Algorithmique numérique et fiabilité des calculs en arithmétique flottante
Introduction à l'arithmétique par intervalles

Cours M2 - ISFA

16 Janvier 2013

Agenda

Expressions, function extensions
- Functions
- Expressions and functions extensions
- Vectors, matrices

Cons and pros
- Cons : overestimation, complexity
- Pros : contractant iterations, Brouwer's theorem

Variants
- affine arithmetic, Taylor models arithmetic

Conclusions
- Historique
- Bibliographie

Definitions : functions

Definition :
an interval extension $\mathbf{f}$ of a function f satisfies

$$\forall \mathbf{x},\ f(\mathbf{x}) \subset \mathbf{f}(\mathbf{x}),\ \text{ and } \forall x,\ f(\{x\}) = \mathbf{f}(\{x\}).$$

Elementary functions : again, use the monotony.

$$\begin{array}{lcl} \exp \mathbf{x} & = & [\exp \underline{x}, \exp \overline{x}] \\ \log \mathbf{x} & = & [\log \underline{x}, \log \overline{x}] \text{ if } \underline{x} \geq 0, [-\infty, \log \overline{x}] \text{ if } \overline{x} > 0 \\ \sin[\pi/6, 2\pi/3] & = & [1/2, 1] \\ \ldots & & \end{array}$$

Agenda

Expressions, function extensions
- Functions
- Expressions and functions extensions
- Vectors, matrices

Cons and pros
- Cons : overestimation, complexity
- Pros : contractant iterations, Brouwer's theorem

Variants
- affine arithmetic, Taylor models arithmetic

Conclusions
- Historique
- Bibliographie

Definitions : function extension

Example : $f(x) = x^2 - x + 1$ with $x \in [-2, 1]$.
$[-2, 1]^2 - [-2, 1] + 1 = [0, 4] + [-1, 2] + 1 = [0, 7]$.
Since $x^2 - x + 1 = x(x - 1) + 1$, we get $[-2, 1] \cdot ([-2, 1] - 1) + 1 = [-2, 1] \cdot [-3, 0] + 1 = [-3, 6] + 1 = [-2, 7]$.
Since $x^2 - x + 1 = (x - 1/2)^2 + 3/4$, we get $([-2, 1] - 1/2)^2 + 3/4 = [-5/2, 1/2]^2 + 3/4 = [0, 25/4] + 3/4 = [3/4, 7] = f([-2, 1])$.

Problem with this definition : infinitely many interval extensions, syntactic use (instead of semantic).

How to choose the best extension ? How to choose a good one ?

Definitions : function extension

Mean value theorem of order 1 (Taylor expansion of order 1) :

$\forall x, \forall y, \exists \xi_{x,y} \in (x, y) \ : \ f(y) = f(x) + (y - x) \cdot f'(\xi_{x,y})$

Interval interpretation :

$\forall y \in \mathbf{x}, \forall \tilde{x} \in \mathbf{x}, \ f(y) \in f(\tilde{x}) + (y - \tilde{x}) \cdot \mathbf{f'}(\mathbf{x})$

$\Rightarrow f(\mathbf{x}) \subset f(\tilde{x}) + (\mathbf{x} - \tilde{x}) \cdot \mathbf{f'}(\mathbf{x})$

Mean value theorem of order 2 (Taylor expansion of order 2) :

$\forall x, \forall y, \exists \xi_{x,y} \in (x, y) : f(y) = f(x) + (y - x) \cdot f'(x) + \frac{(y-x)^2}{2} \cdot f''(\xi_{x,y})$

Interval interpretation :

$\forall y \in \mathbf{x}, \forall \tilde{x} \in \mathbf{x}, \ f(y) \in f(\tilde{x}) + (y - \tilde{x}) \cdot f'(\tilde{x}) + \frac{(y-\tilde{x})^2}{2} \cdot \mathbf{f''}(\mathbf{x})$

$\Rightarrow f(\mathbf{x}) \subset f(\tilde{x}) + (\mathbf{x} - \tilde{x}) \cdot f'(\tilde{x}) + \frac{(\mathbf{x}-\tilde{x})^2}{2} \cdot \mathbf{f''}(\mathbf{x})$

Definitions : function extension

No need to go further :

- it is difficult to compute (automatically) the derivatives of higher order,
  especially for multivariate functions ;
- there is no (theoretical) gain in quality.

Theorem :

- for the natural extension $\mathbf{f}$ of f, it holds $d(f(\mathbf{x}), \mathbf{f}(\mathbf{x})) \leq \mathcal{O}(w(\mathbf{x}))$
- for the first order Taylor extension $\mathbf{f}_{\mathbf{T}_1}$ of f, it holds $d(f(\mathbf{x}), \mathbf{f}_{\mathbf{T}_1}(\mathbf{x})) \leq \mathcal{O}(w(\mathbf{x})^2)$
- getting an order higher than 3 is impossible without the squaring operation, is difficult even with it...

Agenda

Expressions, function extensions
- Functions
- Expressions and functions extensions
- Vectors, matrices

Cons and pros
- Cons : overestimation, complexity
- Pros : contractant iterations, Brouwer's theorem

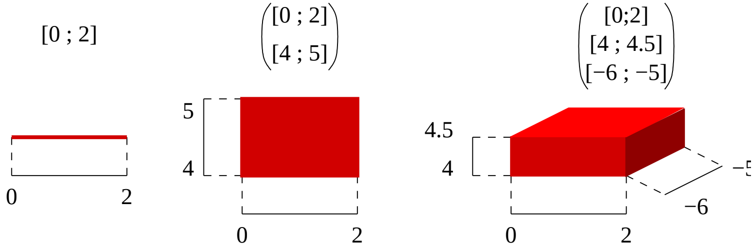
Variants
- affine arithmetic, Taylor models arithmetic

Conclusions
- Historique
- Bibliographie

Definitions : intervals, vectors, matrices

Objects :

- intervals of real numbers = closed connected sets of $\mathbb{R}$
  - interval for π : $[3.14159, 3.14160]$
  - data d measured with an absolute error less than $\pm\varepsilon$: $[d-\varepsilon, d+\varepsilon]$
- interval vector : components = intervals ; also called *box*



- interval matrix : components = intervals.

Agenda

Expressions, function extensions
- Functions
- Expressions and functions extensions
- Vectors, matrices

Cons and pros
- Cons : overestimation, complexity
- Pros : contractant iterations, Brouwer's theorem

Variants
- affine arithmetic, Taylor models arithmetic

Conclusions
- Historique
- Bibliographie

Cons : overestimation (1/2)

The result encloses the true result, but it is too large : overestimation phenomenon.
Two main sources : variable dependency and wrapping effect.

(Loss of) Variable dependency :

$$\mathbf{x} - \mathbf{x} = \{x - y \ : \ x \in \mathbf{x}, y \in \mathbf{x}\} \neq \{x - x \ : \ x \in \mathbf{x}\} = \{0\}.$$

Cons : overestimation (2/2)

Wrapping effect

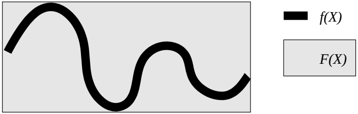
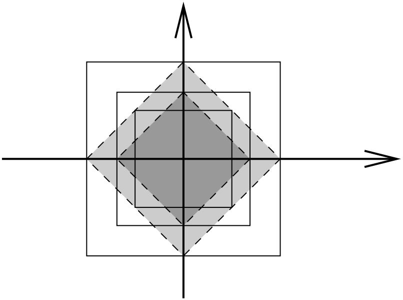


image of $f(\mathbf{x})$
with $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$



2 successive rotations of $\pi/4$
of the little central square

Cons : Complexity : almost every problem is NP-hard

Gaganov 1982, Rohn 1994 ff, Kreinovich...

- evaluate a function on a box (cartesian product of intervals)
- evaluate a function on a box up to ε
- solve a linear system
- solve a linear system up to $1/4n^4$ ($n =$ dim. of the system)
- determine if the solution of a linear system is bounded
- compute the matrix norm $\|\mathbf{A}\|_{\infty,1}$
- determine if an interval matrix ($=$ a matrix with interval coefficients) is regular, i.e. if every possible punctual matrix in it is regular
- ...

Cons : Complexity : Gaganov 1982

evaluation of a multivariate polynomial with rational coeff. on a box is NP-hard

Idea : reduce polynomially CNF-3 to this problem.

On n boolean variables $q_1, \cdots, q_n$, a formula f in CNF-3 is defined by

$$f = \bigwedge_{i=1}^{m} f_i \text{ with } f_i = \bigvee_{j=1}^{1,2 or 3} r_{i,j}$$

with $r_{i,j} = q_{k_{i,j}}$ or $r_{i,j} = \neg q_{k_{i,j}}$.

Cons : Complexity : Gaganov 1982

evaluation of a multivariate polynomial with rational coeff. on a box is NP-hard

To each boolean variable q_i, let us associate a real variable $x_i \in [0, 1]$.
Meaning : $x_i = 0$ if $q_i = F$ and $x_i = 1$ if $q_i = T$.

Goal : get a polynomial which takes only values in $[0, 1]$
i.e. allow only product of terms or of $(1-$ term$)$.

Cons : Complexity : Gaganov 1982

evaluation of a multivariate polynomial with rational coeff. on a box is NP-hard

To each boolean variable q_i, let us associate a real variable $x_i \in [0, 1]$.
Meaning : $x_i = 0$ if $q_i = F$ and $x_i = 1$ if $q_i = T$.

Goal : get a polynomial which takes only values in $[0, 1]$
i.e. allow only product of terms or of ($1-$ term).
A product corresponds to a conjunction and $1 - x$ to a negation

Cons : Complexity : Gaganov 1982

evaluation of a multivariate polynomial with rational coeff. on a box is NP-hard

To each boolean variable q_i, let us associate a real variable $x_i \in [0, 1]$.
Meaning : $x_i = 0$ if $q_i = F$ and $x_i = 1$ if $q_i = T$.

Goal : get a polynomial which takes only values in $[0, 1]$
i.e. allow only product of terms or of $(1-$ term$)$.
A product corresponds to a conjunction and $1 - x$ to a negation
$\Rightarrow$ express f and the f_i using conjonctions and negations

Cons : Complexity : Gaganov 1982

evaluation of a multivariate polynomial with rational coeff. on a box is NP-hard

To each boolean variable q_i, let us associate a real variable $x_i \in [0, 1]$.
Meaning : $x_i = 0$ if $q_i = F$ and $x_i = 1$ if $q_i = T$.

Goal : get a polynomial which takes only values in $[0, 1]$
i.e. allow only product of terms or of $(1-$ term$)$.
A product corresponds to a conjunction and $1 - x$ to a negation
$\Rightarrow$ express f and the f_i using conjonctions and negations
$\Rightarrow$ express the f_i as $\neg \bigwedge_{j=1}^{1,2or3} \neg r_{i,j}$.

Cons : Complexity : Gaganov 1982

evaluation of a multivariate polynomial with rational coeff. on a box is NP-hard

More precisely :

1. to each $r_{i,j}$ let us associate a polynomial $y_{i,j}$ (corresponding to the negation of $r_{i,j}$) defined by

$$\begin{array}{rclcrcl} r_{i,j} & = & q_{k_{i,j}} & \rightarrow & y_{i,j}(x) & = & 1 - x_{k_{i,j}} \\ r_{i,j} & = & \neg q_{k_{i,j}} & \rightarrow & y_{i,j}(x) & = & x_{k_{i,j}} \end{array}$$

Cons : Complexity : Gaganov 1982

evaluation of a multivariate polynomial with rational coeff. on a box is NP-hard

More precisely :

1. to each $r_{i,j}$ let us associate a polynomial $y_{i,j}$ (corresponding to the negation of $r_{i,j}$) defined by
$$\begin{array}{ccccccc} r_{i,j} & = & q_{k_{i,j}} & \rightarrow & y_{i,j}(x) & = & 1 - x_{k_{i,j}} \\ r_{i,j} & = & \neg q_{k_{i,j}} & \rightarrow & y_{i,j}(x) & = & x_{k_{i,j}} \end{array}$$
2. to each f_i, let us associate a polynomial p_i (corresponding to the negation of f_i) defined by
$f_i = \bigvee r_{i,j} = \neg \bigwedge \neg r_{i,j} \ \rightarrow \ p_i(x) = \prod y_{i,j}(x).$

Cons : Complexity : Gaganov 1982

evaluation of a multivariate polynomial with rational coeff. on a box is NP-hard

More precisely :

1. to each $r_{i,j}$ let us associate a polynomial $y_{i,j}$ (corresponding to the negation of $r_{i,j}$) defined by
$$\begin{array}{rclcrcl} r_{i,j} & = & q_{k_{i,j}} & \to & y_{i,j}(x) & = & 1 - x_{k_{i,j}} \\ r_{i,j} & = & \neg q_{k_{i,j}} & \to & y_{i,j}(x) & = & x_{k_{i,j}} \end{array}$$
2. to each f_i, let us associate a polynomial p_i (corresponding to the negation of f_i) defined by
$f_i = \bigvee r_{i,j} = \neg \bigwedge \neg r_{i,j} \ \to \ p_i(x) = \prod y_{i,j}(x)$.
3. to f, let us associate the polynomial p defined by
$f = \bigwedge_{i=1}^{m} f_i \ \to \ p(x) = \prod_{i=1}^{m} (1 - p_i(x))$.

Cons : Complexity : Gaganov 1982

evaluation of a multivariate polynomial with rational coeff. on a box is NP-hard

Lemma :

1. $\forall x \in [0,1]$, $p(x) \in [0,1]$.
2. if α is a boolean vector and β is the associated $0-1$ vector, then
$$\begin{array}{rclcrcl} f(\alpha) & = & T & \Rightarrow & p(\beta) & = & 1 \\ f(\alpha) & = & F & \Rightarrow & p(\beta) & = & 0. \end{array}$$
3. if f is not feasible, then $\forall x \in [0,1]^n$, $p(x) \leq 7/8$.

Cons : Complexity : Gaganov 1982

Proof of (3) : (proving (1) and (2) is easy).
$\forall x \in [0,1]^n$, let us consider β the 0-1 vector obtained by rounding x to the nearest.
Since f is not feasible, $p(\beta) = 0$.
Since $p(x) = \prod_{i=1}^m (1 - p_i(x))$, $\exists i_0$ such that $1 - p_{i_0}(\beta) = 0$.
One can prove that $p_{i_0}(x) \geq 1/8$, using the fact that it is the product of at most three terms, each of them $\geq 1/2$, using the fact that β is the rounding to nearest of x. Thus $1 - p_{i_0}(x) \leq 7/8$.
The remaining factors $1 - p_j(x)$ are less or equal to 1.
Thus $p(x) = \prod_{i=1}^m (1 - p_i(x)) \leq 7/8$.

Consequence : since checking the feasibility of a CNF-3 formula is NP-hard, evaluating a multivariate polynomial (up to a small ε) is NP-hard.

Agenda

Expressions, function extensions
- Functions
- Expressions and functions extensions
- Vectors, matrices

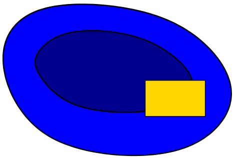
Cons and pros
- Cons : overestimation, complexity
- Pros : contractant iterations, Brouwer's theorem

Variants
- affine arithmetic, Taylor models arithmetic

Conclusions
- Historique
- Bibliographie

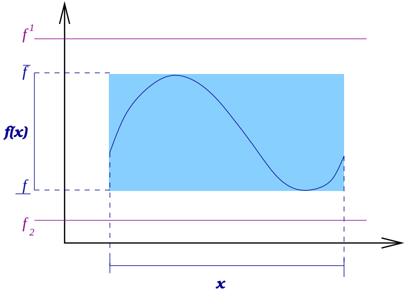
Pros : set computing

Behaviour safe ?
controllable ? dangerous ?



always controllable.

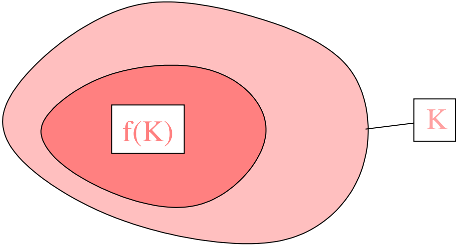
On **x**, are the extrema of the function f $> f^1$, $< f_2$?



No if $f(\mathbf{x}) = [\underline{f}, \overline{f}] \subset [f_2, f^1]$.

Pros : Brouwer-Schauder theorem

A function f which is continuous on the unit ball B and which satisfies $f(B) \subset B$ has a fixed point on B.
Furthermore, if $f(B) \subset \text{int}B$ then f has a unique fixed point on B.



The theorem remains valid if B is replaced by a compact K and in particular an interval.

Agenda

Expressions, function extensions
- Functions
- Expressions and functions extensions
- Vectors, matrices

Cons and pros
- Cons : overestimation, complexity
- Pros : contractant iterations, Brouwer's theorem

Variants
- affine arithmetic, Taylor models arithmetic

Conclusions
- Historique
- Bibliographie

Affine arithmetic

Comba, Stolfi and Figueiredo (1993, 2004)

Definition : each input or computed quantity x is represented by
$x = x_0 + \alpha_1\varepsilon_1 + \alpha_2\varepsilon_2 + \cdots + \alpha_n\varepsilon_n$
where x_0, α_1, $\ldots\alpha_n$ are known real / floating-point numbers,
and ε_1, ε_2 $\ldots\varepsilon_n$ are symbolic variables $\in [-1, +1]$.
Example : $x \in [3, 7]$ is represented by $x = 5 + 2\varepsilon$.

Operations :
$(x + \sum_k \alpha_k\varepsilon_k) + (y + \sum_k \beta_k\varepsilon_k) = (x + y) + \sum_k(\alpha_k + \beta_k)\varepsilon_k$.
$(x+\sum_k \alpha_k\varepsilon_k)\times(y+\sum_k \beta_k\varepsilon_k) = (x\times y)+\sum_k(x\beta_k+y\alpha_k)\varepsilon_k+\gamma_l\varepsilon_l$
with ε_l a new variable.

Roundoff errors : compute δ_l an upper bound of all roundoff errors and add it to γ_l.

Taylor models

Berz, Hoefkens and Makino 1998, Nedialkov, Neher

Principle : represent a function $f(x)$ for $x \in [-1, 1]$ by a polynomial part $p(x)$ and a reminder part (a big bin) I such that $\forall x \in [-1, 1]$, $f(x) \in p(x) + I$.

Operations :

- affine operations : straigthforward ;
- non-affine operations : enclose the nonlinear terms and add this enclosure to the reminder.

Roundoff errors : determine an upper bound b on the roundoff errors and add $[-b, b]$ to the reminder.

Agenda

Expressions, function extensions
- Functions
- Expressions and functions extensions
- Vectors, matrices

Cons and pros
- Cons : overestimation, complexity
- Pros : contractant iterations, Brouwer's theorem

Variants
- affine arithmetic, Taylor models arithmetic

Conclusions
- Historique
- Bibliographie

Historique de l'arithmétique par intervalles

- **1962 :** Ramon Moore définit l'arithmétique par intervalles dans sa thèse de doctorat et complète ce travail dans un livre en 1966

Historique de l'arithmétique par intervalles

- **1962 :** Ramon Moore définit l'arithmétique par intervalles dans sa thèse de doctorat et complète ce travail dans un livre en 1966
- **1958 :** Tsunaga, dans son mémoire de mastère en japonais

Historique de l'arithmétique par intervalles

- **1962 :** Ramon Moore définit l'arithmétique par intervalles dans sa thèse de doctorat et complète ce travail dans un livre en 1966
- **1958 :** Tsunaga, dans son mémoire de mastère en japonais
- **1956 :** Warmus

Historique de l'arithmétique par intervalles

- **1962 :** Ramon Moore définit l'arithmétique par intervalles dans sa thèse de doctorat et complète ce travail dans un livre en 1966
- **1958 :** Tsunaga, dans son mémoire de mastère en japonais
- **1956 :** Warmus
- **1951 :** Dwyer, dans un cas particulier

Historique de l'arithmétique par intervalles

- **1962 :** Ramon Moore définit l'arithmétique par intervalles dans sa thèse de doctorat et complète ce travail dans un livre en 1966
- **1958 :** Tsunaga, dans son mémoire de mastère en japonais
- **1956 :** Warmus
- **1951 :** Dwyer, dans un cas particulier
- **1931 :** Rosalind Cecil Young dans sa thèse de doctorat à Cambridge (RU) utilise quelques formules

Historique de l'arithmétique par intervalles

- **1962 :** Ramon Moore définit l'arithmétique par intervalles dans sa thèse de doctorat et complète ce travail dans un livre en 1966
- **1958 :** Tsunaga, dans son mémoire de mastère en japonais
- **1956 :** Warmus
- **1951 :** Dwyer, dans un cas particulier
- **1931 :** Rosalind Cecil Young dans sa thèse de doctorat à Cambridge (RU) utilise quelques formules
- **1927 :** Bradis, dans un cas encore plus particulier, en russe

Historique de l'arithmétique par intervalles

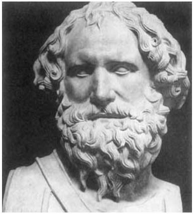
- **1962 :** Ramon Moore définit l'arithmétique par intervalles dans sa thèse de doctorat et complète ce travail dans un livre en 1966
- **1958 :** Tsunaga, dans son mémoire de mastère en japonais
- **1956 :** Warmus
- **1951 :** Dwyer, dans un cas particulier
- **1931 :** Rosalind Cecil Young dans sa thèse de doctorat à Cambridge (RU) utilise quelques formules
- **1927 :** Bradis, dans un cas encore plus particulier, en russe
- **1908 :** Young, dans un autre cas particulier, en italien

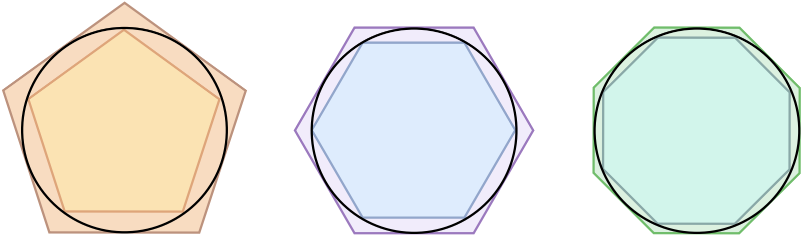
Historique de l'arithmétique par intervalles

- **1962 :** Ramon Moore définit l'arithmétique par intervalles dans sa thèse de doctorat et complète ce travail dans un livre en 1966
- **1958 :** Tsunaga, dans son mémoire de mastère en japonais
- **1956 :** Warmus
- **1951 :** Dwyer, dans un cas particulier
- **1931 :** Rosalind Cecil Young dans sa thèse de doctorat à Cambridge (RU) utilise quelques formules
- **1927 :** Bradis, dans un cas encore plus particulier, en russe
- **1908 :** Young, dans un autre cas particulier, en italien
- **3e siècle avant JC :** Archimède, pour calculer un encadrement de π !

Cf. http://www.cs.utep.edu/interval-comp/, cliquer sur *Early papers by Others*.

Archimède et un encadrement de π





Historical remarks

Childhood until the seventies.

Popularization in the 1980, German school (U. Kulisch).

IEEE-754 standard for floating-point arithmetic in 1985 :
directed roundings are standardized and available (?).

Since the nineties : interval **algorithms**.

IEEE-1788 standard for interval arithmetic in 2014 ?
I hope so. . .

Agenda

Expressions, function extensions
- Functions
- Expressions and functions extensions
- Vectors, matrices

Cons and pros
- Cons : overestimation, complexity
- Pros : contractant iterations, Brouwer's theorem

Variants
- affine arithmetic, Taylor models arithmetic

Conclusions
- Historique
- Bibliographie

References on interval arithmetic

- R. Moore : *Interval Analysis*, Prentice Hall, Englewood Cliffs, 1966.
- A. Neumaier : *Interval methods for systems of equations*, CUP, 1990.
- R. Moore, R.B. Kearfott, M.J. Cloud : *Introduction to interval analysis*, SIAM, 2009.
- W. Tucker : *Validated Numerics : A Short Introduction to Rigorous Computations*, Princeton University Press, 2011.
- S.M. Rump : *Computer-assisted proofs and Self-Validating Methods*, pp. 195-240. Handbook on Accuracy and Reliability in Scientific Computation (B. Einarsson ed.), SIAM, 2005.
- S.M. Rump : *Verification methods : Rigorous results using floating-point arithmetic*, Acta Numerica, vol. 19, pp. 287-449, 2010.

References on interval arithmetic

- J. Rohn : *A Handbook of Results on Interval Linear Problems*, http://www.cs.cas.cz/rohn/handbook 2006.
- E. Hansen and W. Walster : *Global optimization using interval analysis*, MIT Press, 2004.
- R.B. Kearfott : *Rigorous global search : continuous problems*, Kluwer, 1996.
- V. Kreinovich, A. Lakeyev, J. Rohn, P. Kahl : *Computational Complexity and Feasibility of Data Processing and Interval Computations*, Dordrecht, 1997.
- L.H. Figueiredo, J. Stolfi : *Affine arithmetic* http://www.ic.unicamp.br/~stolfi/EXPORT/projects/affine-arith/.
- *Taylor models arith. :* M. Berz and K. Makino, N. Nedialkov, M. Neher.