

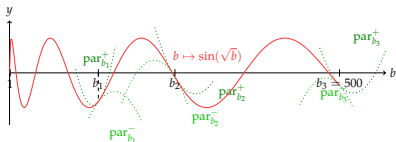
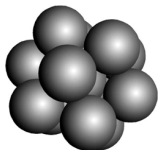
NLCertify: A Tool for Formal Nonlinear Optimization

Victor Magron, Postdoc LAAS-CNRS

6 August 2014

ICMS 2014

Hanyang University, Seoul



Errors and Proofs

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



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

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

Errors and Proofs

- Possible workaround: proof assistants

COQ (Coquand, Huet 1984) 🐣

HOL-LIGHT (Harrison, Gordon 1980)



Built in top of OCAML 🐪

- Tool: Formal Bounds for Global Optimization

- Collaboration with:



Benjamin Werner (LIX Polytechnique)




Stéphane Gaubert (Maxplus Team CMAP/INRIA
Polytechnique)

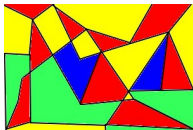


Xavier Allamigeon (Maxplus Team)

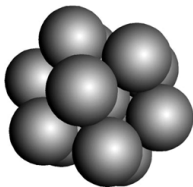
Complex Proofs

- Complex mathematical proofs / mandatory computation

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



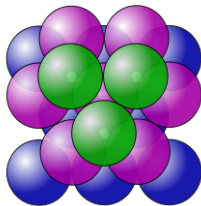
 T. Hales, A Proof of the Kepler Conjecture, 1994.



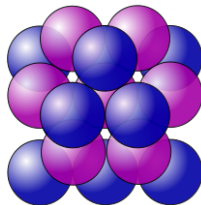
From Oranges Stack...

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]: **Formal Proof of Kepler Conjecture**

...to FLYSPECK Nonlinear Inequalities

- Nonlinear inequalities: quantified reasoning with “ \forall ”

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

- NP-hard optimization problem

A “Simple” Example

In the computational part:

- Multivariate Polynomials:

$$\Delta \mathbf{x} := x_1 x_4 (-x_1 + x_2 + x_3 - x_4 + x_5 + x_6) + x_2 x_5 (x_1 - x_2 + x_3 + x_4 - x_5 + x_6) + x_3 x_6 (x_1 + x_2 - x_3 + x_4 + x_5 - x_6) - x_2 (x_3 x_4 + x_1 x_6) - x_5 (x_1 x_3 + x_4 x_6)$$

A “Simple” Example

In the computational part:

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

$$p(\mathbf{x}) := \partial_4 \Delta \mathbf{x} \quad q(\mathbf{x}) := 4x_1 \Delta \mathbf{x}$$

$$r(\mathbf{x}) := p(\mathbf{x}) / \sqrt{q(\mathbf{x})}$$

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

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

Lemma₉₉₂₂₆₉₉₀₂₈ from Flyspeck:

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

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

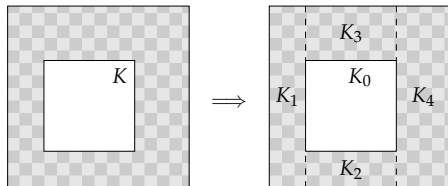
- Certified interval arithmetic in COQ [Melquiond 12]
- Taylor methods in HOL Light [Solovyev thesis 13]
 - Formal verification of floating-point operations
- robust but subject to the **Curse of Dimensionality**

Existing Formal Frameworks

Lemma₉₉₂₂₆₉₉₀₂₈ from Flyspeck:

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

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



Existing Formal Frameworks

Sums of squares techniques

- Formalized in HOL-LIGHT [Harrison 07] COQ [Besson 07]
 - Precise methods but scalability and robustness issues (numerical)
 - powerful: global optimality certificates without branching
- but
- not so robust: handles moderate size problems
 - Restricted to polynomials

Existing Formal Frameworks

Approximation theory: Chebyshev/Taylor models

- mandatory for non-polynomial problems
- hard to combine with SOS techniques (degree of approximation)

Existing Formal Frameworks

Can we develop a new approach with both keeping the respective strength of interval and precision of SOS?

Proving Flyspeck Inequalities is challenging: medium-size and tight

New Framework (in my PhD thesis)

- Certificates for lower bounds of Global Optimization Problems using SOS and new ingredients in Global Optimization:
 - Maxplus approximation (Optimal Control)
 - Nonlinear templates (Static Analysis)
- Verification of these certificates inside COQ

New Framework (in my PhD thesis)

Software Implementation NLCertify:

- <https://forge.ocamlcore.org/projects/nl-certify/>



15 000 lines of OCAML code



4000 lines of COQ code

Introduction

Moment-SOS relaxations and Maxplus approximation

Formal Nonlinear Optimization

Moment-SOS relaxations

- Semialgebraic set $\mathbf{K} := \{\mathbf{x} \in \mathbb{R}^n : g_1(\mathbf{x}) \geq 0, \dots, g_m(\mathbf{x}) \geq 0\}$
- $p^* := \min_{\mathbf{x} \in \mathbf{K}} p(\mathbf{x})$: NP hard
- Sums of squares $\Sigma[\mathbf{x}]$
- $\mathcal{Q}(\mathbf{K}) := \left\{ \sigma_0(\mathbf{x}) + \sum_{j=1}^m \sigma_j(\mathbf{x})g_j(\mathbf{x}), \text{ with } \sigma_j \in \Sigma[\mathbf{x}] \right\}$

Moment-SOS relaxations

- $\mathcal{M}_+(\mathbf{K})$: space of probability measures supported on \mathbf{K}

Polynomial Optimization Problems (POP)

$$\begin{array}{ll} \text{(Primal)} & \text{(Dual)} \\ \inf \int_{\mathbf{K}} p d\mu & = \sup \lambda \\ \text{s.t. } \mu \in \mathcal{M}_+(\mathbf{K}) & \text{s.t. } \lambda \in \mathbb{R}, \\ & p - \lambda \in \mathcal{Q}(\mathbf{K}) \end{array}$$

Moment-SOS relaxations

- Truncated quadratic module $\mathcal{Q}_k(\mathbf{K}) := \mathcal{Q}(\mathbf{K}) \cap \mathbb{R}_{2k}[\mathbf{x}]$

Polynomial Optimization Problems (POP)

(Moment)		(SOS)
$\inf \int_{\mathbf{K}} p d\mu$	\geq	$\sup \lambda$
s.t. $\mu \in \mathcal{M}_+(\mathbf{K})$		s.t. $\lambda \in \mathbb{R}$, $p - \lambda \in \mathcal{Q}_k(\mathbf{K})$

Practical Computation

- Hierarchy of SOS relaxations:

$$\lambda_k := \sup_{\lambda} \left\{ \lambda : p - \lambda \in \mathcal{Q}_k(\mathbf{K}) \right\}$$

- Convergence guarantees $\lambda_k \uparrow p^*$ [Lasserre 01]
- Can be computed with semidefinite programming (SDP) solvers (CSDP, SDPA)
- **No free lunch:** $\binom{n+2k}{n}$ SDP variables
- Extension to semialgebraic functions $r(\mathbf{x}) = p(\mathbf{x}) / \sqrt{q(\mathbf{x})}$ [Lasserre-Putinar 10]

Practical Computation

- *Caprasse* Problem

$$\forall \mathbf{x} \in [-0.5, 0.5]^4, -x_1x_3^3 + 4x_2x_3^2x_4 + 4x_1x_3x_4^2 + 2x_2x_4^3 + 4x_1x_3 + 4x_3^2 - 10x_2x_4 - 10x_4^2 + 5.1801 \geq 0.$$

- `scale_pol = true`: scaled on $[0, 1]^4$

- `relax_order = 2`: SOS of degree at most 4

- `bound_squares_variables = true`:
rsedundant constraints $x_1^2 \leq 1, \dots, x_4^2 \leq 1$

The General “Informal Framework”

Given \mathbf{K} a compact set and f a **transcendental** function, bound $f^* = \inf_{\mathbf{x} \in \mathbf{K}} f(\mathbf{x})$ and prove $f^* \geq 0$

- f is underestimated by a **semialgebraic** function f_{sa}
- Reduce the problem $f_{\text{sa}}^* := \inf_{\mathbf{x} \in \mathbf{K}} f_{\text{sa}}(\mathbf{x})$ to a **polynomial optimization problem (POP)**

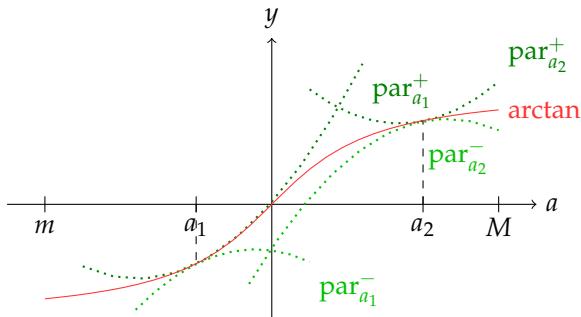
Maxplus Approximation

- Initially introduced to solve Optimal Control Problems [Fleming-McEneaney 00]
- **Curse of dimensionality** reduction [McEaneny Kluberg, Gaubert-McEneaney-Qu 11, Qu 13].
Allowed to solve instances of dim up to 15 (inaccessible by grid methods)
- In our context: approximate **transcendental** functions

Maxplus Approximation

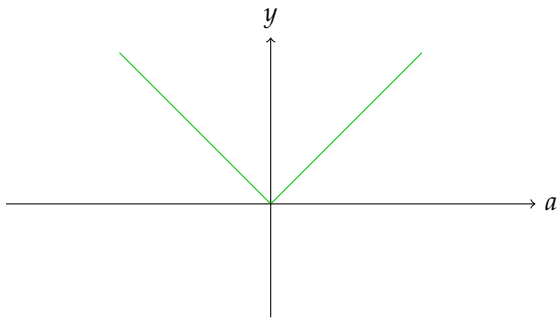
Definition

Let $\gamma \geq 0$. A function $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$ is said to be γ -semiconvex if the function $\mathbf{x} \mapsto \phi(\mathbf{x}) + \frac{\gamma}{2} \|\mathbf{x}\|_2^2$ is convex.



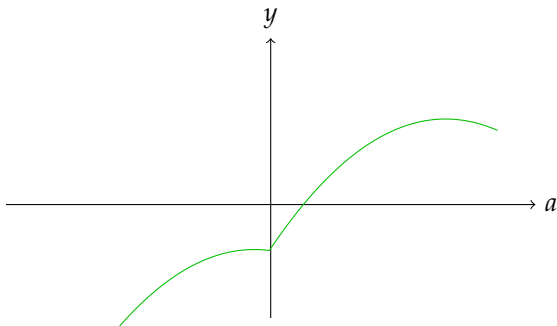
Nonlinear Function Representation

Exact parsimonious maxplus representations



Nonlinear Function Representation

Exact parsimonious maxplus representations



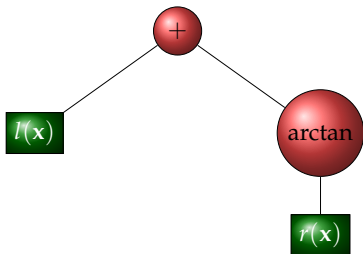
Nonlinear Function Representation

Abstract syntax tree representations of multivariate transcendental functions:

- leaves are **semialgebraic** functions of \mathcal{A}
- nodes are univariate functions of \mathcal{D} or binary operations

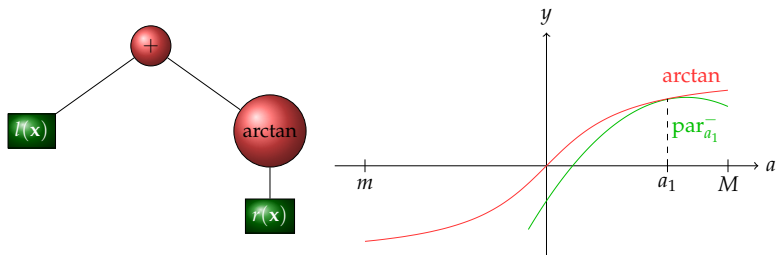
Nonlinear Function Representation

- For the “Simple” Example from Flyspeck:



Maxplus Optimization Algorithm

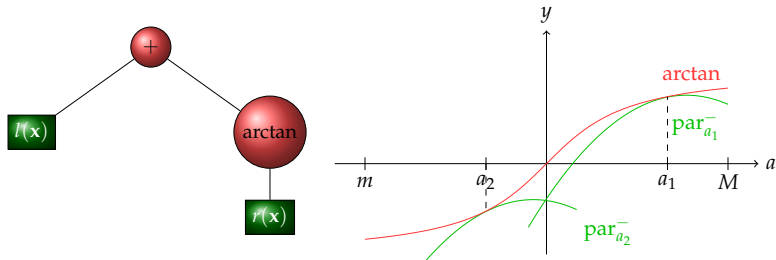
First iteration:



- 1 control point $\{a_1\}$ SOS Computation: $m_1 = -0.0047 < 0$

Maxplus Optimization Algorithm

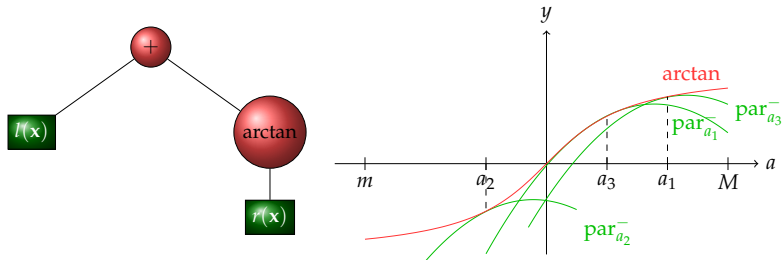
Second iteration:



- 2 2 control points $\{a_1, a_2\}$: $m_2 = -0.000061 < 0$

Maxplus Optimization Algorithm

Third iteration:



3 3 control points $\{a_1, a_2, a_3\}$: $m_3 = 0.0000046 \geq 0$

Maxplus Optimization Algorithm

Input: tree t , box \mathbf{K} , SOS relaxation order k , precision p

Output: bounds m and M , approximations t_2^- and t_2^+

- 1: **if** $t \in \mathcal{A}$ **then** $t^- := t, t^+ := t$
- 2: **else if** $u := \text{root}(t) \in \mathcal{D}$ **with child** c **then**
- 3: $m_c, M_c, c^-, c^+ := \text{samp_approx}(c, \mathbf{K}, k, p)$
- 4: $I := [m_c, M_c]$
- 5: $u^-, u^+ := \text{unary_approx}(u, I, c, p)$
- 6: $t^-, t^+ := \text{compose_approx}(u, u^-, u^+, I, c^-, c^+)$
- 7: **else if** $\text{bop} := \text{root}(t)$ **with children** c_1 **and** c_2 **then**
- 8: $m_i, M_i, c_i^-, c_i^+ := \text{samp_approx}(c_i, \mathbf{K}, k, p)$ **for** $i \in \{1, 2\}$
- 9: $t^-, t^+ := \text{compose_bop}(c_1^-, c_1^+, c_2^-, c_2^+, \text{bop}, [m_2, M_2])$
- 10: **end**
- 11: **return** $\text{min_sa}(t^-, \mathbf{K}, k), \text{max_sa}(t^+, \mathbf{K}, k), t^-, t^+$

Contributions

Published:



X. Allamigeon, S. Gaubert, V. Magron, and B. Werner.
Certification of inequalities involving transcendental functions:
combining sdp and max-plus approximation, *ECC Conference*
2013.



X. Allamigeon, S. Gaubert, V. Magron, and B. Werner.
Certification of bounds of non-linear functions: the templates
method, *CICM Conference*, 2013.

In revision:






X. Allamigeon, S. Gaubert, V. Magron, and B. Werner.
Certification of Real Inequalities – Templates and Sums of
Squares, arxiv:1403.5899, 2014.

Introduction

Moment-SOS relaxations and Maxplus approximation

Formal Nonlinear Optimization

The General “Formal Framework”

-  We check the correctness of SOS certificates for **POP**
-  We build certificates to prove interval bounds for **semialgebraic** functions
-  We bound formally **transcendental** functions with semialgebraic approximations

Formal SOS bounds

When $q \in \mathcal{Q}(\mathbf{K})$, $\sigma_0, \dots, \sigma_m$ is a positivity certificate for q

Check **symbolic polynomial equalities** $q = q'$ in COQ



Existing tactic `ring` [Grégoire-Mahboubi 05]



Polynomials coefficients: arbitrary-size rationals `bigQ`
[Grégoire-Théry 06]



Much simpler to verify certificates using *sceptical approach*



Extends also to **semialgebraic** functions

Bounding the Polynomial Remainder

- Normalized POP ($\mathbf{x} \in [0, 1]^n$)
- $\epsilon_{\text{pop}}(\mathbf{x}) := p(\mathbf{x}) - \lambda_k - \sum_{j=0}^m \sigma_j(\mathbf{x})g_j(\mathbf{x})$
- $\forall \mathbf{x} \in [0, 1]^n, \epsilon_{\text{pop}}(\mathbf{x}) \geq \epsilon_{\text{pop}}^* := \sum_{\epsilon_{\alpha} \leq 0} \epsilon_{\alpha}$

Formal SOS Results

- *POP1*: $\forall \mathbf{x} \in \mathbf{K}, \partial_4 \Delta \mathbf{x} \geq -41$.
- *POP2*: $\forall \mathbf{x} \in \mathbf{K}, \Delta \mathbf{x} \geq 0$.



Problem	n	NLCertify	micromega [Besson 07]
<i>POP1</i>	6	0.08 s	9.00 s
<i>POP2</i>	2	0.09 s	0.36 s
	3	0.39 s	—
	6	13.2 s	—



Sparse SOS relaxations \implies Speedup



Benchmarks for Flyspeck Inequalities

Inequality	#boxes	 Time	 Time
9922699028	39	190 s	2218 s
3318775219	338	1560 s	19136 s

- Comparable with Taylor interval methods in HOL-LIGHT [Hales-Solovyev 13]



No free lunch: SDP informal bottleneck



22 times slower than SDP: $q = q'$ formal bottleneck


Contribution

For more details on the formal side:



X. Allamigeon, S. Gaubert, V. Magron and B. Werner. Formal Proofs for Nonlinear Optimization. Submitted for publication, arxiv:1404.7282

Formal Nonlinear Optimization

- Formal nonlinear optimization: NLCertify 
- Safe solutions for challenging problems, e.g. Flyspeck

Formal Nonlinear Optimization

Further research:



OCAML API



Alternative Polynomials bounds using geometric programming (T. de Wolff, S. Ilman)



COQ tactic



Improve formal polynomial checker

Formal Nonlinear Optimization

Further research:

Generalized problem of moments

(Moment)		(SOS)	
\inf	$\int_{\mathbf{K}} p_0 d\mu$	$\geq \sup$	$\lambda_0 + \sum_i \lambda_i b_i$
s.t.	$\int_{\mathbf{K}} p_i d\mu \leq b_i$	s.t.	$\lambda_0, \lambda_i \leq 0$,
	$\mu \in \mathcal{M}_+(\mathbf{K})$		$p_0 - \lambda_0 - \sum_i \lambda_i p_i \in \mathcal{Q}_k(\mathbf{K})$

- Formal bounds using SDP and `ring`

End

Thank you for your attention!

<http://homepages.laas.fr/vmagron/>