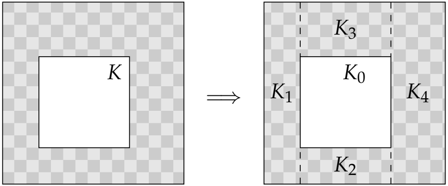
- Certified interval arithmetic in Coq [Melquiond 12]
- Taylor methods in HOL Light [Solovyev thesis 13]
  - Formal verification of floating-point operations
- robust but subject to the **Curse of Dimensionality**

Existing Formal Frameworks

Lemma$_{9922699028}$ from Flyspeck:

$$\forall \mathbf{x} \in \mathbf{K}, \arctan\left(\frac{\partial_4 \Delta \mathbf{x}}{\sqrt{4 x_1 \Delta \mathbf{x}}}\right) + l(\mathbf{x}) \geqslant 0$$

- Dependency issue using Interval Calculus:
  - One can bound $\partial_4 \Delta \mathbf{x} / \sqrt{4 x_1 \Delta \mathbf{x}}$ and $l(\mathbf{x})$ separately
  - Too coarse lower bound: -0.87
  - Subdivide $\mathbf{K}$ to prove the inequality



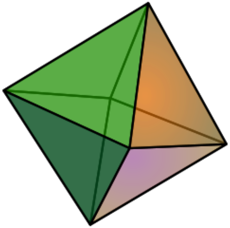
Introduction

Flyspeck Inequalities and Semidefinite Programming

Semidefinite Programming

- Linear Programming (LP):

$$\min_{\mathbf{z}} \quad \mathbf{c}^\top \mathbf{z}$$
$$\text{s.t.} \quad \mathbf{A}\,\mathbf{z} \geqslant \mathbf{d} \ .$$



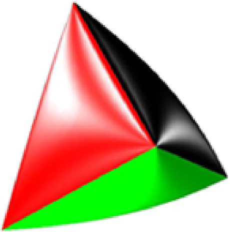
  - Linear cost $\mathbf{c}$
  - Linear inequalities "$\sum_i A_{ij}\, z_j \geqslant d_i$"

Polyhedron

Semidefinite Programming

- Semidefinite Programming (SDP):

$$\min_{\mathbf{z}} \quad \mathbf{c}^\top \mathbf{z}$$
$$\text{s.t.} \quad \sum_i \mathbf{F}_i z_i \succcurlyeq \mathbf{F}_0 \ .$$



Spectrahedron

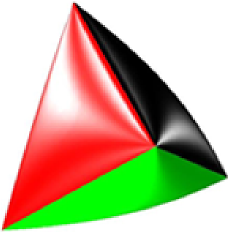
- Linear cost $\mathbf{c}$
- Symmetric matrices $\mathbf{F}_0, \mathbf{F}_i$
- Linear matrix inequalities "$\mathbf{F} \succcurlyeq 0$" ($\mathbf{F}$ has nonnegative eigenvalues)

Semidefinite Programming

- Semidefinite Programming (SDP):

$$\min_{\mathbf{z}} \quad \mathbf{c}^\top \mathbf{z}$$
$$\text{s.t.} \quad \sum_i \mathbf{F}_i z_i \succcurlyeq \mathbf{F}_0 \ , \quad \mathbf{A}\,\mathbf{z} = \mathbf{d} \ .$$

  - Linear cost $\mathbf{c}$
  - Symmetric matrices $\mathbf{F}_0, \mathbf{F}_i$
  - Linear matrix inequalities "$\mathbf{F} \succcurlyeq 0$" ($\mathbf{F}$ has nonnegative eigenvalues)



Spectrahedron

SDP for Polynomial Optimization

- Prove **polynomial inequalities** with SDP:

$$p(a,b) := a^2 - 2ab + b^2 \geqslant 0 \ .$$

- Find $\mathbf{z}$ s.t. $p(a,b) = \begin{pmatrix} a & b \end{pmatrix} \underbrace{\begin{pmatrix} z_1 & z_2 \\ z_2 & z_3 \end{pmatrix}}_{\succcurlyeq 0} \begin{pmatrix} a \\ b \end{pmatrix}$.

- Find $\mathbf{z}$ s.t. $a^2 - 2ab + b^2 = z_1 a^2 + 2 z_2 ab + z_3 b^2$ $\qquad (\mathbf{A}\,\mathbf{z} = \mathbf{d})$

- $\begin{pmatrix} z_1 & z_2 \\ z_2 & z_3 \end{pmatrix} = \underbrace{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}}_{\mathbf{F}_1} z_1 + \underbrace{\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}}_{\mathbf{F}_2} z_2 + \underbrace{\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}}_{\mathbf{F}_3} z_3 \succcurlyeq \underbrace{\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}}_{\mathbf{F}_0}$

SDP for Polynomial Optimization

- Choose a cost $\mathbf{c}$ e.g. $(1, 0, 1)$ and solve:

$$\min_{\mathbf{z}} \quad \mathbf{c}^\top \mathbf{z}$$
$$\text{s.t.} \quad \sum_i \mathbf{F}_i z_i \succcurlyeq \mathbf{F}_0 \;, \quad \mathbf{A}\,\mathbf{z} = \mathbf{d} \;.$$

- Solution $\begin{pmatrix} z_1 & z_2 \\ z_2 & z_3 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \succcurlyeq 0$ (eigenvalues 0 and 1)

- $a^2 - 2ab + b^2 = \begin{pmatrix} a & b \end{pmatrix} \underbrace{\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}}_{\succcurlyeq 0} \begin{pmatrix} a \\ b \end{pmatrix} = (a - b)^2 \,.$

- Solving **SDP** $\implies$ Finding **SUMS OF SQUARES** certificates

Polynomial Optimization 

Semidefinite Programming

$$\begin{pmatrix} 1 & a & b \\ a & 1 & c \\ b & c & 1 \end{pmatrix} \succcurlyeq 0$$

$\rightsquigarrow$ control, polynomial optim (Henrion, Lasserre, Parrilo)

$\rightsquigarrow$ combinatorial optim. electrical engineering(Laurent, Steurers)

Polynomial Optimization 

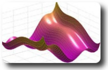
Semidefinite Programming

$$\begin{pmatrix} 1 & a & b \\ a & 1 & c \\ b & c & 1 \end{pmatrix} \succcurlyeq 0$$

$\rightsquigarrow$ control, polynomial optim (Henrion, Lasserre, Parrilo)

$\rightsquigarrow$ combinatorial optim. electrical engineering(Laurent, Steurers)

Theoretical Approach



$$p^* := \inf_{\mathbb{R}^n} p(\mathbf{x}) \ ?$$

$$\text{sup} \quad \lambda$$

$$\Leftarrow \text{with} \quad p - \lambda \geqslant 0$$

INFINITE LP

Polynomial Optimization

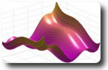
Semidefinite Programming

$$\begin{pmatrix} 1 & a & b \\ a & 1 & c \\ b & c & 1 \end{pmatrix} \succcurlyeq 0$$

$\rightsquigarrow$ control, polynomial optim (Henrion, Lasserre, Parrilo)

$\rightsquigarrow$ combinatorial optim. electrical engineering(Laurent, Steurers)

Practical Approach



$$p^* := \inf_{\mathbb{R}^n} p(\mathbf{x}) \ ?$$

$$\Leftarrow \begin{array}{ll} \sup & \lambda \\ \text{with} & p - \lambda = \textbf{sums of squares of fixed degree} \end{array}$$

FINITE SDP

Polynomial Optimization 

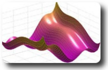
Semidefinite Programming

$$\begin{pmatrix} 1 & a & b \\ a & 1 & c \\ b & c & 1 \end{pmatrix} \succcurlyeq 0$$

$\rightsquigarrow$ control, polynomial optim (Henrion, Lasserre, Parrilo)

$\rightsquigarrow$ combinatorial optim. electrical engineering(Laurent, Steurers)

Practical Approach



$p^* := \inf_{\mathbb{R}^n} p(\mathbf{x})$?

$$\sup \quad \lambda$$

$\Leftarrow$ with $\quad p - \lambda =$ **sums of squares of fixed degree**

FINITE SDP

SDP bounds Hierarchy $\uparrow p^*$

degree d

n variables $\Rightarrow \binom{n+2d}{n}$ variables **SDP**

Polynomial Optimization

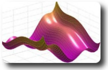
Semidefinite Programming

$$\begin{pmatrix} 1 & a & b \\ a & 1 & c \\ b & c & 1 \end{pmatrix} \succcurlyeq 0$$

$\rightsquigarrow$ control, polynomial optim (Henrion, Lasserre, Parrilo)

$\rightsquigarrow$ combinatorial optim. electrical engineering(Laurent, Steurers)

Practical Approach



$p^* := \inf_{\mathbb{R}^n} p(\mathbf{x})$?

$\sup \quad \lambda$

$\Leftarrow$ with $\quad p - \lambda =$ **sums of squares of fixed degree**

FINITE SDP

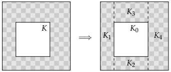
SDP bounds Hierarchy $\uparrow p^*$

degree d
n variables $\quad \Longrightarrow \binom{n+2d}{n}$ variables **SDP**

- **Strengthening** $\quad p - \lambda =$ **sums of squares** $\Longrightarrow p \geqslant \lambda$
- $1 + x_1^4 - 2x_1^2x_2^2 + x_2^4 = 1 + (x_1^2 - x_2^2)^2$

Non-polynomial Optimization

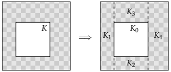
TAYLOR + INTERVALS : ⊕ scalable ⊖ coarse



⇝ **Curse of dimensionality**

Non-polynomial Optimization

TAYLOR + INTERVALS : $\oplus$ scalable $\ominus$ coarse



$\rightsquigarrow$ **Curse of dimensionality**

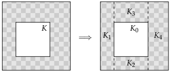
TAYLOR + SUMS OF SQUARES : $\ominus$ not scalable $\oplus$ precise

high degree d
n variables $\Longrightarrow \binom{n+2d}{n}$

$\rightsquigarrow$ **No free lunch**

Non-polynomial Optimization

TAYLOR + INTERVALS : $\oplus$ scalable $\ominus$ coarse



$\rightsquigarrow$ **Curse of dimensionality**

TAYLOR + SUMS OF SQUARES : $\ominus$ not scalable $\oplus$ precise

high degree d
n variables $\Longrightarrow \binom{n+2d}{n}$

$\rightsquigarrow$ **No free lunch**

MAXPLUS + SUMS OF SQUARES : $\oplus$ scalable $\oplus$ precise

Maxplus in control (Akian Gaubert)

$\Updownarrow$

Templates in static analysis (Manna)

$\rightsquigarrow$ **Curse reduction**

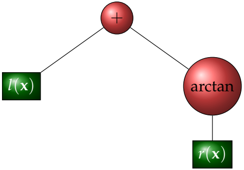
Maxplus Approximations

Approximate $f : \mathbb{R}^n \to \mathbb{R}$ with supremum of quadratic forms.

Non-polynomial Optimization

MAXPLUS + SUMS OF SQUARES: ⊕ scalable ⊕ precise

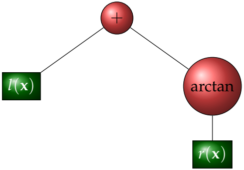
Function from "simple" inequality:

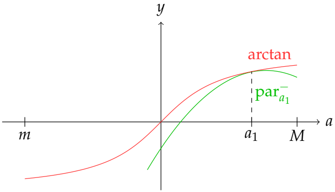


Non-polynomial Optimization

MAXPLUS + SUMS OF SQUARES: ⊕ scalable ⊕ precise

Verification software `NLCertify`, 1st iteration:





1 control point $\{a_1\}$

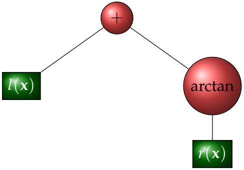
$m_1 = -4.7 \times 10^{-3}$

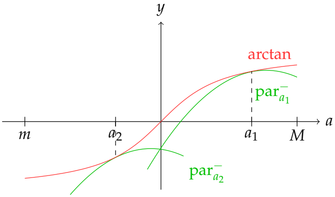
< 0

Non-polynomial Optimization

MAXPLUS + SUMS OF SQUARES: ⊕ scalable ⊕ precise

Verification software NLCertify, 2nd iteration:





2 control points $\{a_1, a_2\}$

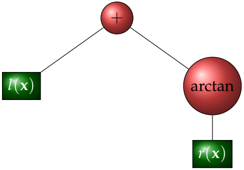
$m_2 = -6.1 \times 10^{-5}$

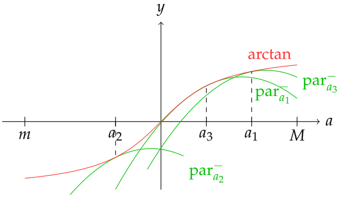
< 0

Non-polynomial Optimization

MAXPLUS + SUMS OF SQUARES: ⊕ scalable ⊕ precise

Verification software NLCertify, 3rd iteration:





3 control points $\{a_1, a_2, a_3\}$

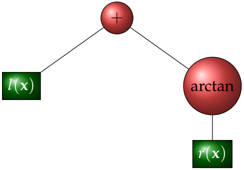
$m_3 = 4.1 \times 10^{-6}$

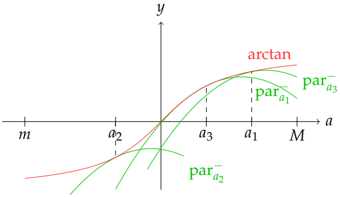
$> \mathbf{0}$

Non-polynomial Optimization

MAXPLUS + SUMS OF SQUARES: $\oplus$ scalable $\oplus$ precise

Verification software `NLCertify`, 3rd iteration:





3 control points $\{a_1, a_2, a_3\}$

$m_3 = 4.1 \times 10^{-6}$

$> \mathbf{0}$

Theorem

The algorithm **converges** to a global optimum and **certifies** inequalities.

$\mathbf{n_{Hales}}$: time ratio between **formal** and **numerical** certification (V. Vœvodsky)

$\leadsto \mathbf{n_{Hales}} \lesssim 10$ (Maxplus + Sums of squares) $\ll 2000$ (Taylor + Intervals)

Contributions

CERTIFICATION MAXPLUS–SUMS OF SQUARES: **NUMERIC** OR **FORMAL**

Journals

Magron, Allamigeon, Gaubert & Werner, ***Journal*** *Math. Prog. Ser. B* 2014

Magron, Allamigeon, Gaubert & Werner, ***Journal*** *of Formalized Reasoning* 2015

Conferences

Allamigeon, Gaubert, Magron & Werner, *Calculemus* ***Conference*** 2013

Allamigeon, Gaubert, Magron & Werner, *European Control* ***Conference*** 2013

Magron, *ICMS* ***Conference*** 2014 **+ software** `NLCertify`

A FORMAL PROOF OF KEPLER CONJECTURE

Hales, Adams, Bauer, Dang, Harrison, Hoang, Kaliszyk, Magron, Mclaughlin, Nguyen, Nguyen, Nipkow, Obua, Pleso, Rute, Solovyev, Ta, Tran, Trieu, Urban, Vu & Zumkeller, *Prepublication, submitted Sigma/Pi* ***Journal*** 2015

End

Thank you for your attention!

cas.ee.ic.ac.uk/people/vmagron