LARGE SCALE SECOND-ORDER OPTIMIZATION

Alvaro F. M. Azevedo

Faculty of Engineering
Department of Civil Engineering
UNIVERSITY OF PORTO
PORTUGAL

STRUCTURAL OPTIMIZATION

- What is that?
  - Design of Civil Engineering structures:
    - Buildings
    - Bridges
    - Dams
  - Design of Mechanical Engineering devices:
    - Automobiles
    - Air planes

STRUCTURAL OPTIMIZATION

- Find the least cost solution the satisfies all the requirements
  - Structure must be safe - unlikely collapse
  - Structure must have quality – stiffness, comfort
  - These requirements must prevail for an extended period of time - durability
  - Structural behavior is governed by the laws of nature – equilibrium, compatibility, relationship between forces and displacements

OPTIMIZATION ALGORITHMS

- Genetic algorithms
  - Derivative free
  - Robust in global optimization
  - Can be easily parallelized
  - Inefficient when the number of variables is high

OPTIMIZATION ALGORITHMS

- First order methods
  - Structural analysis / Sensitivity analysis / Redesign
  - First order sensitivity analysis
  - Adequate for a moderate number of design variables

OPTIMIZATION ALGORITHMS

- Second order method presented here
  - Integrated formulation (all the variables are present)
  - First and second derivatives are symbolically determined (all the functions must be explicitly available)
  - Adequate for problems with a large number of design variables
  - Penalized by the presence of a large number of behavior variables

NONLINEAR PROGRAMMING

$$\text{Minimize } f(\underset{\sim}{x})$$

subject to

$$\underset{\sim}{g}(\underset{\sim}{x}) \leq \underset{\sim}{0} \quad \rightarrow \quad g_i(\underset{\sim}{x}) + s_i^2 = 0$$

$$\underset{\sim}{h}(\underset{\sim}{x}) = \underset{\sim}{0}$$

- Variables / functions ⇨ real and continuous
- All the functions are generalized polynomials, such as:

$$f(\underset{\sim}{x}) = 2.1 x_1^2 x_4^{-3} - 3.1 x_2 + 2.7 x_1^{-1} x_3 x_5^2 - 1.8$$

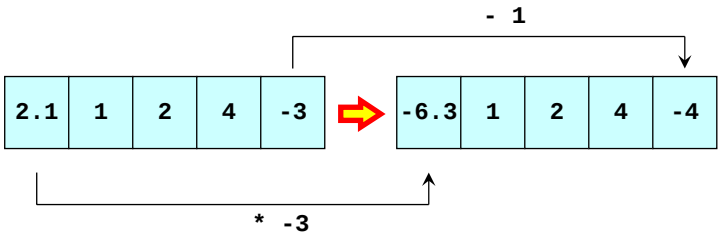
GENERALIZED POLYNOMIALS

$$f(\underset{\sim}{x}) = 2.1 x_1^2 x_4^{-3} - 3.1 x_2 + 2.7 x_1^{-1} x_3 x_5^2 - 1.8$$

- A symbolic manipulation is performed
- Expression parsing and evaluation is simplified
- Exact first and second derivatives can be easily calculated
- All these operations can be efficiently performed

DERIVATIVE EXAMPLE

$$f(x_1,\ldots,x_4) = 2.1\, x_1^2\, x_4^{-3} \quad \rightarrow \quad \frac{\partial f}{\partial x_4} = -6.3\, x_1^2\, x_4^{-4}$$



INPUT FILE

```
### Main title of the nonlinear program
      Symmetric truss with two load cases (kN,cm)
Min.
    +565.685 * t5 ^ 2 + 100 * t8 ^2 ; # truss volume (cm3)

s.t.i.c.
    Min. area 4:    - t4 ^ 2 + 0.15 < 0 ;

s.t.e.c.
    Equil 16:  + 141.421 * t5 ^ 2 * disp16 - 100 = 0 ;

END_OF_FILE
```

- All the software is coded in ANSI C

DENOMINATOR SUBSTITUTION

$$\frac{x^3 - 5.7}{9.8\,y^2 - z} - 3.1 \le 0 \quad \Leftrightarrow \quad \begin{cases} -D + 9.8\,y^2 - z = 0 \\ x^3\,D^{-1} - 5.7\,D^{-1} - 3.1 \le 0 \end{cases}$$

$$D = 9.8\,y^2 - z \qquad \text{(D is a new variable)}$$

LAGRANGIAN

$$L\left(\underset{\sim}{X}\right) = f\left(\underset{\sim}{x}\right) + \sum_{k=1}^{m} \lambda_k^g \left[g_k\left(\underset{\sim}{x}\right) + s_k^2 \right] + \sum_{k=1}^{p} \lambda_k^h h_k\left(\underset{\sim}{x}\right)$$

VARIABLES

$$\underset{\sim}{X} = \left(\underset{\sim}{s}, \underset{\sim}{\lambda}^g, \underset{\sim}{x}, \underset{\sim}{\lambda}^h\right)$$

SOLUTION

- Stationary point of the Lagrangian

SYSTEM OF NONLINEAR EQUATIONS

$$\nabla L\left(\underset{\sim}{X}\right) = \underset{\sim}{0} \quad \Rightarrow \quad \begin{cases} 2 s_i \lambda_i^g = 0 & (i = 1,\ldots,m) \\ g_i + s_i^2 = 0 & (i = 1,\ldots,m) \\ \dfrac{\partial f}{\partial x_i} + \displaystyle\sum_{k=1}^{m} \lambda_k^g \dfrac{\partial g_k}{\partial x_i} + \sum_{k=1}^{p} \lambda_k^h \dfrac{\partial h_k}{\partial x_i} = 0 & (i = 1,\ldots,n) \\ h_i = 0 & (i = 1,\ldots,p) \end{cases}$$

- The solution of the system is a KKT solution when

$$\underset{\sim}{\lambda}^g \geq 0$$

LAGRANGE-NEWTON METHOD

- The system of nonlinear equations

$$\nabla L\,(\underset{\sim}{X}) = \underset{\sim}{0}$$

is solved by the Newton method

- In each iteration the following system of linear equations has to be solved

$$H\left(\underset{\sim}{X}^{q-1}\right) \Delta \underset{\sim}{X}^{q} + \nabla L\left(\underset{\sim}{X}^{q-1}\right) = \underset{\sim}{0}$$

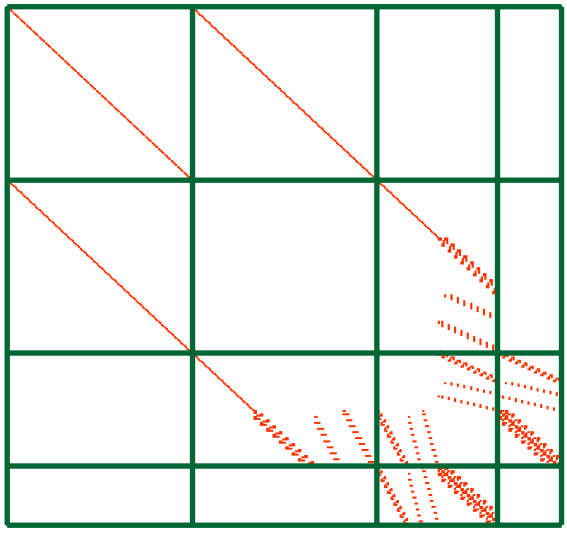
HESSIAN MATRIX

$$\underset{\sim}{H} = \begin{array}{c|cccc} & (m) & (m) & (n) & (p) \\ \hline (m) & Diag\left(2\lambda_i^g\right) & Diag\left(2s_i\right) & \underset{\sim}{0} & \underset{\sim}{0} \\ (m) & & \underset{\sim}{0} & \dfrac{\partial g_i}{\partial x_j} & \underset{\sim}{0} \\ (n) & & & \bullet & \dfrac{\partial h_j}{\partial x_i} \\ (p) & SYMMETRIC & & & \underset{\sim}{0} \end{array}$$

$$\bullet \quad \frac{\partial^2 f}{\partial x_i \partial x_j} + \sum_{k=1}^{m} \lambda_k^g \frac{\partial^2 g_k}{\partial x_i \partial x_j} + \sum_{k=1}^{p} \lambda_k^h \frac{\partial^2 h_k}{\partial x_i \partial x_j}$$

HESSIAN MATRIX SPARSITY PATTERN

$$\underset{\sim}{H} =$$



SYSTEM OF LINEAR EQUATIONS

- Gaussian elimination
  - adapted to the sparsity pattern of the Hessian matrix
- Conjugate gradients
  - diagonal preconditioning
  - adapted to an indefinite Hessian matrix

$$\underset{\sim}{H}^T \underset{\sim}{H} \Delta \underset{\sim}{X} + \underset{\sim}{H}^T \nabla L = \underset{\sim}{0}$$

  - DRAWBACK: condition number increases significantly

LINE SEARCH

$$\underset{\sim}{X}^q = \underset{\sim}{X}^{q-1} + \alpha \, \Delta \underset{\sim}{X}^q$$

- When the value of α minimizes the error in $\Delta \underset{\sim}{X}^q$ direction
  - the value of α is often close to one
  - faster convergence
  - process may fail
- When the value of α is made considerably smaller (e.g. $\alpha = 0.1$)
  - stable convergence
  - more iterations - slower

NEWTOP COMPUTER CODE

- Elementary equality constraints are substituted:

$$x_i = c\,x_j \qquad \text{or} \qquad x_i = c$$

- All the variables are scaled
- Objective function and constraints are normalized
- The NLP is simplified ⇨ easier to solve
- Problems with a large number of variables can be solved

  (e.g., 4 000 design variables and 20 000 constraints)

SCALING

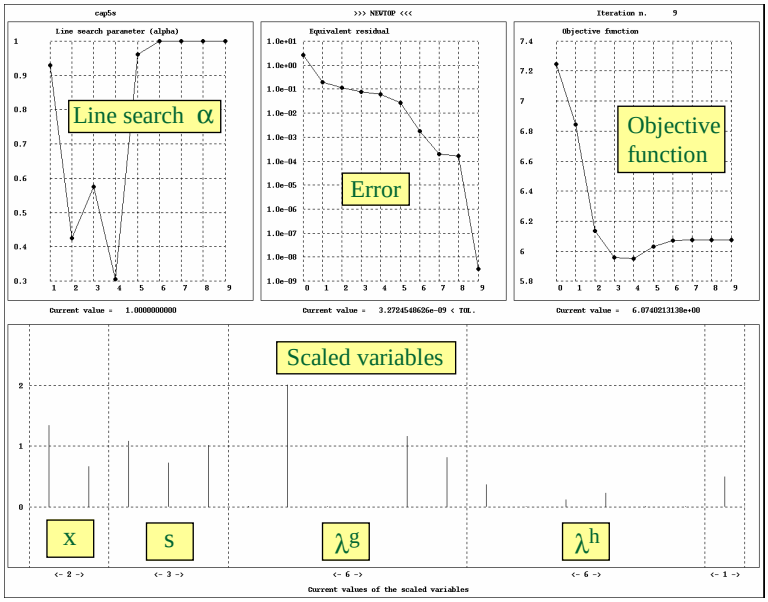
Initial solution:
$\left(500, 0.1, \sqrt{300}, \sqrt{0.1}\right)$

- Variable substitution: $x_i = Z_i\ \bar{x}_i$
- Functions f, g_i and h_i are multiplied by factors such that, for the initial solution, all the gradient norms = 1

$$\begin{aligned} &\textit{Min.}\ 2000\, x_1 \\ &\textit{subject to} \\ &-x_1 + 200 + x_3^2 = 0 \\ &x_2 - 0.2 + x_4^2 = 0 \\ &-10\, x_1\, x_2 + 500 = 0 \end{aligned}$$

⇨

$$\begin{aligned} &\textit{Min.}\ \bar{x}_1 \\ &\textit{subject to} \\ &-0.640\,\bar{x}_1 + 0.256 + 0.384\,\bar{x}_3^{\,2} = 0 \\ &0.447\,\bar{x}_2 - 0.894 + 0.447\,\bar{x}_4^{\,2} = 0 \\ &-0.707\,\bar{x}_1\,\bar{x}_2 + 0.707 = 0 \end{aligned}$$



TRUSS OPTIMIZATION

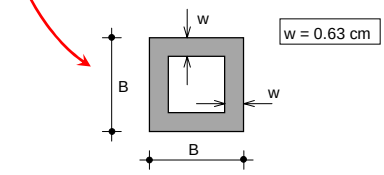
- Cost minimization (often similar to volume minimization)

- Sizing ⇨ cross-sectional areas may change
- Shape optimization ⇨ nodal coordinates may change

Simultaneously

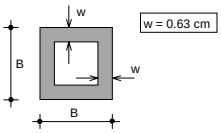
VARIABLES

- Integrated formulation
- Design variables and behavior variables simultaneously present in the nonlinear program
  - Cross-section dimensions (e.g., width, diameter, area)
  - Some nodal coordinates
  - Nodal displacements



SUBSTITUTED VARIABLES

- In most cases the area (A) and the moment of inertia (I) depend on a single parameter (B)



$$A = C_0^A + C_1^A B + C_2^A B^2$$

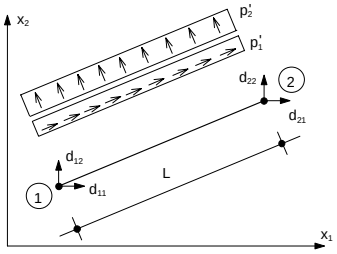
$$I = C_0^I + C_1^I B + C_2^I B^2 + C_3^I B^3 + C_4^I B^4$$

  - Coefficients C_i^A and C_j^I are fixed
  - Variables A and I can be substituted in all the functions that define the mathematical program

ADDITIONAL VARIABLES

$$k_{ij} = \cdots + EAL^{-1} + \cdots$$

$$L = \sqrt{\left(x_{21} - x_{11}\right)^2 + \left(x_{22} - x_{12}\right)^2}$$



♦ Additional variables ⇨ L_i

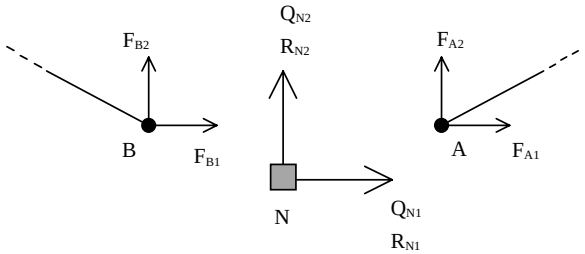
$$-L^2 + x_{11}^2 + x_{12}^2 + x_{21}^2 + x_{22}^2 - 2x_{11}x_{21} - 2x_{12}x_{22} = 0$$

♦ Additional equality constraints ⇨ L_i definition

EQUILIBRIUM EQUATIONS

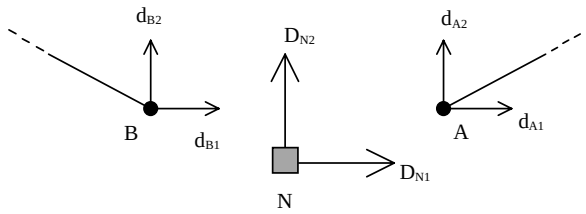
• Equality constraints:

$$\underset{\sim}{F}_A + \cdots + \underset{\sim}{F}_B + \cdots = \underset{\sim}{Q} + \underset{\sim}{R}$$



• Reactions are only present in constrained dof 's

COMPATIBILITY EQUATIONS



$$\underset{\sim}{d}_A = \underset{\sim}{D}_N$$
$$\underset{\sim}{d}_B = \underset{\sim}{D}_N$$

- Variables d are substituted
- D_{Ni} is fixed in constrained dof 's

NON LINEAR PROGRAM

- Objective function: cost ⇨ $f(\underset{\sim}{x}) = \sum_{i=1}^{NB} c_i A_i L_i$
- Equality constraints:
  - for each bar with variable length:
    - ➔ one equation defining L
  - for each non-prescribed degree of freedom:
    - ➔ one equilibrium equation

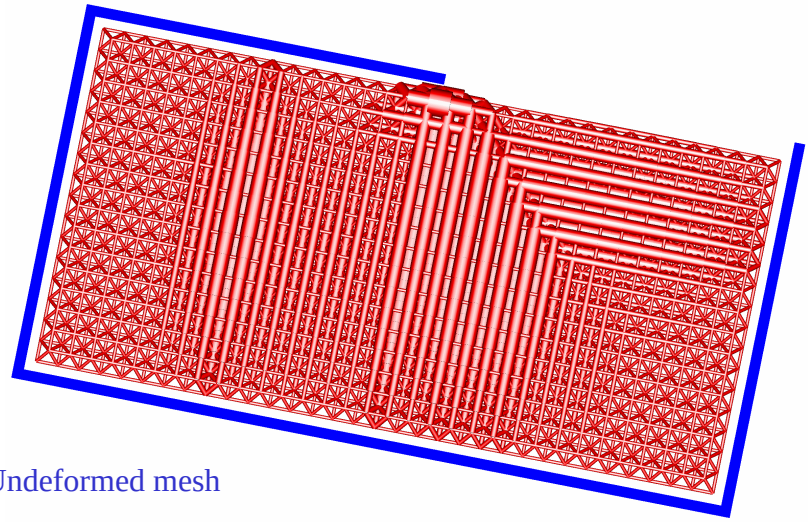
NON LINEAR PROGRAM

- Inequality constraints:
  - minimum width ⇨ $B \geq B_{min}$
  - allowable stress (tension and compression)
  - local Euler buckling
  - side constraints in nodal coordinates ⇨ $x_{min} \leq x_i \leq x_{max}$
  - side constraints in some displacements ⇨ $D_{min} \leq D_{N_i} \leq D_{max}$

LARGE SCALE OPTIMIZATION PROBLEM

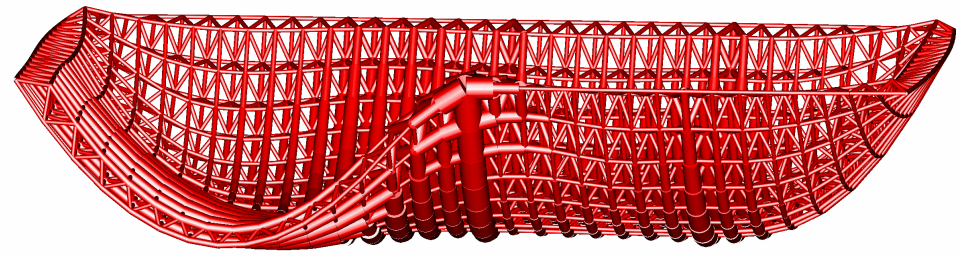
- 3D truss sizing
- Number of bars = 4 096
- Number of degrees of freedom = 3 135
- Number of decision variables = 7 231
- Number of inequality constraints = 19 038
- No variable linking
- No active set strategy

BUILDING ROOF - OPTIMAL SOLUTION



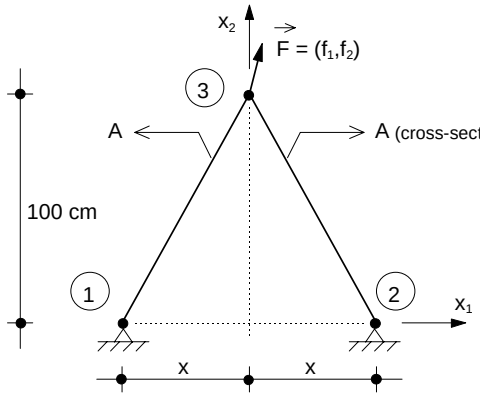
Undeformed mesh

BUILDING ROOF - OPTIMAL SOLUTION



Deformed mesh

SHAPE OPTIMIZATION TEST PROBLEM



$$\| \vec{F} \| = 200\ kN$$

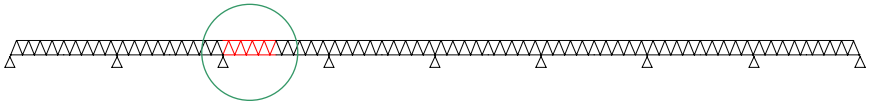
$$f_2 = 8\ f_1$$

$$\sigma_{max} = 100\ kN/cm^2$$

- Variables: A , x
- Svanberg's solution confirmed

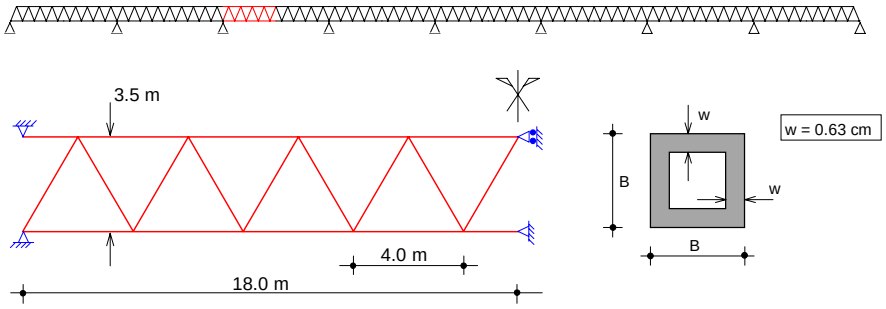
SHAPE OPTIMIZATION PROBLEM

- Minimize the cost of a steel bridge
- Member sizing and shape optimization
- Linear elastic structural behavior
- Fixed nodes (normal direction)
- Local Euler buckling
- Portuguese structural codes



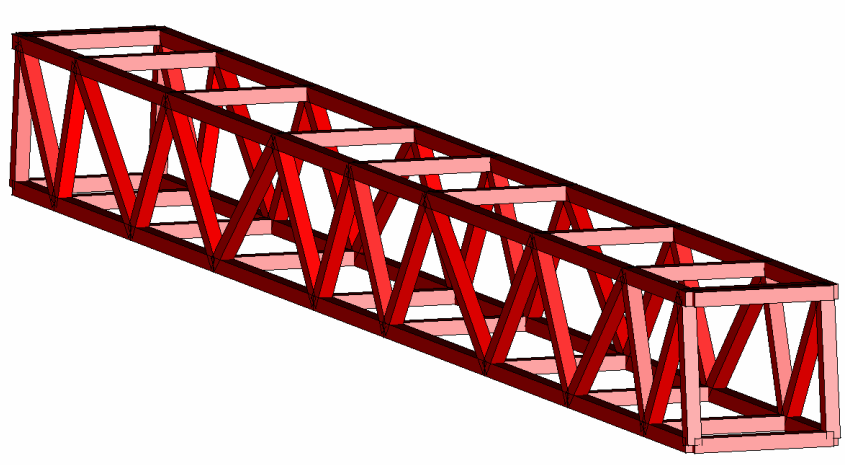
STEEL BRIDGE

Vertical distributed load

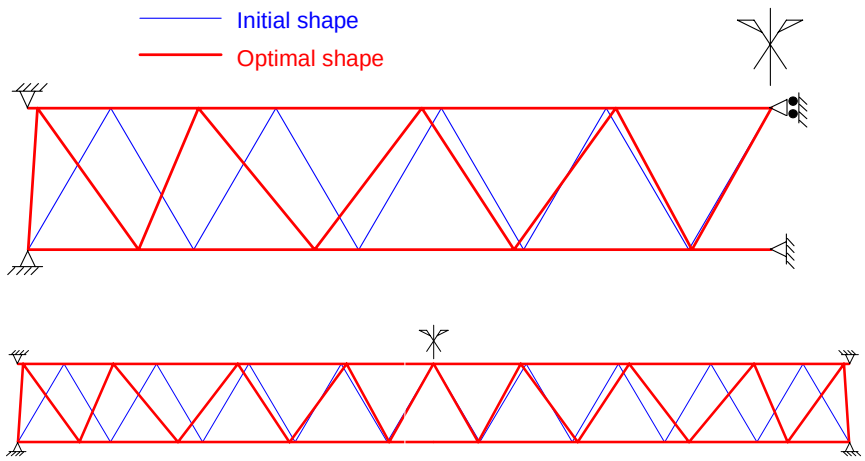


- Group I - horizontal bars
- Group II - diagonal bars

STEEL BRIDGE



OPTIMAL SHAPE

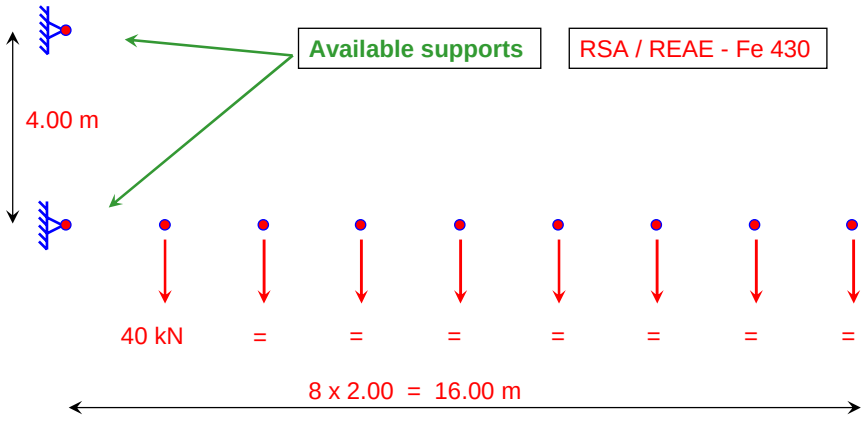


NUMERICAL RESULTS

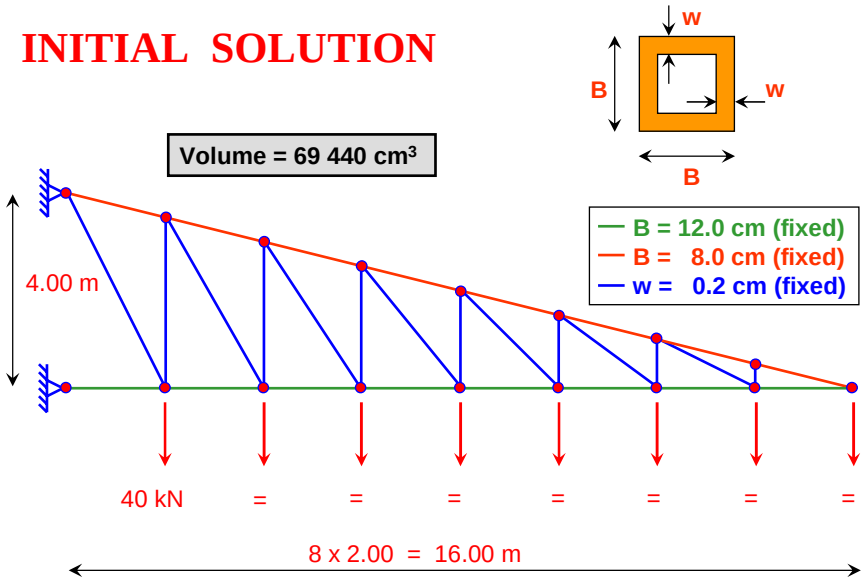
- Optimal solution - sizing only
  - Volume = 170 dm^3
- Optimal solution - sizing and shape optimization
  - Volume = 146 dm^3 (14 % smaller)
  - CPU time (PC): less than 10 seconds

PROBLEM

- Create a structure to hold 8 loads of 40 kN each

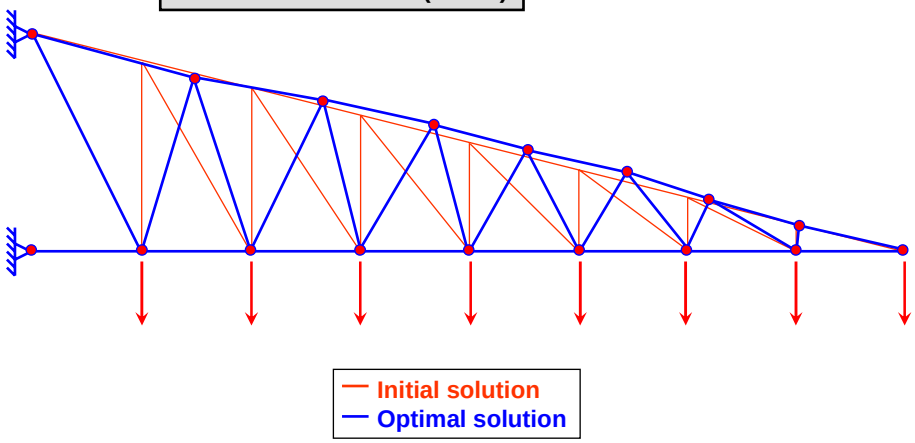


INITIAL SOLUTION



OPTIMAL SOLUTION

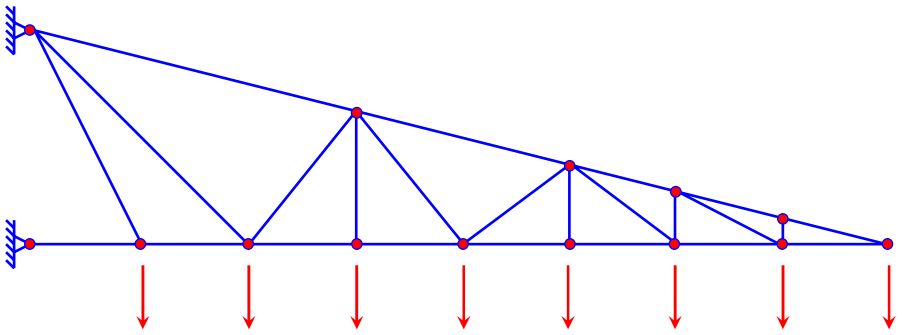
Volume = 66 674 cm³ (- 4 %)



— Initial solution
— Optimal solution

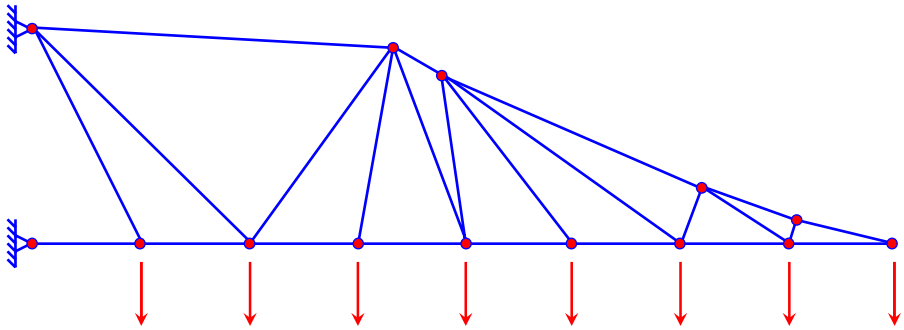
NEW INITIAL SOLUTION

- Same problem
- Distinct topology

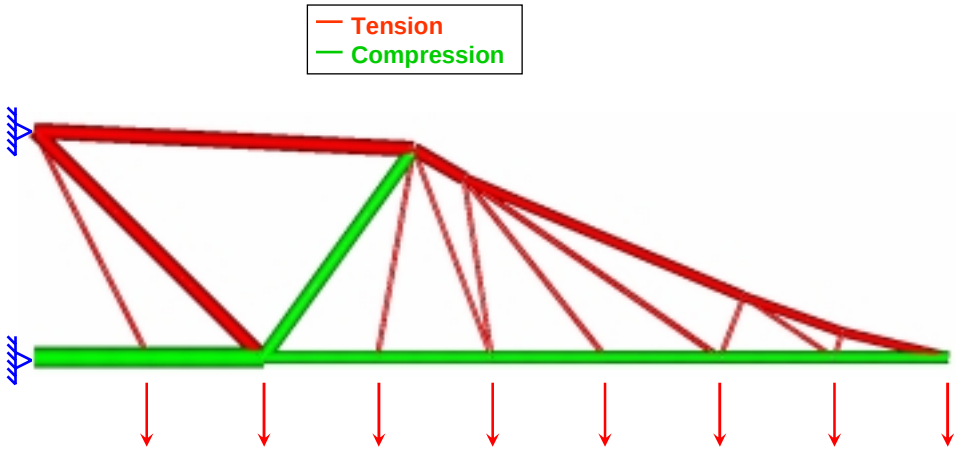


OPTIMAL SOLUTION

Volume = 58 934 cm^3 (- 15 %)



OPTIMAL SOLUTION



CONCLUSIONS

- ↑ • Applicable to large scale optimization problems
- ↑ • Very accurate and efficient
- ↑ • Can be used in realistic truss optimization problems
- ↓ • A large number of behavior variables and/or load cases reduces efficiency
- ↓ • Friendly user interface is still required

INTERNET URLs

- My home page:
  - http://www.fe.up.pt/~alvaro
- My list of papers (in English):
  - http://civil.fe.up.pt/alvaro/public_eng.htm
- Selected papers about this subject:
  - http://civil.fe.up.pt/alvaro/pdf/1995_OPTI_Paper_Sec_Order_Struct_Opt.pdf
  - http://civil.fe.up.pt/alvaro/pdf/1997_OPTI_Paper_Opt_Frames_Nonlin.pdf
  - http://civil.fe.up.pt/alvaro/pdf/1999_OPTI_Paper_Shap_Opt_Steel_Bridge.pdf
  - http://civil.fe.up.pt/alvaro/pdf/2001_OPTI_Paper_OO_Sec_Order_Opt.pdf