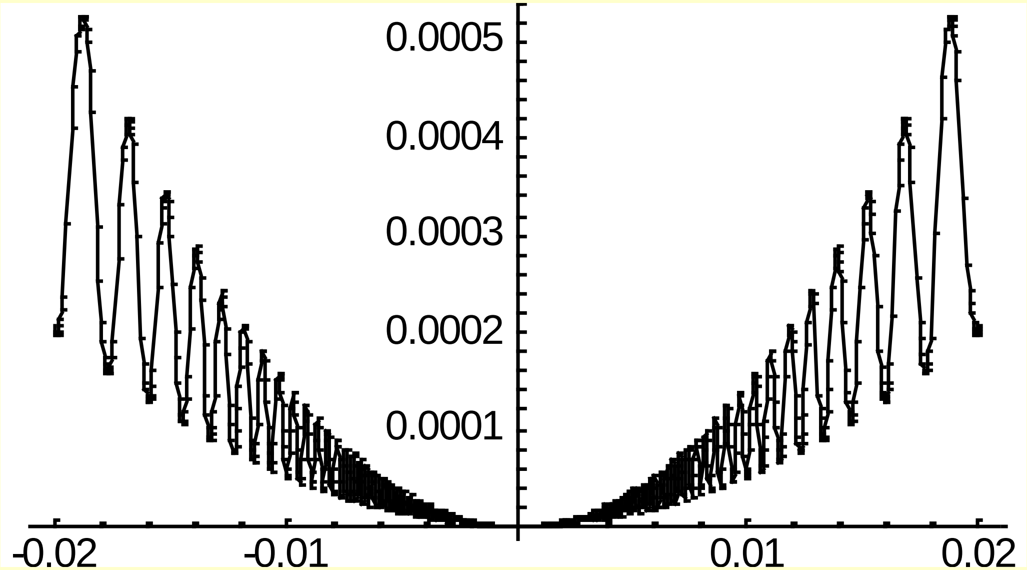
New inclusion functions in interval global optimization of engineering structures

Andrzej Pownuk
Chair of Theoretical Mechanics
Faculty of Civil Engineering
Silesian University of Technology

Global optimization

- Gradient methods
- Stochastic methods
- Special methods
  (e.g. linear programming)
- Analytical methods
- **Branch and bound methods**

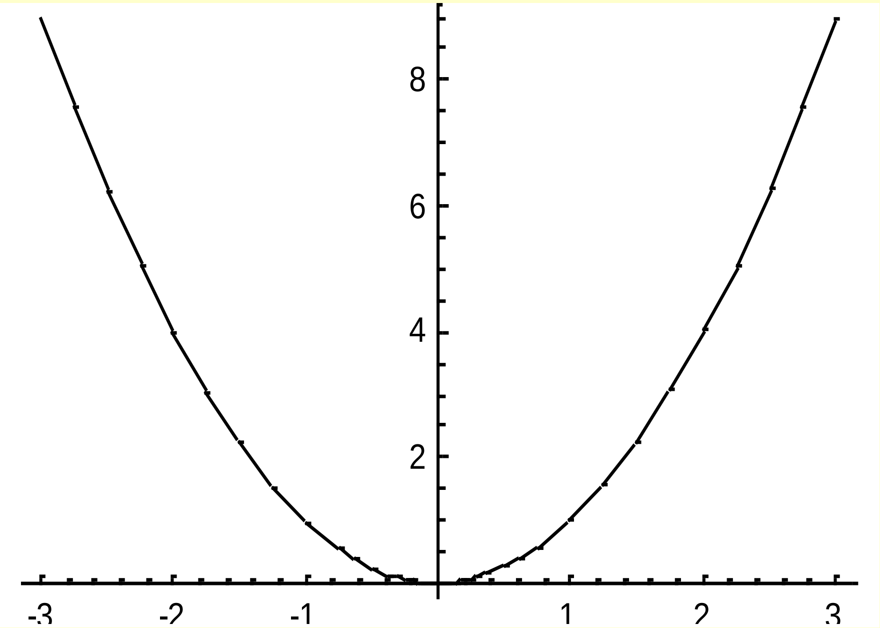
Problem with gradient methods



Many local minima

$$f(x) = x^2 \cdot \left(1 - 0.5 \cdot \cos\left(\frac{1}{x}\right)\right)$$

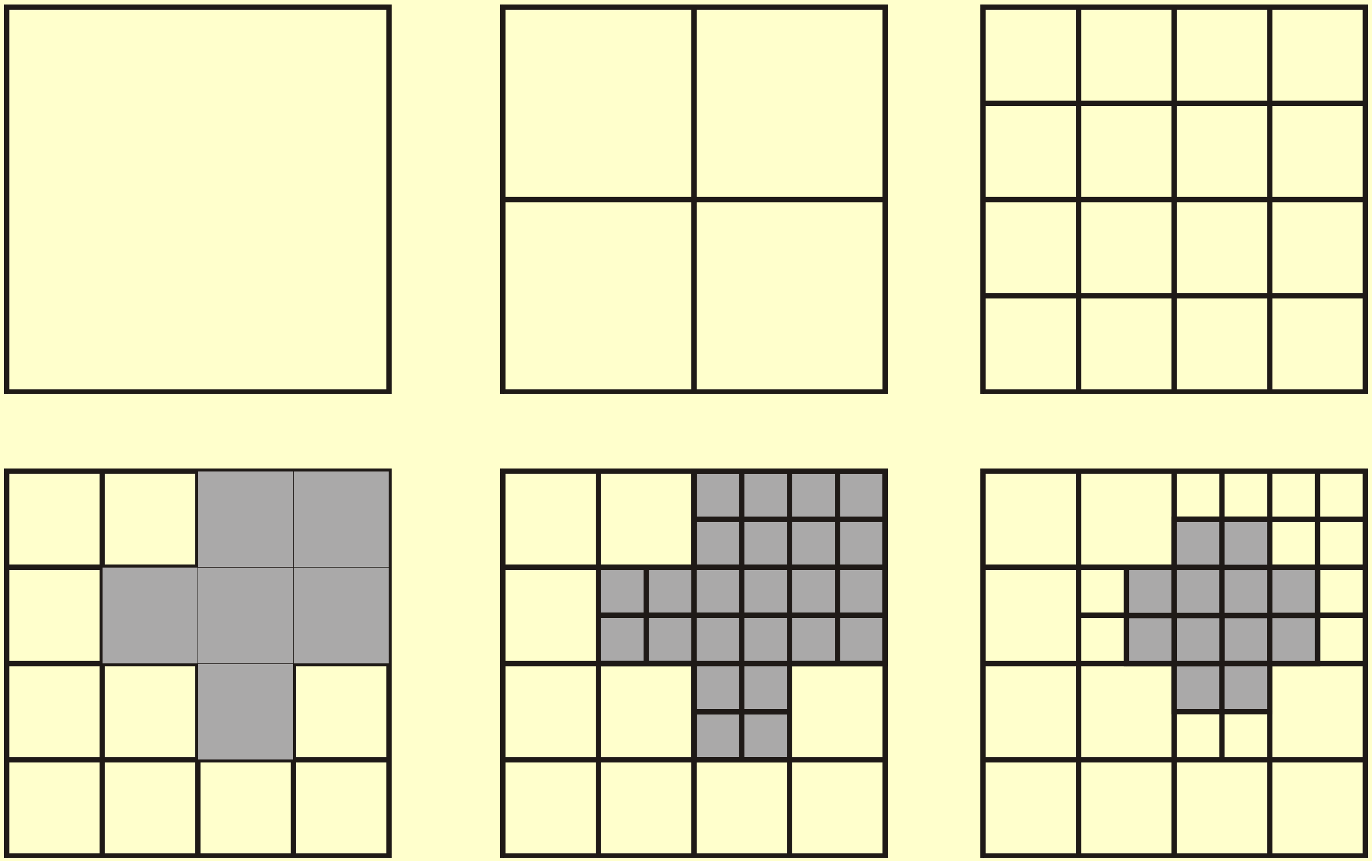
Problem with stochastic methods



Convergence criteria

$$y = x^2$$

Branch and bound methods



Contents

Properties of interval global optimization

Interval arithmetic

Acceleration devices

New inclusion functions

Examples of applications

Conclusions

Properties of interval global optimization

The algorithm guarantees that all stationary global solutions (in the initial interval) have been found.

The bounds on the solution(s) are guaranteed to be correct.

Error from all sources are accounted for.

The algorithm can solve the global optimization problem when the objective function is nondifferentiable or even not continuous.

Software

GlobSol - SUN Microsystems and Marquette University

Numerica - Pascal Van Hentenryck, Laurent Michel, i Yves Deville

UniCalc - Russian Research Institute of Artificial Intelligence

GLOBOPT - Arnold Neumaier, University of Vienna

Interval arithmetic

$$\bar{x} = [x^-, x^+] = \{x \in R : x^- \le x \le x^+\}$$

Interval operations

$$\bar{x} \circ \bar{y} = \left\{ x \circ y : x \in \bar{x}, y \in \bar{y} \right\}$$

for example

$$\bar{x} + \bar{y} = [x^- + y^-, x^+ + y^+]$$

Interval extension

$$f(x) = x^2 - x$$

Interval extension

$$\hat{f}(x) = \overline{x} \cdot \overline{x} - \overline{x}$$

Interval function

$$f(\overline{\mathbf{x}}) = \{ f(\overline{\mathbf{x}}) : \mathbf{x} \in \overline{\mathbf{x}} \}$$

Interval extension

$$\hat{f}([-1,2]) = [-1,2]\cdot[-1,2] - [-1,2] = [-4,5]$$

$$[-1,2]\cdot[-1,2] = [-2,4]$$

$$[-2,4] - [-1,2] = [-2,4] + [-2,1] = [-4,5]$$

$$f([-1,2]) = \{ f(x) : x \in [-1,2]\} = [-\frac{1}{4},2]$$

$$\hat{f}([-1,2]) = [-1,2]\cdot[-1,2] - [-1,2] = [-4,5]$$

$$f([-1,2]) = \{f(x) : x \in [-1,2]\} = [-\frac{1}{4},2]$$

$$f([-1,2]) \subseteq \hat{f}([-1,2])$$

Fundamental property of interval arithmetic

$$f(\overline{\mathbf{x}}) \subseteq \hat{f}(\overline{\mathbf{x}})$$

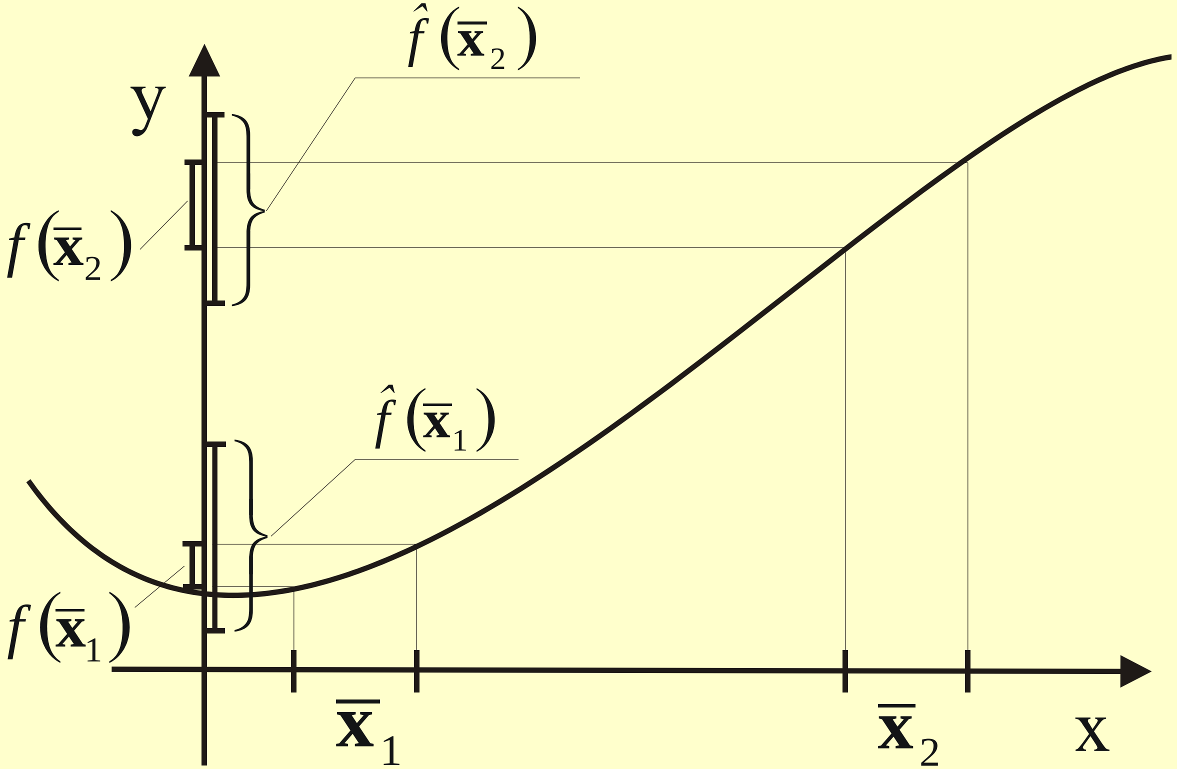
$$f(x) = x^2 - x$$

$$\min_{x\in[-1,2]} f(x) = -\frac{1}{4} \qquad \max_{x\in[-1,2]} f(x) = 2$$

$$\min \hat{f}([-1,2]) = -4 \qquad \max \hat{f}([-1,2]) = 5$$

$$\min \hat{f}(\overline{x}) \le \min_{x\in\overline{x}} f(x) \le \max_{x\in\overline{x}} f(x) \le \max \hat{f}(\overline{x})$$

Basic algorithm



$$\hat{f}(\bar{\mathbf{x}}_1)^{+} < \hat{f}(\bar{\mathbf{x}}_2)^{-}$$

Example

$$f(x) = x^2 - x, \quad \overline{x} = [-10, 10]$$

$$x_{opt} = \frac{1}{2}, \quad f_{opt} = -\frac{1}{4}$$

$$L = \{\overline{x}\}$$

Iteration 1

$\overline{x}$	[-10, 10]
$\hat{f}(\overline{x})$	[-110, 110]
$f(mid(\overline{x}))$	0

$$L = \{\overline{x}_1, \ldots, \overline{x}_n\}$$

$$f^- = min\{inf\ \hat{f}(\overline{x}_1), \ldots, inf\ \hat{f}(\overline{x}_n)\}$$

$$f^+ = max\{sup\ \hat{f}(\overline{x}_1), \ldots, sup\ \hat{f}(\overline{x}_n)\}$$

$$f_{opt} = min\{f(mid(\overline{x}_1)), \ldots, f(mid(\overline{x}_n))\}$$

$$f^- \leq f_{opt} \leq f^+$$

$$x^- = min\{inf\ \overline{x}_1, \ldots, inf\ \overline{x}_n\}$$

$$x^+ = max\{sup\ \overline{x}_1, \ldots, sup\ \overline{x}_n\}$$

$$x^- \leq x_{opt} \leq x^+$$

Iteration 1

$\overline{x}$	[-10, 10]
$\hat{f}(\overline{x})$	[-110, 110]
$f(mid(\overline{x}))$	0

$$f^- \le f_{opt} \le f^+$$

$$f^- = -110 \quad f_{opt} = 0 \quad f^+ = 110$$

$$x^- \le x_{opt} \le x^+$$

$$x^- = -10 \quad x_{opt} = 0 \quad x^+ = 10$$

$$f^+ - f^- = 220$$

$$x^+ - x^- = 20$$

Iteration 2

$\overline{x}$	$\hat{f}(\overline{x})$	$f(mid(\overline{x}))$
[-10, 0]	[0, 110]	30
[0, 10]	[-10, 100]	20

$$f^{-} = -10 \quad f_{opt} = 0 \quad f^{+} = 110$$

$$x^{-} = -10 \quad x_{opt} = 0 \quad x^{+} = 10$$

$$f^{+} - f^{-} = 120$$

$$x^{+} - x^{-} = 20$$

Iteration 3

$\overline{x}$	$\hat{f}(\overline{x})$	$f(mid(\overline{x}))$
[-10, -5]	[**30**, 110]	63.75
[-5, 0]	[0, 30]	8.75
[0, 5]	[-5, 25]	3.75
[5, 10]	[**15**, 95]	48.75

$$f^{-}=-5,\ f_{opt}=0,\ f^{+}=30$$

$$x^{-}=-5,\ x_{opt}=0,\ x^{+}=5$$

$$f^{+}-f^{-}=35$$

$$x^{+}-x^{-}=10$$

Iteration 4

$\overline{x}$	$\hat{f}(\overline{x})$	$f(mid(\overline{x}))$
[-5, -2.5]	[**8.75**, 30]	17.81
[-2.5, 0]	[0, 8.75]	2.81
[0, 2.5]	[-2.5, 6.25]	0.31
[2.5, 5]	[**1.25**, 22.5]	10.31

$$f^- = -2.5,\; f_{opt} = 0,\; f^+ = 8.75$$

$$x^- = -2.5,\; x_{opt} = 0,\; x^+ = 2.5$$

$$f^+ - f^- = 11.25$$

$$x^+ - x^- = 5$$

Iteration 5

$\bar{x}$	$\hat{f}(\bar{x})$	$f(mid(\bar{x}))$
[-2.5, -1.25]	[**2.81**, 8.75]	5.39
[-1.25, 0]	[**0**, 2.81]	1.02
[0, 1.25]	[-1.25, 1.56]	**-0.23**
[1.25, 2.5]	[-0.93, 5]	1.64

$$f^- = -1.25,\ f_{opt} = -0.23,\ f^+ = 5$$

$$x^- = 0,\ x_{opt} = 0.625,\ x^+ = 2.5$$

$$f^+ - f^- = 6.25$$

$$x^+ - x^- = 2.5$$

Iteration 6

$\overline{x}$	$\hat{f}(\overline{x})$	$f(mid(\overline{x}))$
[0, 0.625]	[-0.625, 0.39]	-0.214
[0.625, 1.25]	[-0.859, 0.93]	-0.059
[1.25,1.875]	[-0.31, 2.26]	0.878
[1.875, 2.5]	[**1.01**, 4.37]	1.64

$$f^- = -0.859,\ f_{opt} = -0.23,\ f^+ = 2.26$$

$$x^- = 0,\ x_{opt} = 0.625,\ x^+ = 1.875$$

$$f^+ - f^- = 3.119$$

$$x^+ - x^- = 1.875$$

Basic algorithm

Step 0 Set $[\mathbf{y}] = [\mathbf{x}]$ and $y = \hat{f}([\mathbf{x}])^-$. Initialize the list L = $([\mathbf{y}], y)$) and the cut-off level $z = \hat{f}([\mathbf{x}])^+$.

Step 1 Choose a coordinate direction $k \in \{1,2,\ldots,n\}$.

Step 2 Bisect [**y**] in direction k: $[\mathbf{y}] = [\mathbf{v}_1] \cup [\mathbf{v}_2]$.

Step 3 Calculate $\hat{f}([\mathbf{v}_1])$ and $\hat{f}([\mathbf{v}_2])$ and set $v_i = \hat{f}([\mathbf{v}_i])^-$ for i=1,2 and $z = \min\{z, \hat{f}([\mathbf{v}_1])^+, \hat{f}([\mathbf{v}_2])^+\}$.

Step 4 Remove $([\mathbf{y}], y)$ from the list L.

Step 5 Cutoff test: discard the pair $([\mathbf{v}_i], v_i)$ if $v_i > z$ (where i=1, 2).

Step 6 Add any remaining pair(s) to the list L. If the list becomes empty then STOP.

Step 7 Denote the pair with the smallest second element by $([\mathbf{y}], y)$.

Step 8 If the width of $\hat{f}([\mathbf{y}])$ is less than ε, then print $\hat{f}([\mathbf{y}])$ and $[\mathbf{y}]$, STOP.

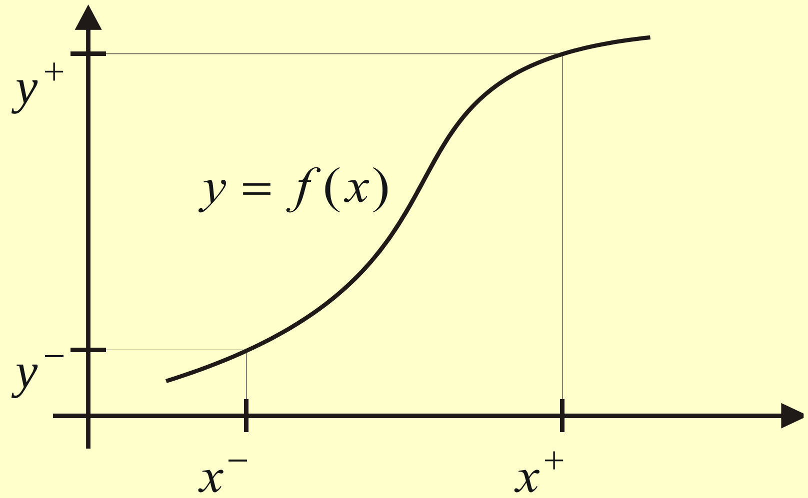
Step 9 Go to step 1.

Acceleration devices

Monotonicity test

$$0 \notin \frac{\partial \hat{f}\left(\overline{\mathbf{x}}\right)}{\partial \mathrm{x}_j} \quad \Longrightarrow \quad \forall \mathbf{x} \in \overline{\mathbf{x}}, \frac{\partial f\left(\mathbf{x}\right)}{\partial x_i} \neq 0$$

and function $y = f(\ldots, x_i, \ldots)$ is monotone



Acceleration devices

Monotonicity test

$$\text{If } \frac{df}{dx} > 0, \quad \text{then } y^- = f(x^-), \quad y^+ = f(x^+)$$

$$\text{If } \frac{df}{dx} < 0, \quad \text{then } y^- = f(x^+), \quad y^+ = f(x^-)$$

it can be written

$$y^- = f(x^{-sign\left(\frac{df}{dx}\right)}), \quad y^+ = f(x^{sign\left(\frac{df}{dx}\right)})$$

Acceleration devices

Midpoint test

If $f\left(mid(\overline{\mathbf{x}}_1)\right) < inf \quad \hat{f}\left(\overline{\mathbf{x}}_2\right)$

then $\forall \mathbf{x} \in \overline{\mathbf{x}}_2,\ f(mid(\mathbf{x}_1)) < f(\mathbf{x})$

and $\mathbf{x}_{opt} \notin \overline{\mathbf{x}}_2$

$$f(\mathbf{x}_{opt}) = \inf_{\mathbf{x}:\, \mathbf{x} \in \overline{\mathbf{x}}} f(\mathbf{x})$$

The interval $\overline{\mathbf{x}}_2$ can be neglected in a future calculation

Acceleration devices

A nonconvexity check

If the function f
has unconstrained minimum at $\mathbf{x}^*$,
then f must be convex
in some neighborhood of $\mathbf{x}^*$.

Hence, the Hessian $\mathbf{H}$
of f must be positive semidefinite at $\mathbf{x}^*$.

A necessary condition for this
is that the diagonal elements H_{ii} (i=1,..., n)
be nonnegative.

Acceleration devices

A nonconvexity check

Consider an interval $\overline{\mathbf{x}}$.

If $\hat{H}_{ii}(\overline{\mathbf{x}})^{+} < 0$ for some i = 1,..., n

then $H_{ii}(\mathbf{x}) < 0$ for all $\mathbf{x} \in \overline{\mathbf{x}}$.

Hence, **H** cannot be positive semidefinitive for any point in $\overline{\mathbf{x}}$.

Therefore, f cannot have a stationary minimum in $\overline{\mathbf{x}}$ and $\overline{\mathbf{x}}$ can be deleted.

Acceleration devices

The interval Newton method

1) Let $x^* \in \bar{x}_0$
2) $x_n = mid\left(\bar{x}_n\right)$
3) $N\left(x_n, \bar{x}_n\right) = x_n - \dfrac{f\left(x_n\right)}{\hat{f}'\left(\bar{x}_n\right)}$
4) $\bar{x}_{n+1} = \bar{x}_n \cap N\left(x_n, \bar{x}_n\right)$
5) Go to 1

Using this method we can find
all solutions of algebraic equations

Acceleration devices

The interval Newton method

If $\forall x \in \bar{x}, grad\ f(x) \neq 0,$
then $x_{opt} \notin \bar{x}$

The interval $\bar{x}$ can be neglected
in a future calculation

Acceleration devices

Finding a function value as small as possible

If $f\left(\mathbf{x}^{*}\right) < inf \quad \hat{f}\left(\overline{\mathbf{x}}_2\right)$ where $x^{*} \in \overline{x}$

then $\forall \mathbf{x} \in \overline{\mathbf{x}}_2,\ f\,(\mathbf{x}^{*}) < f\,(\mathbf{x})$

Midpoint test gives better results
if the number $f\left(\mathbf{x}^{*}\right)$
is as small as possible.
The point $\mathbf{x}^{*}$ can be found
using any optimization method.

Acceleration devices

Use a good inclusion function

In a calculation it is better to use
an inclusion function rather
than the natural interval extension

Fundamental property
of interval arithmetic

$$f(\overline{x}) \subseteq \hat{f}(\overline{x})$$

or

$$w(f(\overline{x})) \leq w(\hat{f}(\overline{x}))$$

where $w(\overline{x}) = x^{+} - x^{-}$

Interval extension gives
overestimated results.

New inclusion function

$$0 \notin \frac{\partial \hat{f}(\overline{\mathbf{x}})}{\partial x_i}$$

First order monotonicity test

$$\mathbf{y}^{-} = f\left(\mathbf{x}^{-sign\left(\frac{\partial \hat{f}(\overline{\mathbf{x}})}{\partial \mathbf{x}}\right)}\right),\ \mathbf{y}^{+} = f\left(\mathbf{x}^{sign\left(\frac{\partial \hat{f}(\overline{\mathbf{x}})}{\partial \mathbf{x}}\right)}\right)$$

$$f(\overline{\mathbf{x}}) = \overline{\mathbf{y}} = [\mathbf{y}^{-}, \mathbf{y}^{+}]$$

Exact result

Second order monotonicity test

$$0 \in \frac{\partial \hat{f}(\overline{\mathbf{x}})}{\partial x_i} \qquad 0 \notin \frac{\partial^2 \hat{f}(\overline{\mathbf{x}})}{\partial x_i \partial x_j}$$

$$\mathbf{y}_1^- = \frac{\partial f}{\partial \mathbf{x}}\left(\mathbf{x}^{-sign\left(\frac{\partial^2 \hat{f}(\overline{\mathbf{x}})}{\partial \mathbf{x}^2}\right)}\right), \quad \mathbf{y}_1^+ = \frac{\partial f}{\partial \mathbf{x}}\left(\mathbf{x}^{sign\left(\frac{\partial^2 \hat{f}(\overline{\mathbf{x}})}{\partial \mathbf{x}^2}\right)}\right)$$

$$\overline{\mathbf{y}}_1 = [\mathbf{y}_1^-, \mathbf{y}_1^+] \text{ (exact result)}$$

If $0 \notin \overline{y}_{1i}$, then

$$\mathbf{y}^- = f(\mathbf{x}^{-sign(\overline{\mathbf{y}}_1)}), \; \mathbf{y}^+ = f(\mathbf{x}^{sign(\overline{\mathbf{y}}_1)})$$

N-th order monotonicity test

$$\mathbf{y}_N^- = inf\,\frac{\partial^N \hat{f}(\overline{\mathbf{x}})}{\partial \mathbf{x}^N}, \quad \mathbf{y}_N^+ = sup\,\frac{\partial^N \hat{f}(\overline{\mathbf{x}})}{\partial \mathbf{x}^N}$$

$$\overline{\mathbf{y}}_N = [\overline{\mathbf{y}}_N^-, \overline{\mathbf{y}}_N^+]$$

$$\mathbf{y}_k^- = \frac{\partial^k f}{\partial \mathbf{x}^k}(\mathbf{x}^{-sign(\overline{\mathbf{y}}_{k+1})}), \; \mathbf{y}_k^+ = \frac{\partial^k f}{\partial \mathbf{x}^k}(\mathbf{x}^{sign(\overline{\mathbf{y}}_{k+1})})$$

If $\mathbf{0} \notin \overline{\mathbf{y}}_k$, then
the method gives exact result

Taylor series expansion

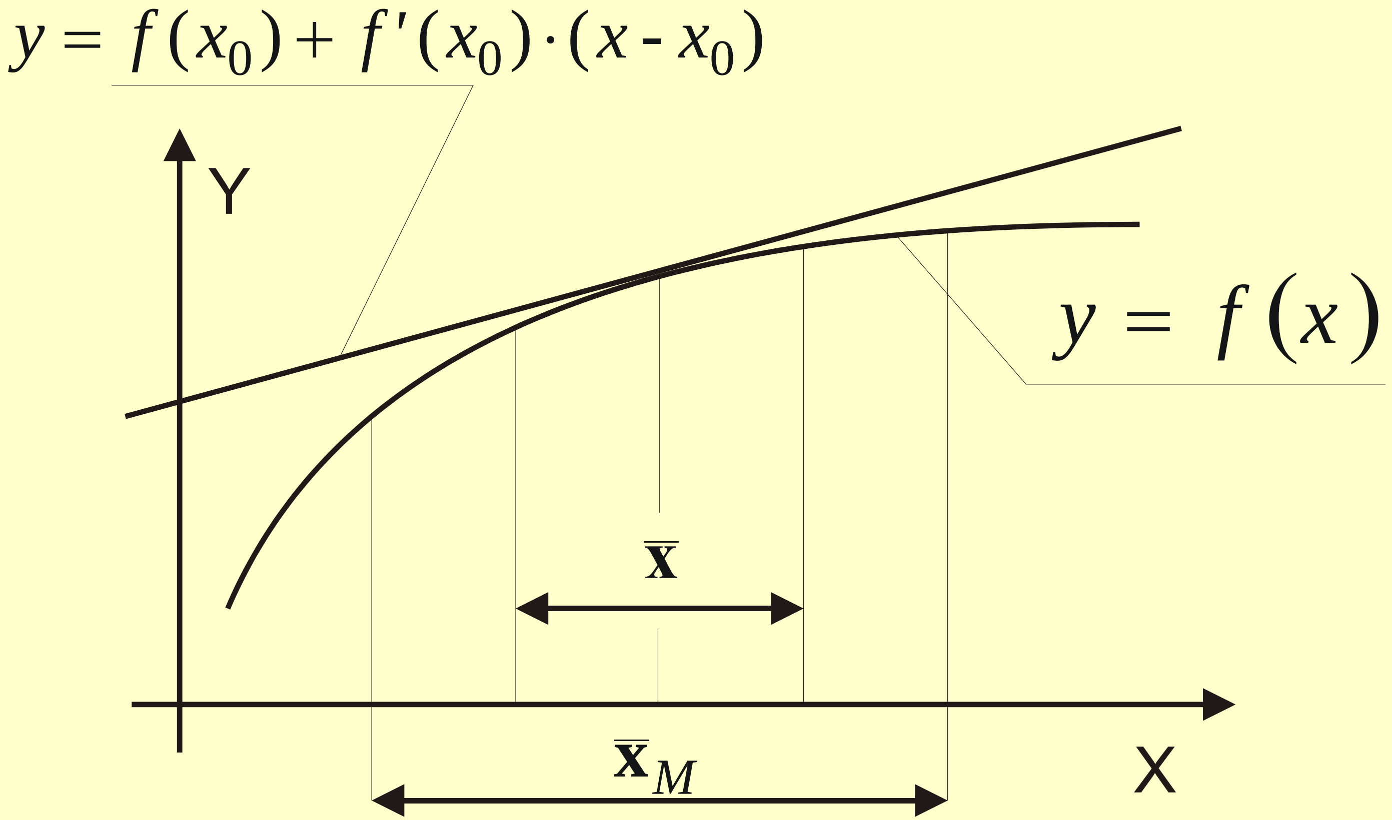
$$f(\mathbf{x}) \approx f_N(\mathbf{x}) = \sum_{i=1}^{N} \frac{1}{i!} \frac{\partial^i f(\mathbf{x}_0)}{\partial \mathbf{x}^i} (\mathbf{x} - \mathbf{x}_0)^k$$

To calculation the coefficients $\frac{\partial^k f(\mathbf{x}_0)}{\partial \mathbf{x}^k}$

we can apply sensitivity analysis.

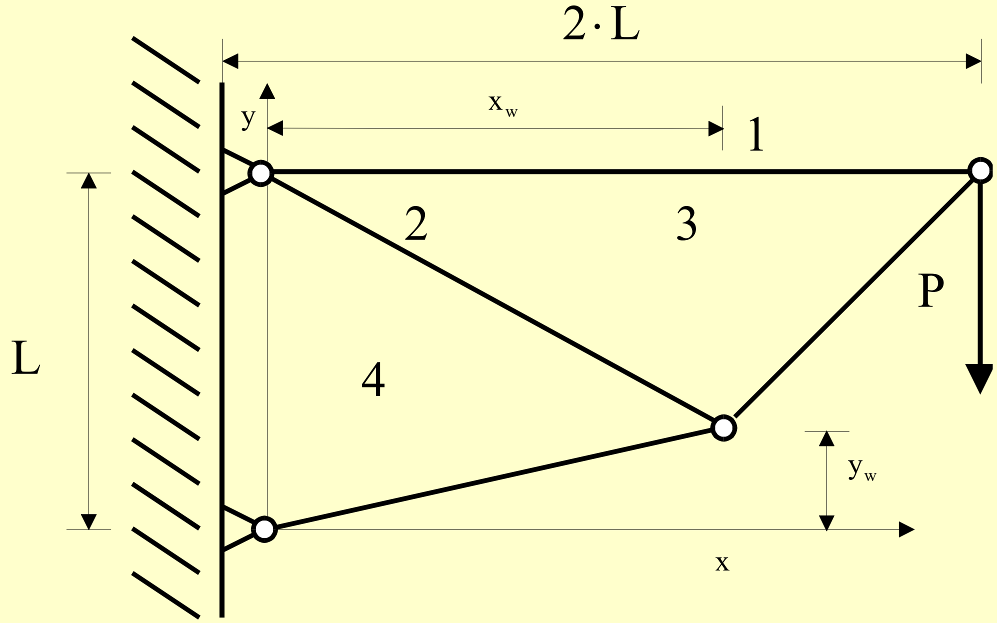
This method is more efficient than pure interval method.

This method cannot get results with guaranteed accuracy.

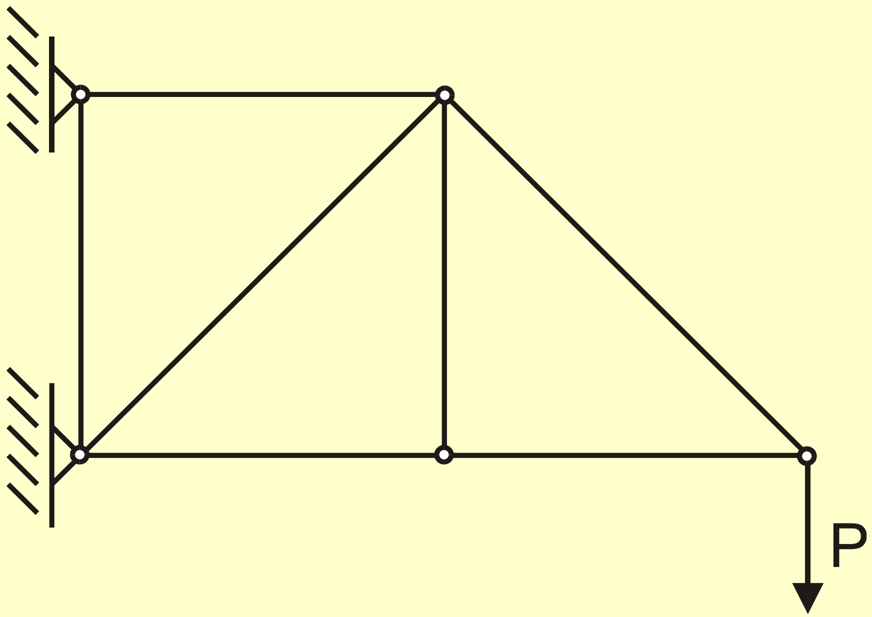


$$\overline{x} \subseteq \overline{x}_M \qquad 2\cdot w(\overline{x}) = w(\overline{x}_M)$$

Shape optimisation of truss

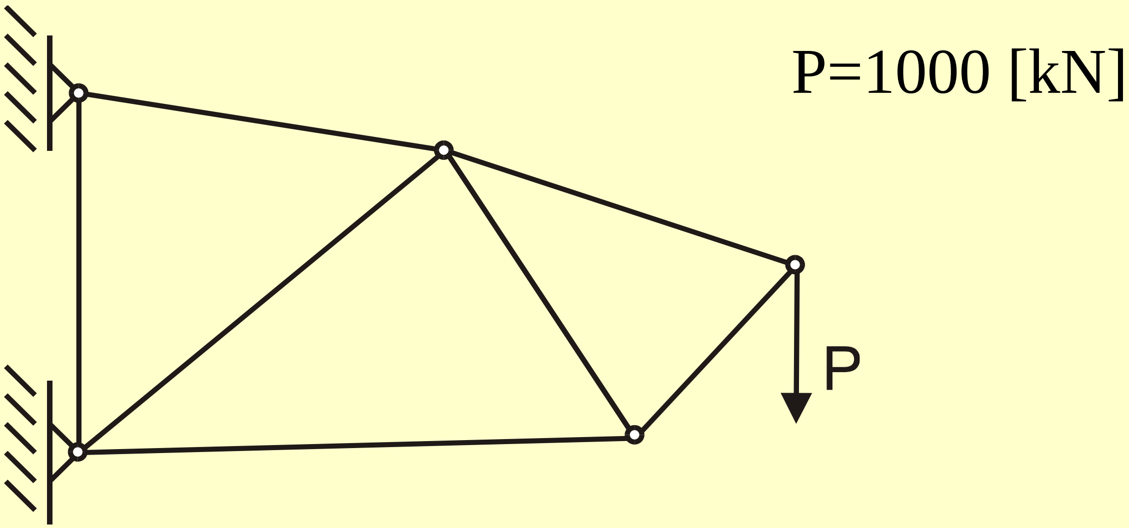


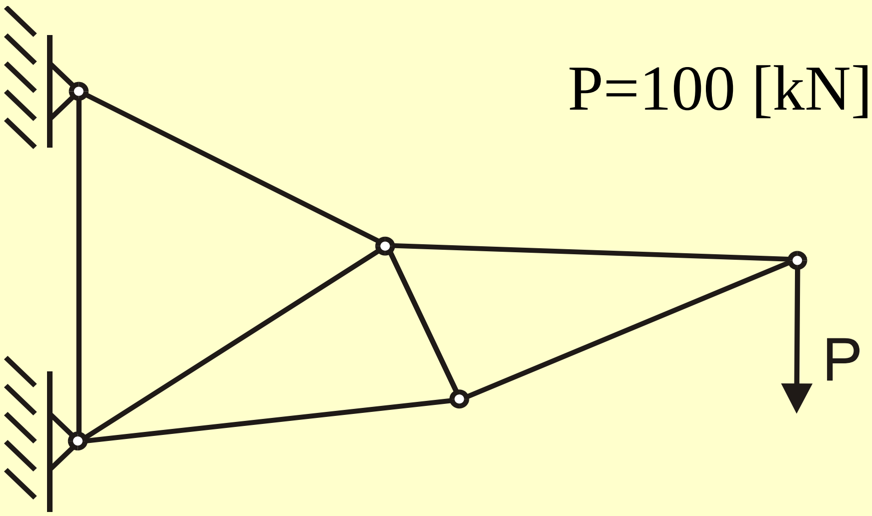
Shape optimisation of truss



Initial shape

Final shape





Objective function

$$f(\mathbf{x}) = \sum_{e=1}^{Ne} A_e L_e = \frac{1}{\sigma_0} \sum_{e=1}^{Ne} \left| N_e \right| \cdot L_e$$

Constraints

$$\sigma_0 - \sigma_e \geq 0 \quad \text{(stress)}$$

$$\left| N_e \right| \leq P_{st} \quad \text{(stability)}$$

Conclusions

The algorithm guarantees that all stationary global solutions (in the initial interval) have been found.

The bounds on the solution(s) are guaranteed to be correct.

Error from all sources are accounted for.

The algorithm can solve the global optimization problem when the objective function is nondifferentiable or even not continuous.

Conclusions

In order to check monotonicity
of the objective function sensitivity analysis
be can applied.

If we apply the Taylor expansion then the
calculation are more efficient but we loss
guaranteed accuracy of the results.

Presented algorithm is more efficient
than method based on
the natural interval extension.