Machine-Checkable Correctness Proofs: Formalizing Taylor Models

Roland Zumkeller

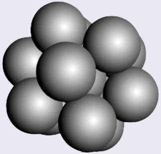
Project TypiCal, INRIA
École Polytechnique, Paris

Project MathComponents
Microsoft Research / INRIA Joint Lab, Paris

Taylor Model Methods V, May 2008, Toronto

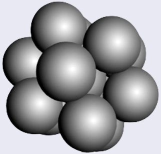
Conjecture (Johannes Kepler, 1611)

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



Conjecture (Johannes Kepler, 1611)

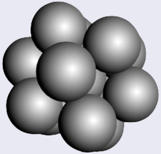
The maximal density of sphere packings in 3-space is $\frac{\pi}{\sqrt{18}}$.



Proof (Thomas Hales, 1998)

Conjecture (Johannes Kepler, 1611)

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



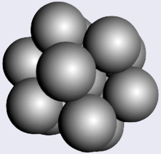
Proof (Thomas Hales, 1998)

300 pages

- Geometry
- Analysis

Conjecture (Johannes Kepler, 1611)

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



Proof (Thomas Hales, 1998)

300 pages

- Geometry
- Analysis

40.000 lines, several weeks

- Graph Enumeration
- Linear Optimization
- Non-linear Optimization

Inside the Proof: Slicing and Measuring Space

Lemma 751442360

$$2.51^2 \le x_1 \le 2.696^2 \rightarrow \quad 4 \le x_4 \le 2.51^2 \rightarrow$$
$$4 \le x_2 \le 2.168^2 \rightarrow \quad 4 \le x_5 \le 2.51^2 \rightarrow$$
$$4 \le x_3 \le 2.168^2 \rightarrow \quad 4 \le x_6 \le 2.51^2 \rightarrow$$

$$\frac{\begin{array}{c} -x_1x_3 - x_2x_4 + x_1x_5 + x_3x_6 - x_5x_6 + \\ x_2(-x_2 + x_1 + x_3 - x_4 + x_5 + x_6) \end{array}}{\sqrt{4x_2\left(\begin{array}{c} x_2x_4(-x_2 + x_1 + x_3 - x_4 + x_5 + x_6) + \\ x_1x_5(x_2 - x_1 + x_3 + x_4 - x_5 + x_6) + \\ x_3x_6(x_2 + x_1 - x_3 + x_4 + x_5 - x_6) \\ - x_1x_3x_4 - x_2x_3x_5 - x_2x_1x_6 - x_4x_5x_6 \end{array}\right)}} < \tan\left(\frac{\pi}{2} - 0.74\right)$$

Inside the Proof: Slicing and Measuring Space

Lemma 751442360

$$2.51^2 \leq x_1 \leq 2.696^2 \to \quad 4 \leq x_4 \leq 2.51^2 \to$$
$$4 \leq x_2 \leq 2.168^2 \to \quad 4 \leq x_5 \leq 2.51^2 \to$$
$$4 \leq x_3 \leq 2.168^2 \to \quad 4 \leq x_6 \leq 2.51^2 \to$$

$$\frac{\begin{array}{c} -x_1x_3 - x_2x_4 + x_1x_5 + x_3x_6 - x_5x_6 + \\ x_2(-x_2 + x_1 + x_3 - x_4 + x_5 + x_6) \end{array}}{\sqrt{4x_2 \left(\begin{array}{c} x_2x_4(-x_2 + x_1 + x_3 - x_4 + x_5 + x_6) + \\ x_1x_5(x_2 - x_1 + x_3 + x_4 - x_5 + x_6) + \\ x_3x_6(x_2 + x_1 - x_3 + x_4 + x_5 - x_6) \\ - x_1x_3x_4 - x_2x_3x_5 - x_2x_1x_6 - x_4x_5x_6 \end{array} \right)}} < \tan\left(\frac{\pi}{2} - 0.74\right)$$

Proof 1

Homegrown, Refined Interval Arithmic

Inside the Proof: Slicing and Measuring Space

Lemma 751442360

$$2.51^2 \leq x_1 \leq 2.696^2 \to \quad 4 \leq x_4 \leq 2.51^2 \to$$
$$4 \leq x_2 \leq 2.168^2 \to \quad 4 \leq x_5 \leq 2.51^2 \to$$
$$4 \leq x_3 \leq 2.168^2 \to \quad 4 \leq x_6 \leq 2.51^2 \to$$

$$\frac{\begin{array}{c}-x_1x_3 - x_2x_4 + x_1x_5 + x_3x_6 - x_5x_6 + \\ x_2(-x_2 + x_1 + x_3 - x_4 + x_5 + x_6)\end{array}}{\sqrt{4x_2\left(\begin{array}{c}x_2x_4(-x_2 + x_1 + x_3 - x_4 + x_5 + x_6) + \\ x_1x_5(x_2 - x_1 + x_3 + x_4 - x_5 + x_6) + \\ x_3x_6(x_2 + x_1 - x_3 + x_4 + x_5 - x_6) \\ - x_1x_3x_4 - x_2x_3x_5 - x_2x_1x_6 - x_4x_5x_6\end{array}\right)}} < \tan\left(\frac{\pi}{2} - 0.74\right)$$

Proof 1

Homegrown, Refined Interval Arithmic

Proof 2

Computer Algebra System ...

Inside the Proof: Slicing and Measuring Space

Lemma 751442360

$$2.51^2 \leq x_1 \leq 2.696^2 \rightarrow \qquad 4 \leq x_4 \leq 2.51^2 \rightarrow$$
$$4 \leq x_2 \leq 2.168^2 \rightarrow \qquad 4 \leq x_5 \leq 2.51^2 \rightarrow$$
$$4 \leq x_3 \leq 2.168^2 \rightarrow \qquad 4 \leq x_6 \leq 2.51^2 \rightarrow$$

$$\frac{\begin{array}{c} -x_1x_3 - x_2x_4 + x_1x_5 + x_3x_6 - x_5x_6 + \\ x_2(-x_2 + x_1 + x_3 - x_4 + x_5 + x_6) \end{array}}{\sqrt{4x_2\left(\begin{array}{c} x_2x_4(-x_2 + x_1 + x_3 - x_4 + x_5 + x_6) + \\ x_1x_5(x_2 - x_1 + x_3 + x_4 - x_5 + x_6) + \\ x_3x_6(x_2 + x_1 - x_3 + x_4 + x_5 - x_6) \\ - x_1x_3x_4 - x_2x_3x_5 - x_2x_1x_6 - x_4x_5x_6 \end{array}\right)}} < \tan\left(\frac{\pi}{2} - 0.74\right)$$

Proof 1

Homegrown, Refined Interval Arithmic

Proof 2

Computer Algebra System . . .

Proof 3

Proof Assistant: “Flyspeck” project

What is a proof?

Theorem

$$\forall n \in \mathbb{N}.\ \sum_{k=0}^{n} k = n(n+1)/2$$

Proof.

$$\begin{array}{ccccc} 1 & + & \dots & + & n \\ n & + & \dots & + & 1 \\ \hline (n+1) & + & \dots & + & (n+1) \end{array}$$

□

Theorem

$$\forall n \in \mathbb{N}.\ \sum_{k=0}^{n} k^2 = n(n^2+1)/2$$

Proof.

$$\begin{array}{ccccc} 1 & + & \ldots & + & n^2 \\ n^2 & + & \ldots & + & 1 \\ \hline (n^2+1) & + & \ldots & + & (n^2+1) \end{array}$$

□

Theorem

$$\forall n \in \mathbb{N}.\ \sum_{k=0}^{n} k^2 = n(n^2+1)/2$$

Proof.

$$\begin{array}{ccccc} 1 & + & \dots & + & n^2 \\ n^2 & + & \dots & + & 1 \\ \hline (n^2+1) & + & \dots & + & (n^2+1) \end{array}$$

□

Example

$$1 + 4 + 9 = 3 \cdot (9+1)/2,$$

i.e.

$$14 = 15.$$

Not a Theorem!

$$\forall n \in \mathbb{N}.\ \sum_{k=0}^{n} k^2 = n(n^2+1)/2$$

Proof by intimidation.

$$\begin{array}{ccccccc} 1 & + & 2 & + & \dots & + & n^2 \\ n^2 & + & (n-1)^2 & + & \dots & + & 1 \\ \hline (n^2+1) & + & (n^2-2n+3) & + & \dots & + & (n^2+1) \end{array}$$

□

Example

$$1+4+9 = 3 \cdot (9+1)/2,$$

i.e.

$$14 = 15.$$

Theorem

$$\forall n \in \mathbb{N}.\ \sum_{k=0}^{n} k = n(n+1)/2$$

A More Detailed Proof.

By induction on *n*.

- Basis: $0 = 0$
- Step: Suppose $\sum_{k=0}^{n} k = n(n+1)/2$. Then

$$\begin{aligned} \sum_{k=0}^{n+1} k &= \sum_{k=0}^{n} k + (n+1) \\ &= n(n+1)/2 + (n+1) && \text{by hypothesis} \\ &= (n+1)(n+2)/2 && \text{by algebra} \end{aligned}$$

Computer-Assisted Proofs

- What is a proof?
  $\Rightarrow$ An object that can *in principle* be refined to a formal proof.

Computer-Assisted Proofs

- What is a proof?
  $\Rightarrow$ An object that can *in principle* be refined to a formal proof.
- What is a formal proof? $\Rightarrow$ A proof in a formal language:
  - Frege's Begriffsschrift (1879)

Computer-Assisted Proofs

- What is a proof?
  $\Rightarrow$ An object that can *in principle* be refined to a formal proof.
- What is a formal proof? $\Rightarrow$ A proof in a formal language:
  - Frege's Begriffsschrift (1879)
  - de Bruijn's Automath system (1967)

Computer-Assisted Proofs

- What is a proof?
  $\Rightarrow$ An object that can *in principle* be refined to a formal proof.
- What is a formal proof? $\Rightarrow$ A proof in a formal language:
  - Frege's Begriffsschrift (1879)
  - de Bruijn's Automath system (1967)
  - Coq system

Computer-Assisted Proofs

- What is a proof?
  $\Rightarrow$ An object that can *in principle* be refined to a formal proof.
- What is a formal proof? $\Rightarrow$ A proof in a formal language:
  - Frege's Begriffsschrift (1879)
  - de Bruijn's Automath system (1967)
  - Coq system
- Computers can *assist* us to ...
  - ... find proofs.
  - ... <u>check</u> proofs.

Computer-Assisted Proofs

- What is a proof?
  $\Rightarrow$ An object that can *in principle* be refined to a formal proof.
- What is a formal proof? $\Rightarrow$ A proof in a formal language:
  - Frege's Begriffsschrift (1879)
  - de Bruijn's Automath system (1967)
  - Coq system
- Computers can *assist* us to ...
  - ... find proofs.
  - ... check proofs.
- Proof assistents are software themselves, so why should we trust them?

Computer-Assisted Proofs

- What is a proof?
  $\Rightarrow$ An object that can *in principle* be refined to a formal proof.
- What is a formal proof? $\Rightarrow$ A proof in a formal language:
  - Frege's Begriffsschrift (1879)
  - de Bruijn's Automath system (1967)
  - Coq system
- Computers can *assist* us to ...
  - ... find proofs.
  - ... check proofs.
- Proof assistents are software themselves, so why should we trust them?
  - Architecture: small, well-tested kernel

Computer-Assisted Proofs

- What is a proof?
  ⇒ An object that can *in principle* be refined to a formal proof.
- What is a formal proof? ⇒ A proof in a formal language:
  - Frege's Begriffsschrift (1879)
  - de Bruijn's Automath system (1967)
  - Coq system
- Computers can *assist* us to ...
  - ... find proofs.
  - ... <u>check</u> proofs.
- Proof assistents are software themselves, so why should we trust them?
  - Architecture: small, well-tested kernel
  - "Coq in Coq"

Big Proofs

Theorem

$$\forall x \in [0;1].\ 0 \leq f\,x$$

Proof.

Assume $x \in [0;1]$. Let $X_i := [(i-1)/n; i/n]$. Then

$$x \in X_1 \vee \ldots \vee x \in X_n.$$

In each of these cases $0 \leq \hat{f}\,X_i$ and thus $0 \leq f\,x$. □

Big Proofs

Theorem

$$\forall x \in [0;1].\ 0 \leq f\,x$$

Proof.

Assume $x \in [0;1]$. Let $X_i := [(i-1)/n; i/n]$. Then

$$x \in X_1 \vee \ldots \vee x \in X_n.$$

In each of these cases $0 \leq \hat{f}\,X_i$ and thus $0 \leq f\,x$. □

- The necessary n depends on f. Is there a largest n such that this a proof?

Big Proofs

Theorem

$$\forall x \in [0;1].\ 0 \leq f\,x$$

Proof.

Assume $x \in [0;1]$. Let $X_i := [(i-1)/n; i/n]$. Then

$$x \in X_1 \vee \ldots \vee x \in X_n.$$

In each of these cases $0 \leq \hat{f}\,X_i$ and thus $0 \leq f\,x$. □

- The necessary n depends on f. Is there a largest n such that this a proof?
- Non-toy examples with quite large "n": Four Color Theorem, Pocklington Prime Numbers

Taylor Models and Chebyshev Balls

Definition

Taylor models: $\mathbb{T}[n] := \mathbb{R}[n] \times \mathbb{I}$.
For $f : D \to \mathbb{R}$ (where $D \subseteq \mathbb{R}^n$),

$$f \mathrel{\tilde{\in}} (p, \Delta) :\Leftrightarrow \forall x \in D.\ f\,x - p\,x \in \Delta.$$

Taylor Models and Chebyshev Balls

Definition

Taylor models: $\mathbb{T}[n] := \mathbb{R}[n] \times \mathbb{I}$.
For $f : D \to \mathbb{R}$ (where $D \subseteq \mathbb{R}^n$),

$$f \mathrel{\tilde{\in}} (p, \Delta) :\Leftrightarrow \forall x \in D.\ f\,x - p\,x \in \Delta.$$

Definition

Chebyshev balls: $Ч[n] := \mathbb{R}[n] \times \mathbb{R}$.
For $f : D \to \mathbb{R}$ (where $D \subseteq \mathbb{R}^n$),

$$f \mathrel{\tilde{\in}} (p, \delta) :\Leftrightarrow \|f \mathbin{\mathring{-}} p\|_\infty \leq \delta.$$

Taylor Models and Chebyshev Balls

Definition

Taylor models: $\mathbb{T}[n] := \mathbb{R}[n] \times \mathbb{I}$.
For $f : D \to \mathbb{R}$ (where $D \subseteq \mathbb{R}^n$),

$$f \mathrel{\tilde{\in}} (p, \Delta) :\Leftrightarrow \forall x \in D.\ f\,x - p\,x \in \Delta.$$

Definition

Chebyshev balls: $\mathrm{Ч}[n] := \mathbb{R}[n] \times \mathbb{R}$.
For $f : D \to \mathbb{R}$ (where $D \subseteq \mathbb{R}^n$),

$$f \mathrel{\tilde{\in}} (p, \delta) :\Leftrightarrow \|f \mathbin{\mathring{-}} p\|_\infty \leq \delta.$$

- Chebyshev balls are centered Taylor models:
$$f \mathrel{\tilde{\in}} (p, \Delta) \Leftrightarrow f \mathrel{\tilde{\in}} \left(p + \mathrm{m}\,\Delta, \frac{|\Delta|}{2}\right)$$

Taylor Models and Chebyshev Balls

Definition

Taylor models: $\mathbb{T}[n] := \mathbb{R}[n] \times \mathbb{I}$.
For $f : D \to \mathbb{R}$ (where $D \subseteq \mathbb{R}^n$),

$$f \mathrel{\tilde{\in}} (p, \Delta) :\Leftrightarrow \forall x \in D.\ f\,x - p\,x \in \Delta.$$

Definition

Chebyshev balls: $\text{Ч}[n] := \mathbb{R}[n] \times \mathbb{R}$.
For $f : D \to \mathbb{R}$ (where $D \subseteq \mathbb{R}^n$),

$$f \mathrel{\tilde{\in}} (p, \delta) :\Leftrightarrow \|f \mathbin{\mathring{-}} p\|_\infty \leq \delta.$$

- Chebyshev balls are centered Taylor models:
$$f \mathrel{\tilde{\in}} (p, \Delta) \Leftrightarrow f \mathrel{\tilde{\in}} \left(p + \mathrm{m}\,\Delta, \frac{|\Delta|}{2}\right)$$
- Economy: Lemmas about $\|\cdot\|_\infty$ can be reused.

Extensions and Lifts

Definition

$$\begin{array}{ccccccccc} g & : & (\mathbb{R}^{n_1} \to \mathbb{R}) & \to & \ldots & \to & (\mathbb{R}^{n_r} \to \mathbb{R}) & \to & (\mathbb{R}^{n_{r+1}} \to \mathbb{R}) \\ G & : & Ч[n_1] & \to & \ldots & \to & Ч[n_r] & \to & Ч[n_{r+1}] \end{array}$$

G is an *extension* of g $:\Leftrightarrow$

$$\forall f, F.\ f_1 \mathrel{\tilde{\in}} F_1 \to \ldots \to f_r \mathrel{\tilde{\in}} F_r \to g\, f_1 \ldots f_r \mathrel{\tilde{\in}} G\, F_1 \ldots F_r.$$

Extensions and Lifts

Definition

$$\begin{array}{ccccccccc} g & : & (\mathbb{R}^{n_1} \to \mathbb{R}) & \to & \ldots & \to & (\mathbb{R}^{n_r} \to \mathbb{R}) & \to & (\mathbb{R}^{n_{r+1}} \to \mathbb{R}) \\ G & : & Ч[n_1] & \to & \ldots & \to & Ч[n_r] & \to & Ч[n_{r+1}] \end{array}$$

G is an *extension* of g $:\Leftrightarrow$

$$\forall f, F.\ f_1 \mathrel{\tilde{\in}} F_1 \to \ldots \to f_r \mathrel{\tilde{\in}} F_r \to g\, f_1 \ldots f_r \mathrel{\tilde{\in}} G\, F_1 \ldots F_r.$$

Definition

$$\begin{array}{ccc} g & : & \mathbb{R}^r \to \mathbb{R} \\ G & : & (Ч[n])^r \to Ч[n] \end{array}$$

G is a *lift* of g $:\Leftrightarrow$
G extends $f_1 \ldots f_r\, x_1 \ldots x_n \mapsto g\,(f_1\, x_1 \ldots x_n) \ldots (f_r\, x_1 \ldots x_n)$

Arithmetic

Definition

$$(p_1, \Delta_1) \tilde{+} (p_2, \Delta_2) := (p_1 + p_2, \Delta_1 \hat{+} \Delta_2)$$

$$(p_1, \Delta_1) \tilde{\cdot} (p_2, \Delta_2) := ((p_1 p_2)_{\leq l}, \overline{(p_1 p_2)_{>l} + \urcorner_1 p_2 + p_1 \urcorner_2 + \urcorner_1 \urcorner_2})$$

where $\urcorner_1 \in \Delta_1$ and $\urcorner_2 \in \Delta_2$ are fresh variables.

Arithmetic

Definition

$$(p_1, \Delta_1) \tilde{+} (p_2, \Delta_2) := (p_1 + p_2, \Delta_1 \hat{+} \Delta_2)$$

$$(p_1, \Delta_1) \tilde{\cdot} (p_2, \Delta_2) := ((p_1 p_2)_{\leq l}, \overline{(p_1 p_2)_{>l} + \daleth_1 p_2 + p_1 \daleth_2 + \daleth_1 \daleth_2})$$

where $\daleth_1 \in \Delta_1$ and $\daleth_2 \in \Delta_2$ are fresh variables.

Lemma

$\tilde{+}$ and $\tilde{\cdot}$ are lifts of $+$ and $\cdot$.

Arithmetic

Definition

$$(p_1, \Delta_1) \,\tilde{+}\, (p_2, \Delta_2) := (p_1 + p_2, \Delta_1 \,\hat{+}\, \Delta_2)$$

$$(p_1, \Delta_1) \,\tilde{\cdot}\, (p_2, \Delta_2) := ((p_1 p_2)_{\leq l}, \overline{(p_1 p_2)_{>l} + \daleth_1 p_2 + p_1 \daleth_2 + \daleth_1 \daleth_2})$$

where $\daleth_1 \in \Delta_1$ and $\daleth_2 \in \Delta_2$ are fresh variables.

Lemma

$\tilde{+}$ and $\tilde{\cdot}$ are lifts of $+$ and $\cdot$.

Proof (for $\tilde{\cdot}$).

Assume $f_1 \,\tilde{\in}\, (p_1, \Delta_1)$ and $f_2 \,\tilde{\in}\, (p_2, \Delta_2)$.
Let $d_1 := f_1 \,\mathring{-}\, p_1$ and $d_2 := f_2 \,\mathring{-}\, p_2$.

$$f_1 f_2 = (p_1 \,\mathring{+}\, d_1)(p_2 \,\mathring{+}\, d_2) = p_1 p_2 \,\mathring{+}\, p_1 d_2 \,\mathring{+}\, d_1 p_2 \,\mathring{+}\, d_1 d_2$$

□

Arithmetic

Definition

$$(p_1, \Delta_1) \tilde{+} (p_2, \Delta_2) := (p_1 + p_2, \Delta_1 \hat{+} \Delta_2)$$

$$(p_1, \Delta_1) \tilde{\cdot} (p_2, \Delta_2) := ((p_1 p_2)_{\leq l}, \overline{(p_1 p_2)_{>l} + \daleth_1 p_2 + p_1 \daleth_2 + \daleth_1 \daleth_2})$$

where $\daleth_1 \in \Delta_1$ and $\daleth_2 \in \Delta_2$ are fresh variables.

Lemma

$\tilde{+}$ and $\tilde{\cdot}$ are lifts of $+$ and $\cdot$.

Proof (for $\tilde{\cdot}$).

Assume $f_1 \mathrel{\tilde{\in}} (p_1, \Delta_1)$ and $f_2 \mathrel{\tilde{\in}} (p_2, \Delta_2)$.
Let $d_1 := f_1 \mathbin{\mathring{-}} p_1$ and $d_2 := f_2 \mathbin{\mathring{-}} p_2$.

$$f_1 f_2 = (p_1 \mathbin{\mathring{+}} d_1)(p_2 \mathbin{\mathring{+}} d_2) = p_1 p_2 \mathbin{\mathring{+}} p_1 d_2 \mathbin{\mathring{+}} d_1 p_2 \mathbin{\mathring{+}} d_1 d_2$$

$$\tilde{\in} ((p_1 p_2)_{\leq l}, \overline{(p_1 p_2)_{>l} + \daleth_1 p_2 + p_1 \daleth_2 + \daleth_1 \daleth_2})$$

□

Extending Function Composition

Definition

$$(p, \delta) \,\tilde{\circ}\, F := [p]\, F \,\tilde{+}\, (0, \delta)$$

Extending Function Composition

Definition

$$(p,\delta)\,\tilde{\circ}\,F := [p]\,F\,\tilde{+}\,(0,\delta)$$

Lemma

$\tilde{\circ} : Ч[1] \to Ч[n] \to Ч[n]$ *is an extension of* $\circ : (\mathbb{R} \to \mathbb{R}) \to (\mathbb{R}^n \to \mathbb{R}) \to (\mathbb{R}^n \to \mathbb{R})$.

Extending Function Composition

Definition

$$(p, \delta) \mathbin{\tilde{\circ}} F := [p]\, F \mathbin{\tilde{+}} (0, \delta)$$

Lemma

$\tilde{\circ} : Ч[1] \to Ч[n] \to Ч[n]$ is an extension of $\circ : (\mathbb{R} \to \mathbb{R}) \to (\mathbb{R}^n \to \mathbb{R}) \to (\mathbb{R}^n \to \mathbb{R})$.

Proof.

Assume $g \mathrel{\tilde{\in}} (p, \delta)$ and $f \mathrel{\tilde{\in}} F$. Then

$$\|g \mathbin{\mathring{-}} [p]\|_\infty \leq \delta.$$

Extending Function Composition

Definition

$$(p, \delta) \mathbin{\tilde{\circ}} F := [p]\, F \mathbin{\tilde{+}} (0, \delta)$$

Lemma

$\tilde{\circ} : \text{Ч}[1] \to \text{Ч}[n] \to \text{Ч}[n]$ is an extension of
$\circ : (\mathbb{R} \to \mathbb{R}) \to (\mathbb{R}^n \to \mathbb{R}) \to (\mathbb{R}^n \to \mathbb{R})$.

Proof.

Assume $g \mathrel{\tilde{\in}} (p, \delta)$ and $f \mathrel{\tilde{\in}} F$. Then

$$\|g \circ f \mathbin{\mathring{-}} [p] \circ f\|_\infty \leq \|g \mathbin{\mathring{-}} [p]\|_\infty \leq \delta.$$

Extending Function Composition

Definition

$$(p, \delta) \mathbin{\tilde{\circ}} F := [p]\, F \mathbin{\tilde{+}} (0, \delta)$$

Lemma

$\tilde{\circ} : \text{Ч}[1] \to \text{Ч}[n] \to \text{Ч}[n]$ is an extension of
$\circ : (\mathbb{R} \to \mathbb{R}) \to (\mathbb{R}^n \to \mathbb{R}) \to (\mathbb{R}^n \to \mathbb{R})$.

Proof.

Assume $g \mathrel{\tilde{\in}} (p, \delta)$ and $f \mathrel{\tilde{\in}} F$. Then

$$\|g \circ f \mathbin{\mathring{-}} [p] \circ f\|_\infty \leq \|g \mathbin{\mathring{-}} [p]\|_\infty \leq \delta.$$

Furthermore $[p] \circ f = [p]^\circ f \mathrel{\tilde{\in}} [p]^\sim F$, hence

$$g \circ f \mathrel{\tilde{\in}} [p]^\sim F \mathbin{\tilde{+}} (0, \delta) = (p, \delta) \mathbin{\tilde{\circ}} F.$$

□

Lifting Elementary Functions

Lemma

For $g : \mathbb{R} \mapsto \mathbb{R}$ *and* $G : Ч[1]$, *if* $g \mathrel{\tilde{\in}} G$ *then* $F \mapsto G \mathbin{\tilde{\circ}} F$ *lifts* g.

Lifting Elementary Functions

Lemma

For $g : \mathbb{R} \mapsto \mathbb{R}$ and $G : Ч[1]$, if $g \mathrel{\tilde{\in}} G$ then $F \mapsto G \mathbin{\tilde{\circ}} F$ lifts g.

Proof.

By definition of lift this means that $F \mapsto G \mathbin{\tilde{\circ}} F$ extends $f\,x \mapsto g\,(f\,x) = f \mapsto g \circ f$. The extension property is preserved by partial application. □

Lifting Elementary Functions

Lemma

For $g : \mathbb{R} \mapsto \mathbb{R}$ and $G : Ч[1]$, if $g \tilde{\in} G$ then $F \mapsto G \,\tilde{\circ}\, F$ lifts g.

Proof.

By definition of lift this means that $F \mapsto G \,\tilde{\circ}\, F$ extends $f\,x \mapsto g\,(f\,x) = f \mapsto g \circ f$. The extension property is preserved by partial application. □

Remaining Problem: Polynomial Approximation

For a given $g : X \mapsto \mathbb{R}$ (where $X \subset \mathbb{R}$) find $G : Ч[1]$ such that $g \tilde{\in} G$.

Lifting Elementary Functions

Lemma

For $g : \mathbb{R} \mapsto \mathbb{R}$ and $G : Ч[1]$, if $g \tilde{\in} G$ then $F \mapsto G \tilde{\circ} F$ lifts g.

Proof.

By definition of lift this means that $F \mapsto G \tilde{\circ} F$ extends $f\,x \mapsto g\,(f\,x) = f \mapsto g \circ f$. The extension property is preserved by partial application. □

Remaining Problem: Polynomial Approximation

For a given $g : X \mapsto \mathbb{R}$ (where $X \subset \mathbb{R}$) find $G : Ч[1]$ such that $g \tilde{\in} G$.

- Bernstein
- Taylor
- Chebyshev
- Remez

Lifting Elementary Functions

Lemma

For $g : \mathbb{R} \mapsto \mathbb{R}$ and $G : Ч[1]$, if $g \tilde{\in} G$ then $F \mapsto G \,\tilde{\circ}\, F$ lifts g.

Proof.

By definition of lift this means that $F \mapsto G \,\tilde{\circ}\, F$ extends $f\,x \mapsto g\,(f\,x) = f \mapsto g \circ f$. The extension property is preserved by partial application. □

Remaining Problem: Polynomial Approximation

For a given $g : X \mapsto \mathbb{R}$ (where $X \subset \mathbb{R}$) find $G : Ч[1]$ such that $g \tilde{\in} G$.

Bernstein	slow convergence
Taylor	
Chebyshev	
Remez	

Lifting Elementary Functions

Lemma

For $g : \mathbb{R} \mapsto \mathbb{R}$ and $G : Ч[1]$, if $g \mathrel{\tilde{\in}} G$ then $F \mapsto G \mathbin{\tilde{\circ}} F$ lifts g.

Proof.

By definition of lift this means that $F \mapsto G \mathbin{\tilde{\circ}} F$ extends $f\,x \mapsto g\,(f\,x) = f \mapsto g \circ f$. The extension property is preserved by partial application. □

Remaining Problem: Polynomial Approximation

For a given $g : X \mapsto \mathbb{R}$ (where $X \subset \mathbb{R}$) find $G : Ч[1]$ such that $g \mathrel{\tilde{\in}} G$.

Bernstein	slow convergence
Taylor	easy to implement, good *local* convergence
Chebyshev	
Remez	

Lifting Elementary Functions

Lemma

For $g : \mathbb{R} \mapsto \mathbb{R}$ *and* $G : Ч[1]$, *if* $g \mathrel{\tilde{\in}} G$ *then* $F \mapsto G \mathbin{\tilde{\circ}} F$ *lifts* g.

Proof.

By definition of lift this means that $F \mapsto G \mathbin{\tilde{\circ}} F$ extends $f\,x \mapsto g\,(f\,x) = f \mapsto g \circ f$. The extension property is preserved by partial application. □

Remaining Problem: Polynomial Approximation

For a given $g : X \mapsto \mathbb{R}$ (where $X \subset \mathbb{R}$) find $G : Ч[1]$ such that $g \mathrel{\tilde{\in}} G$.

Bernstein	slow convergence
Taylor	easy to implement, good *local* convergence
Chebyshev	Is there a good Jackson theorem?
Remez	

Lifting Elementary Functions

Lemma

For $g : \mathbb{R} \mapsto \mathbb{R}$ and $G : Ч[1]$, if $g \tilde{\in} G$ then $F \mapsto G \,\tilde{\circ}\, F$ lifts g.

Proof.

By definition of lift this means that $F \mapsto G \,\tilde{\circ}\, F$ extends $f\,x \mapsto g\,(f\,x) = f \mapsto g \circ f$. The extension property is preserved by partial application. □

Remaining Problem: Polynomial Approximation

For a given $g : X \mapsto \mathbb{R}$ (where $X \subset \mathbb{R}$) find $G : Ч[1]$ such that $g \tilde{\in} G$.

Bernstein	slow convergence
Taylor	easy to implement, good *local* convergence
Chebyshev	Is there a good Jackson theorem?
Remez	difficult to implement, but optimal

Lifting Elementary Functions

Lemma

For $g : \mathbb{R} \mapsto \mathbb{R}$ and $G : Ч[1]$, if $g \tilde{\in} G$ then $F \mapsto G \tilde{\circ} F$ lifts g.

Proof.

By definition of lift this means that $F \mapsto G \tilde{\circ} F$ extends $f\,x \mapsto g\,(f\,x) = f \mapsto g \circ f$. The extension property is preserved by partial application. □

Remaining Problem: Polynomial Approximation

For a given $g : X \mapsto \mathbb{R}$ (where $X \subset \mathbb{R}$) find $G : Ч[1]$ such that $g \tilde{\in} G$.

Bernstein	slow convergence
Taylor	easy to implement, good *local* convergence
Chebyshev	Is there a good Jackson theorem?
Remez	difficult to implement, but optimal

Lifting Elementary Functions: by Taylor/Lagrange

$$\mathrm{T}_a^l\, g\, x := \sum_{k=0}^{l} \frac{\partial^k g\, a}{k!}(x - a)^k$$

$$\mathrm{R}_a^l\, g := g \mathbin{\mathring{-}} \mathrm{T}_a^l\, g$$

$$\mathrm{L}_a^l\, g\, X := \frac{\partial^{l+1} g\, X}{(l+1)!}(X \mathbin{\hat{-}} a)^{l+1}$$

Taylor's Theorem with Lagrange remainder

$$\forall x \in X.\ \mathrm{R}_a^l\, g\, x \in \mathrm{L}_a^l\, g\, X$$

Lifting Elementary Functions: by Taylor/Lagrange

$$\mathrm{T}_a^l\, g\, x := \sum_{k=0}^{l} \frac{\partial^k g\, a}{k!}(x - a)^k$$

$$\mathrm{R}_a^l\, g := g \mathbin{\mathring{-}} \mathrm{T}_a^l\, g$$

$$\mathrm{L}_a^l\, g\, X := \frac{\partial^{l+1} g\, X}{(l+1)!}(X \mathbin{\hat{-}} a)^{l+1}$$

Taylor's Theorem with Lagrange remainder

$$\forall x \in X.\ \mathrm{R}_a^l\, g\, x \in \mathrm{L}_a^l\, g\, X$$

$$g \mathrel{\tilde{\in}} (\mathrm{T}_a^l\, g, \mathrm{L}_a^l\, g\, X)$$

Lifting Elementary Functions: by Taylor/Lagrange

$$\mathrm{T}_a^l\, g\, x := \sum_{k=0}^{l} \frac{\partial^k g\, a}{k!}(x - a)^k$$

$$\mathrm{R}_a^l\, g := g \mathbin{\mathring{-}} \mathrm{T}_a^l\, g$$

$$\mathrm{L}_a^l\, g\, X := \frac{\partial^{l+1} g\, X}{(l+1)!}(X \mathbin{\hat{-}} a)^{l+1}$$

Taylor's Theorem with Lagrange remainder

$$\forall x \in X.\ \mathrm{R}_a^l\, g\, x \in \mathrm{L}_a^l\, g\, X$$

$$g \mathrel{\tilde{\in}} (\mathrm{T}_a^l\, g, \mathrm{L}_a^l\, g\, X)$$

- No addition theorem needed. Move the value *a* instead.

Lifting Elementary Functions: by Taylor/Lagrange

$$\mathrm{T}_a^l\, g\, x := \sum_{k=0}^{l} \frac{\partial^k g\, a}{k!}(x - a)^k$$

$$\mathrm{R}_a^l\, g := g \mathbin{\mathring{-}} \mathrm{T}_a^l\, g$$

$$\mathrm{L}_a^l\, g\, X := \frac{\partial^{l+1} g\, X}{(l+1)!}(X \mathbin{\hat{-}} a)^{l+1}$$

Taylor's Theorem with Lagrange remainder

$$\forall x \in X.\ \mathrm{R}_a^l\, g\, x \in \mathrm{L}_a^l\, g\, X$$

$$g \mathrel{\tilde{\in}} (\mathrm{T}_a^l\, g, \mathrm{L}_a^l\, g\, X)$$

- No addition theorem needed. Move the value *a* instead.
- Taking the argument's constant part for *a* yields the same result as in [Makino-PhD] etc.

Lifting Elementary Functions: with Sharp Remainder

Observation

If

$$\forall x \in [x_1, x_2].\ \mathrm{sgn}(\partial\,(\mathrm{R}_a^l\, g)\, x) \geq 0$$

then

$$\forall x \in [x_1, x_2].\ \mathrm{R}_a^l\, g\, x \in \left[\mathrm{R}_a^l\, g\, x_1 ; \mathrm{R}_a^l\, g\, x_2\right].$$

Observation

If

$$\forall x \in [x_1, x_2].\ \mathrm{sgn}(\partial\,(\mathrm{R}^l_a\,g)\,x) \geq 0$$

then

$$\forall x \in [x_1, x_2].\ \mathrm{R}^l_a\,g\,x \in \left[\mathrm{R}^l_a\,g\,x_1 ; \mathrm{R}^l_a\,g\,x_2\right].$$

$\mathrm{sgn}(\partial\,(\mathrm{R}^l_a\,g)\,x) = \mathrm{sgn}(\mathrm{R}^{l-1}_a\,(\partial g)\,x)$ R and ∂ commute

Observation

If

$$\forall x \in [x_1, x_2].\ \mathrm{sgn}(\partial\,(\mathrm{R}_a^l\, g)\, x) \geq 0$$

then

$$\forall x \in [x_1, x_2].\ \mathrm{R}_a^l\, g\, x \in \left[\mathrm{R}_a^l\, g\, x_1 ; \mathrm{R}_a^l\, g\, x_2\right].$$

$$\begin{aligned} \mathrm{sgn}(\partial\,(\mathrm{R}_a^l\, g)\, x) &= \mathrm{sgn}(\mathrm{R}_a^{l-1}\,(\partial g)\, x) && \text{R and } \partial \text{ commute} \\ &\subseteq \mathrm{sgn}(\mathrm{L}_a^{l-1}\,(\partial g)\, X) && \text{Lagrange remainder} \end{aligned}$$

Observation

If
$$\forall x \in [x_1, x_2].\ \mathrm{sgn}(\partial\,(\mathrm{R}^l_a\, g)\, x) \geq 0$$
then
$$\forall x \in [x_1, x_2].\, \mathrm{R}^l_a\, g\, x \in \left[\mathrm{R}^l_a\, g\, x_1 ; \mathrm{R}^l_a\, g\, x_2\right].$$

$$\begin{aligned}
\mathrm{sgn}(\partial\,(\mathrm{R}^l_a\, g)\, x) &= \mathrm{sgn}(\mathrm{R}^{l-1}_a\,(\partial g)\, x) && \text{R and } \partial \text{ commute} \\
&\subseteq \mathrm{sgn}(\mathrm{L}^{l-1}_a\,(\partial g)\, X) && \text{Lagrange remainder} \\
&= \mathrm{sgn}\left(\frac{1}{l!} \cdot \partial^l g\, X \cdot (X \mathbin{\hat{-}} a)^l\right)
\end{aligned}$$

Observation

If

$$\forall x \in [x_1, x_2].\ \mathrm{sgn}(\partial\,(\mathrm{R}_a^l\, g)\, x) \geq 0$$

then

$$\forall x \in [x_1, x_2].\ \mathrm{R}_a^l\, g\, x \in \left[\mathrm{R}_a^l\, g\, x_1 ; \mathrm{R}_a^l\, g\, x_2\right].$$

$$\begin{aligned}
\mathrm{sgn}(\partial\,(\mathrm{R}_a^l\, g)\, x) &= \mathrm{sgn}(\mathrm{R}_a^{l-1}\,(\partial g)\, x) && \text{R and } \partial \text{ commute} \\
&\subseteq \mathrm{sgn}(\mathrm{L}_a^{l-1}\,(\partial g)\, X) && \text{Lagrange remainder} \\
&= \mathrm{sgn}\left(\frac{1}{l!} \cdot \partial^l g\, X \cdot (X \mathbin{\hat{-}} a)^l\right) \\
&= \mathrm{sgn}(\partial^l g\, X) \cdot \mathrm{sgn}(X \mathbin{\hat{-}} a)^l
\end{aligned}$$

Lemma

$$\partial \circ \mathrm{R}_a^l = \mathrm{R}_a^{l-1} \circ \partial$$

Proof.

$$\begin{aligned}
\partial \mathrm{R}_a^l g &= \partial x \mapsto g\,x - \sum_{k=0}^{l} \frac{\partial^k g\,a}{k!}(x-a)^k \\
&= x \mapsto \partial g\,x - \sum_{k=1}^{l} \frac{\partial^k g\,a}{(k-1)!}(x-a)^{k-1} \\
&= x \mapsto \partial g\,x - \sum_{k=0}^{l-1} \frac{\partial^k (\partial g)\,a}{k!}(x-a)^k \\
&= \mathrm{R}_a^{l-1}(\partial g)
\end{aligned}$$

□

Lifting Elementary Functions: by Remez

Remaining Problem: Polynomial Approximation

For a given $g : X \mapsto \mathbb{R}$ (where $X \subset \mathbb{R}$) find $G : Ч[1]$ such that $g \mathrel{\tilde{\in}} G$.

- This problem has an optimal solution: the Remez polynomial

Lifting Elementary Functions: by Remez

Remaining Problem: Polynomial Approximation

For a given $g : X \mapsto \mathbb{R}$ (where $X \subset \mathbb{R}$) find $G : Ч[1]$ such that $g \,\tilde{\in}\, G$.

- This problem has an optimal solution: the Remez polynomial
- The correctness proof is hard, but we don't need it: Once we have obtained $G = (p, \delta)$ we can *compute* $\|g - [p]\|_\infty$ by interval arithmetic.

Lifting Elementary Functions: by Remez

Remaining Problem: Polynomial Approximation

For a given $g : X \mapsto \mathbb{R}$ (where $X \subset \mathbb{R}$) find $G : \text{Ч}[1]$ such that $g \tilde{\in} G$.

- This problem has an optimal solution: the Remez polynomial
- The correctness proof is hard, but we don't need it: Once we have obtained $G = (p, \delta)$ we can *compute* $\|g - [p]\|_\infty$ by interval arithmetic.
- The polynomial p can be obtained from outside the proof assistant: Sollya system by Arenaire in Lyon

Lifting Elementary Functions: by Remez

Remaining Problem: Polynomial Approximation

For a given $g : X \mapsto \mathbb{R}$ (where $X \subset \mathbb{R}$) find $G : Ч[1]$ such that $g \mathrel{\tilde{\in}} G$.

- This problem has an optimal solution: the Remez polynomial
- The correctness proof is hard, but we don't need it: Once we have obtained $G = (p, \delta)$ we can *compute* $\|g - [p]\|_\infty$ by interval arithmetic.
- The polynomial p can be obtained from outside the proof assistant: Sollya system by Arenaire in Lyon
- Remez is slower than Taylor: build a reusable database for different domains and degrees

Summary

- Formal proofs are necesseary if we want to rely on software.

Summary

- Formal proofs are necesseary if we want to rely on software.
- Generalized Taylor models don't depend on Taylor's theorem.

Summary

- Formal proofs are necesseary if we want to rely on software.
- Generalized Taylor models don't depend on Taylor's theorem.
- Chebyshev balls simplify proofs.
- Don't use the Lagrange remainders if deratives' signs are constant.

Summary

- Formal proofs are necesseary if we want to rely on software.
- Generalized Taylor models don't depend on Taylor's theorem.
- Chebyshev balls simplify proofs.
- Don't use the Lagrange remainders if deratives' signs are constant.