

## STRUCTURAL SHAPE OPTIMIZATION USING GRADIENT-BASED METHOD

Espath, Luis Felipe da Rosa<sup>a</sup>, Linn, Renato Vaz<sup>b</sup> and Awruch, Armando Miguel<sup>c</sup>

<sup>a</sup>Graduate Student PPGEC/UFRGS [espath@gmail.com](mailto:espath@gmail.com)

<sup>b</sup>Undergraduate Student PPGEC/UFRGS [renatolinn@gmail.com](mailto:renatolinn@gmail.com)

<sup>c</sup>Graduate Program in Civil Engineering (PPGEC), Federal University of Rio Grande do Sul (UFRGS),  
Av. Osvaldo Aranha, 99, 3o andar, 90035-190, Porto Alegre, RS, Brazil, [amawruch@ufrgs.br](mailto:amawruch@ufrgs.br),  
<http://www6.ufrgs.br/engcivil/>

**Keywords:** Mathematical Optimization, Automatic Differentiation, NURBS, Finite Element Method, Structural Analysis.

**Abstract.** The consolidation of the link among four fields in computational mathematics and mechanics is the main objective of this work. Surfaces based on non-uniform rational B-spline (NURBS), mathematical optimization, the finite element method in structural analysis and automatic differentiation are applied to shape optimization of shells. This problem is performed taking into account the fact that the material mechanical characteristics influence both, the structural shape and the thickness variation, in order to obtain the best performance with respect to a specific criterium. These two variables, shape and thickness variation, have an essential role considering that the minimum material cost, a specific frequency and a pure membrane stress state are typical design objectives. Suitable shapes and thickness variation are intrinsic concepts of shell optimization. Therefore, some techniques were implemented to modify the shell geometry conserving the same parameterization without a new finite element mesh generation. The shape modification is taken by an optimization code and it is based in the data obtained by a finite element analysis and gradients evaluation. Some examples are analyzed and discussed. As a consequence of the shape optimization, shells with high structural performance and esthetically beautiful shapes can be obtained.

## 1 INTRODUCTION

The design of structures with optimal behavior with respect to a specific criterium has been constantly investigated and improved. In ancient times, designs were improved by ingenious engineers through generations. The various books and huge number of papers published recently reflect the increasing interest in this area. The development of the structural optimization can be traced back to Galileo Galilei (1564-1642) and to his clamped beam problem (Galilei, 1638). The cantilever beam subjected to constant shear and bending as formulated by Galileo was an optimal project of minimal weight with an uniform stress constraint. Galileo's problem was probably one of the first structural optimization problems. Joseph Louis Lagrange (1736-1813) had also formulated an optimal design of a column for over 200 years. However, most of the modern structural optimization studies are based on efficient and reliable structural analysis, which became possible only after the advent of modern computers and the finite element method in 1960's. The development of mathematical programming techniques and theories has laid a solid foundation for structural optimization as well.

Sensitivity analysis can be performed by finite differences in most of the problems. However it needs excessive function evaluations (expensive) and it is not precise. Analytical sensitivities is the opposite. It is exact and cheap but an explicit representation is difficult to obtain. The automatic differentiation method is an efficient alternative way to perform the sensitivity analysis, being cheaper and exact. Thus, in the present implementation, the automatic differentiation tool TAPENADE AD, developed by INRIA (2002), is used.

The parameterization description is well suitable for shape modification instead of an explicit or implicit representation. A Bézier patch is a good alternative for shape description and modification and it was used by Ramm et al. (1993). The solid description is the most general way to define the thickness and a rational Bézier solid description was used by Kegl and Brank (2006) for geometry description and for shell thickness definition.

Shell structures are a very wide type of structure with intensive use, for example, in the design of airplane's and submarine's fuselage, metallic silos, hangar roofs, building structures, aerospace and automotive components, among others. The large spans, low weight, and stiffness are some advantage proprieties obtained with this structural type. The shape and thickness distribution play a dominate role in the shell structure response. The best structural performance can be obtained with material and structural optimization.

The structural optimization system implemented in the present work is composed by a shape modification tool based on NURBS description, a gradient-based sequential quadratic programming using automatic reverse differentiation for the sensitivity analysis and a finite element code for shell analysis (using the Discrete Kirchhoff Triangle combined with the Constant Stress Triangle).

A smooth and precise geometry representation is important when the shape of the structure has to be found and modified. A NURBS parameterization provides a robust description of the surface representing accurately complex shapes with an efficient mathematical implementation. For shape optimization purpose, NURBS are well suitable due to the easy control of the surface shape by control points in homogeneous coordinates and knot vectors, providing almost any geometry representation. More details about NURBS may be found in Piegl and Tiller (1997).

The mathematical optimization method has to be chosen based on the specific case analyzed. Optimization procedures can easy handle some types of problems, but they can be unsatisfying or just can not handle other optimization problems. For shape optimization of shell structures, an important point that needs to be noticed is that the constraints used are, in general, highly

nonlinear. In this way, a SQP algorithm was chosen and implemented here. A penalty method would be prohibitive in this case. The SQP is a robust algorithm for nonlinear continuous optimization and the DONLP2 algorithm (Spellucci, 2001) is used in this work. The algorithm needs function and constraints evaluations, as well as its Fréchet derivatives, in order to guide its unidimensional search procedure to find the optimal solution iteratively.

For the structural analysis, a finite element code is implemented here using a 3-nodes triangular flat shell element. A suitable mesh discretization is needed for the optimization algorithm. Mesh distortions, poor mesh quality and other errors can yield bad numerical values, leading the optimization algorithm to non-reliable results.

In Ramm et al. (1993) was analyzed the imperfections effect according to the first mode of linear buckling.

## 2 THEORETICAL ASPECTS

A set of statements referred to the geometrical procedures, the derivatives evaluation, the mathematical programming for the numerical optimization and the structural analysis are presented.

The geometrical procedures are based on NURBS description with a shape modification ruled by control points, weight and knot vectors. The Fréchet derivatives are evaluated using finite differences (FD) and automatic differentiation (AD) in the reverse mode. The numerical optimization is performed by the Sequential Quadratic Programming (SQP) algorithm. The structural analysis of shells is implemented using the Discrete Kirchhoff triangle (DKT) for plate bending combined with Constant Stress Triangle (CST) for membrane effects.

### 2.1 NON-UNIFORM RATIONAL B-SPLINE (NURBS)

In order to represent a complex surface, a parametric representation is used. NURBS parameterization are well suitable for shape optimization in any physical problem involving curves, surfaces and solids. For shell structures, a surface representation in the form

$$\mathbf{S}(u, v) = (x(u, v), y(u, v), z(u, v)) \quad (u, v) \in [0, 1] \times [0, 1]$$

is looked for.

The NURBS surface in homogeneous coordinates is defined as

$$\mathbf{S}^w(u, v) = \sum_{i=0}^n \sum_{j=0}^m N_{i,p}(u) N_{j,q}(v) \mathbf{P}_{i,j}^w \quad (1)$$

where  $(p, q)$ ,  $(n, m)$ ,  $(N_{i,p}(u), N_{j,q}(v))$  are the degrees, the numbers of basis functions, and the basis functions in  $(u, v)$  directions, respectively.  $\mathbf{P}_{i,j}^w = (w_{i,j}x_{i,j}, w_{i,j}y_{i,j}, w_{i,j}z_{i,j}, w_{i,j})$  are the control points in homogeneous coordinates. The computational algorithms in homogeneous coordinates are very efficient (Piegl and Tiller, 1997).

The basis functions in recursive form are defined as

$$N_{i,0}(u) = \begin{cases} 1 & \text{if } u_i \leq u \leq u_{i+1} \\ 0 & \text{otherwise} \end{cases}$$

$$N_{i,p}(u) = \frac{u - u_i}{u_{i+p} - u_i} N_{i,p-1}(u) + \frac{u_{i+p+1} - u}{u_{i+p+1} - u_{i+1}} N_{i+1,p-1}(u)$$

over the following knot vectors, in  $(u, v)$  directions

$$\mathcal{U} = \left\{ \underbrace{0, \dots, 0}_{p+1}, u_{p+1}, \dots, u_{r-p-1}, \underbrace{1, \dots, 1}_{p+1} \right\}$$

$$\mathcal{V} = \left\{ \underbrace{0, \dots, 0}_{q+1}, v_{q+1}, \dots, v_{s-q-1}, \underbrace{1, \dots, 1}_{q+1} \right\}$$

In the shape modification when a control point, a weight and/or a knot position is modified the parametric mesh remain the same; however, the mesh in Euclidean space assume the new position according with the new geometry. This mapping may be not the best choice, because the mesh distortion is not controlled. Although, this problem is easy to get around with a convenient initial mesh.

In order to obtain the best result in an optimization process, the control points in homogeneous coordinates  $\mathbf{P}^w$  and the knot vectors components  $(u_i, v_i)$  are used as optimization variables in this work. Note that the control points  $\mathbf{P}_{i,j}$  and the weights  $w_{i,j}$  are included in  $\mathbf{P}_{i,j}^w$ .

## 2.2 FIRST-ORDER SENSITIVITY DERIVATIVES (FO SDs)

In order to obtain the first-order sensitivity derivatives the finite difference method and automatic differentiation technique was implemented. In one of the examples presented here finite differences was used to compare with automatic differentiation (see section 3.4) and all examples were solved with automatic differentiation.

The finite difference method has been used in three forms: forward, backward and centered differences. In this work forward difference were used. Therefore:

$$\frac{\partial \mathcal{H}(X)}{\partial X_i} \approx \frac{\mathcal{H}(X_i + \delta X_i) - \mathcal{H}(X_i)}{\delta X_i} \quad (2)$$

where  $\mathcal{H}$  is the objective function,  $X$  is the independent variable vector,  $\delta X_i$  is a small variation of  $i$ -th independent variable and the index  $i$  vary from 1 to the number of variables,  $n$ .

In general, values of the perturbation  $\delta X_i$  varies between  $10^{-5}$  and  $10^{-8}$ . In this work  $10^{-5}$  was adopted. The computational time in a gradient evaluation is equivalent to  $n + 1$  evaluations of the objective function.

Regarding the automatic differentiation, a few statements are presented in order to explain this approach. It is easy to understand a numerical algorithm as a composition of functions, where each step of the algorithm is a simple function. The composition of those functions must be differentiated in order to get a new algorithm. The differentiated algorithm must properly evaluate the derivatives of the output variables with respect to the input ones. Using the differentiation of compositions of functions theorem, i.e., the chain rule, the differentiated algorithm results in a multiple product of matrices, where each product belong to a particular step.

Let  $\mathcal{H}$  be a function, defined by a numerical algorithm.

$$\begin{aligned} \mathcal{H} : \mathbb{R}^n &\rightarrow \mathbb{R}^m \\ X &\rightarrow Y \end{aligned} \quad (3)$$

where  $X \in \mathbb{R}^n$  is the domain in a  $n$ -dimensional real space and  $Y \in \mathbb{R}^m$  is the co-domain in a  $m$ -dimensional real space. Unlike the whole algorithm (as a one-unit single structure), it is

possible to find an explicit representation for each step of the numerical algorithm. This way, it is decomposed into  $k \in \mathbb{N}$  steps, where  $\mathbb{N}$  is the set of natural numbers.

$$\begin{aligned} \mathcal{H}^l : \mathbb{R}^{n_{l-1}} &\rightarrow \mathbb{R}^{n_l} & 1 \leq l \leq k \\ Z^{l-1} &\rightarrow Z^l \end{aligned} \quad (4)$$

where  $Z^l$  are intermediate variable, being  $Z^0 \stackrel{\text{def}}{=} X$  and  $Z^k \stackrel{\text{def}}{=} Y$ , where  $\stackrel{\text{def}}{=}$  is the statement definition.

Using the composition operator,  $\odot$ , the composition  $\mathcal{H}$  is stated as

$$\mathcal{H}(X) = \bigodot_{l=1}^k \mathcal{H}^l(X) \stackrel{\text{def}}{=} (\mathcal{H}^k \circ \dots \circ \mathcal{H}^1)(X) \quad (5)$$

being  $\mathcal{H}^l$  a differentiable function. It is possible to evaluate its Jacobian at the point  $X_0$  as

$$A_{ij}(X_0) \stackrel{\text{def}}{=} \left. \frac{\partial \mathcal{H}_i(X)}{\partial X_j} \right|_{X=X_0} \quad 1 \leq i \leq m \quad \wedge \quad 1 \leq j \leq n \quad (6)$$

Applying the chain rule, the following equation is found

$$A(X_0) = \left. \frac{\partial \mathcal{H}^k}{\partial Z^{k-1}} \right|_{Z^{k-1} = \bigodot_{l=1}^{k-1} \mathcal{H}^l(X_0)} \cdot \dots \cdot \frac{\partial \mathcal{H}^1}{\partial Z^0} \quad (7)$$

The differentiation strategies are based on the associative matrix multiplication. Computationally, the two most common strategies that arise in this context are: forward multiplication and backward multiplication. According to [Giering and Kaminski \(1996\)](#) more efficient strategies may arise when the Jacobian sparsity is taken into account.

In the forward mode, the matrices multiplication are evaluated in the same order as the composition, i.e.,

$$\left( \frac{\partial \mathcal{H}^k}{\partial Z^{k-1}} \cdot \left( \dots \cdot \left( \frac{\partial \mathcal{H}^3}{\partial Z^2} \cdot \left( \frac{\partial \mathcal{H}^2}{\partial Z^1} \cdot \frac{\partial \mathcal{H}^1}{\partial Z^0} \right) \right) \right) \right) \quad (8)$$

and all intermediate results have  $n$  columns. In reverse mode, the matrices multiplication are evaluated backwards (in the reverse order), i.e.,

$$\left( \left( \left( \left( \frac{\partial \mathcal{H}^k}{\partial Z^{k-1}} \cdot \frac{\partial \mathcal{H}^{k-1}}{\partial Z^{k-2}} \right) \cdot \frac{\partial \mathcal{H}^{k-2}}{\partial Z^{k-3}} \right) \cdot \dots \right) \cdot \frac{\partial \mathcal{H}^1}{\partial Z^0} \right) \quad (9)$$

and all intermediate results have  $m$  rows.

In numerical optimization, generally,  $n > m$  and  $m = 1$ , and the reverse mode is preferable because it needs less numerical computation. Although, for each case, the computational time must be analyzed taking into account the memory requirement.

Taking the following definition for an intermediate result in a scalar-valued functional

$$Z_0^l \stackrel{\text{def}}{=} \bigodot_{i=1}^l \mathcal{H}^i(X_0) \quad 1 \leq l \leq k \quad (10)$$

the intermediate variable variation is function of the previous intermediate variable variation, and with  $\delta Z^0 \stackrel{\text{def}}{=} \delta X$ , the proposition is stated as

$$\delta Z^l = \left. \frac{\partial \mathcal{H}^l (Z^{l-1})}{\partial Z^{l-1}} \right|_{Z^{l-1}=Z_0^{l-1}} \cdot \delta Z^{l-1} \quad (11)$$

The adjoint operator defined over an appropriate inner product,  $(\cdot, \cdot)$ , is stated in linear algebra as  $(v, T(u)) = (T^*(v), u)$ , where  $u, v \in \mathcal{V}$  space and  $T^*$  is the adjoint of transformation  $T : \mathcal{V} \rightarrow \mathcal{V}$ . For the Euclidean inner product the adjoint operator is the transposed matrix.

The adjoint of an intermediate result is stated as the gradient of  $\mathcal{H}$  with respect to the intermediate result

$$\delta^* Z^l \stackrel{\text{def}}{=} \nabla_{Z^l} \left( \bigodot_{i=l+1}^k \mathcal{H}^i (Z^l) \right) \Big|_{Z^l=Z_0^l} \quad (12)$$

Denoting the inner product in an Euclidean space by  $\langle \cdot, \cdot \rangle$ , and with the gradient definition stated as  $\delta \mathcal{H} = (\nabla_X \mathcal{H} (X_0), \delta X)$ , it is obtained

$$\delta \mathcal{H} = \langle \delta^* Z^l, \delta Z^l \rangle \quad (13)$$

It is natural that eq. (13) holds for any  $l$  and

$$\begin{aligned} \langle \delta^* Z^{l-1}, \delta Z^{l-1} \rangle &= \langle \delta^* Z^l, \delta Z^l \rangle \\ &= \left\langle \delta^* Z^l, \left( \frac{\partial \mathcal{H}^l (Z^{l-1})}{\partial Z^{l-1}} \right) \Big|_{Z^{l-1}=Z_0^{l-1}} \cdot \delta Z^{l-1} \right\rangle \\ &= \left\langle \left( \frac{\partial \mathcal{H}^l (Z^{l-1})}{\partial Z^{l-1}} \right)^* \Big|_{Z^{l-1}=Z_0^{l-1}} \cdot \delta^* Z^l, \delta Z^{l-1} \right\rangle \\ \delta^* Z^{l-1} &= \left( \frac{\partial \mathcal{H}^l (Z^{l-1})}{\partial Z^{l-1}} \right)^* \Big|_{Z^{l-1}=Z_0^{l-1}} \cdot \delta^* Z^l \end{aligned} \quad (14)$$

According to the adjoint operator definition as a transpose matrix, the eq. (14) can be written as

$$\delta^* Z_i^{l-1} = \sum_{j=1}^{n_l} \frac{\partial \mathcal{H}_j^l (Z^{l-1})}{\partial Z_i^{l-1}} \Big|_{Z^{l-1}=Z_0^{l-1}} \cdot \delta^* Z_j^l \quad (15)$$

and this eq. (15) exemplifies a simple step of adjoint algorithm.

The gradient of  $\mathcal{H}$  is evaluated in the last step of the adjoint algorithm, then

$$\delta^* Z^0 = \delta^* X = \nabla_X \mathcal{H} \quad (16)$$

The complete recipes for adjoint code construction are presented by [Giering and Kaminski \(1996\)](#). Another way to get the same result can be obtained by oriented graph theory ([Griewank and Walther, 2008](#)).

## 2.3 NUMERICAL OPTIMIZATION

In the context of the numerical shape optimization, the SQP method is used. The optimization problem is stated as

$$\min_{X \in \mathbb{R}^n} \mathcal{H}(X) \quad \text{subject to} \quad \begin{cases} C_i(\mathbf{x}) = 0, & i \in \mathcal{E} \\ C_i(\mathbf{x}) \geq 0, & i \in \mathcal{I} \end{cases} \quad (17)$$

where  $\mathcal{H}$ ,  $C_i$  are defined in  $\mathbb{R}^n$ ; and  $\mathcal{E}$  and  $\mathcal{I}$  are two finite index sets.  $X$  are the independent variables,  $\mathcal{H}$  is the objective function,  $C_i, i \in \mathcal{E}$  are equality constraints and  $C_i, i \in \mathcal{I}$  are inequality constraints.

The Lagrangian is defined as

$$\mathcal{L}(X, \lambda) = \mathcal{H}(X) - \sum_{i \in \mathcal{E} \cup \mathcal{I}} \lambda_i C_i(\mathbf{x}) \quad (18)$$

where  $\lambda_i$  are the Lagrange multipliers. If the linear independence constraint qualification holds, then the optimum  $(X^*, \lambda^*)$  must satisfy the *Karush–Kuhn–Tucker (KKT)* condition, i.e.,

$$\nabla_X \mathcal{L}(X^*, \lambda^*) = \mathbf{0}, \quad (19a)$$

$$C_i(X^*) = 0, \quad \forall i \in \mathcal{E}, \quad (19b)$$

$$C_i(X^*) \geq 0, \quad \forall i \in \mathcal{I}, \quad (19c)$$

$$\lambda_i^* \geq 0, \quad \forall i \in \mathcal{I}, \quad (19d)$$

$$\lambda_i^* C_i(X^*) = 0, \quad \forall i \in \mathcal{E} \cup \mathcal{I} \quad (19e)$$

The proof and the complete theory related is shown by [Nocedal and Wright \(1999\)](#). The *KKT* conditions, eqs. (19a) to (19e), must be satisfied for a given tolerance range.

## 2.4 STRUCTURAL ANALYSIS

The structural analysis is performed using a triangular flat shell element with 18 degrees of freedom combining the Discrete Kirchhoff Triangle (DKT) and the Constant Stress Triangle (CST). This element was developed by [Bathe and Ho \(1981\)](#). Geometrically non linear problems are solved using the generalized displacement control method ([Yang and Shieh, 1990](#)).

The thickness variation implemented is quadratic and the variation is orthogonal to the axis of symmetry, being constant in each element. There is not thickness variation in parallel directions to the axis of symmetry.

## 3 SHAPE OPTIMIZATION

The main objective is to find an optimum shape with respect to a specified criterium. The function  $\mathcal{H}$  to be minimized here is: internal strain energy in linear elastic problem.

$$\mathcal{H} = \int_{\Omega} \underline{\underline{\sigma}} : \underline{\underline{\epsilon}} d\Omega = U^T K U \quad (20)$$

$$(21)$$

where  $\Omega$  is the domain,  $\underline{\underline{\sigma}}$  is the stress tensor,  $\underline{\underline{\epsilon}}$  is the strain tensor,  $U$  is the displacement vector and  $K$  is the stiffness matrix. The optimization problem is performed using relative objective functions values, i.e., referred to an initial value.

This minimization of  $\mathcal{H}$  is constrained by geometrical or mechanical functions, such as

$$\mathcal{C} = 1 - \frac{V}{V_{ini}} = 0 \quad (22)$$

$$\mathcal{C} = 1 - \frac{\|u_{idof}\|_2}{\|u_p\|_2} \geq 0 \quad (23)$$

where  $V$  is the current volume,  $V_{ini}$  is the initial volume,  $u_{idof}$  is the displacement at a selected node and  $u_p$  is the maximum allowed value.

These functions are evaluated by FEM, and the gradients by AD.

### 3.1 NUMERICAL ALGORITHM

The general algorithm (fig. 1) is a union between FEM, AD, SQP and NURBS. Basically the computation tool has three types of input data: the optimization problem description, the finite element model and the surface description using NURBS.

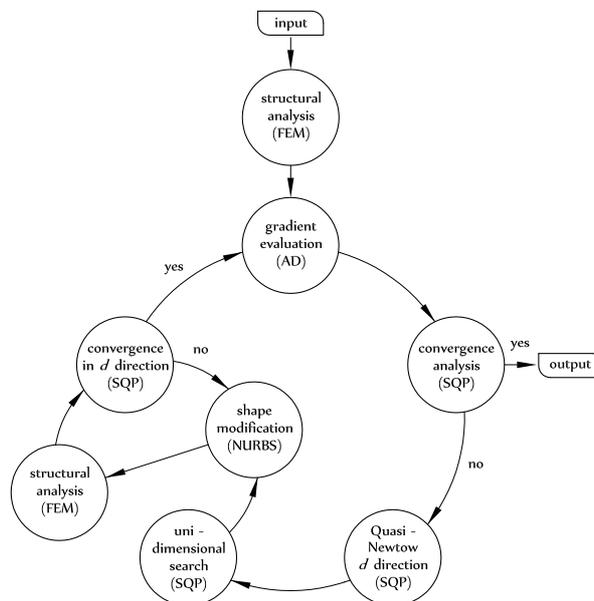


Figure 1: General algorithm.

After the initial structural analysis on the initial shape is performed, the main loop begins. At the beginning of each loop the gradients are evaluated. Then, *KKT* conditions are verified. When the *KKT* conditions are satisfied the main loop finishes. The quadratic problem is solved and the Quasi-Newton direction  $d$  is obtained. On  $d$  direction the uni-dimensional minimization is performed. The geometry is modified along  $d$  and the functions involved are evaluated. The uni-dimensional minimization is not exact. The loop converges on the  $d$  direction when a reasonable point is obtained. A reasonable point is referred to the fact that the value of the objective function decreases and this point is more feasible than the previous one, assuming that constraints are satisfied.

### 3.2 APPLICATIONS

The elastic material properties used are given by: the Young's modulus  $E$ , the Poisson coefficient  $\nu$ , the specific weight  $\gamma$  and the thickness  $e$ . The system units is the SI.

If the initial geometry is plane and the load is orthogonal to this plane, then the membrane internal strain energy is avoided. Thus, a perturbation is inserted in the center of the geometry to avoid this critical point. A central control point is displaced  $1 \cdot 10^{-5}$  in the normal direction.

### 3.3 EXAMPLE 1: CODE VALIDATION

This example is based on the half cylinder and an analytic solution is well-know for this optimization problem. For this reason it may be taken as a code validation case. The circular arch has a radius of 1 and it is hinged supported at the extremes with a load shown in fig. 2. The material properties are:  $E = 2.1 \cdot 10^{11}$ ,  $\nu = 0.3$ ,  $e = 1 \cdot 10^{-2}$ . This problem is performed

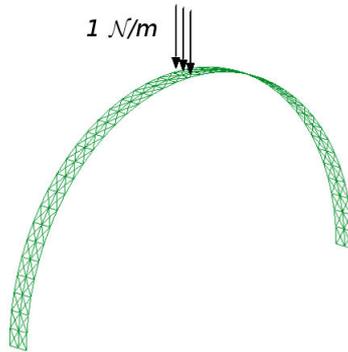


Figure 2: Initial mesh.

with two variables, the  $(xw, zw)$  of the blue control point shown in fig. 3a, and the objective is to minimize the relative internal strain energy. The geometrical constraints are: the geometry must hold the initial symmetry and the planar projection.

The control points  $\mathbf{P}^w$  in the initial geometry are given in tab. 1. The basis functions are

$\mathbf{P}^w$	$xw$	$yw$	$zw$	$w$		$\mathbf{P}^w$	$xw$	$yw$	$zw$	$w$
(0,0)	-1	-0.05	0	1		(0,1)	-1	0.05	0	1
(1,0)	$\frac{-1}{\sqrt{2}}$	$\frac{-0.05}{\sqrt{2}}$	$\frac{-1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}$		(1,1)	$\frac{-1}{\sqrt{2}}$	$\frac{0.05}{\sqrt{2}}$	$\frac{-1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}$
(2,0)	0	-0.05	1	1		(2,1)	0	0.05	1	1
(3,0)	$\frac{1}{\sqrt{2}}$	$\frac{-0.05}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}$		(3,1)	$\frac{1}{\sqrt{2}}$	$\frac{0.05}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}$	$\frac{1}{\sqrt{2}}$
(4,0)	1	-0.05	0	1		(4,1)	1	0.05	0	1

Table 1: Control points  $\mathbf{P}^w$  in the initial geometry.

defined over the knot vectors

$$\mathcal{U} = \{0, 0, 0, 0.5, 0.5, 1, 1, 1\}$$

$$\mathcal{V} = \{0, 0, 1, 1\}$$

In the optimization process, the inital functional value is  $\mathcal{H}_{ini} = 1$  and, at the end, the functional value is  $\mathcal{H}_{opt} = 6.43 \cdot 10^{-4}$ .

The solution obtained in this problem is  $xw = -0.265$  and  $zw = -0.443$  for  $\mathbf{P}^w(1, 0)$ , shown in fig. 3b. The analytical solution of this problem is  $x + z = -1 \forall (x, z) \in [-1, 0] \times [-1, 0]$  for any  $w \neq 0$ . Naturally, this proof the non uniqueness of solution, where the solution is obtained for any  $(x, z)$  and  $w$  that satisfies the previous equation.

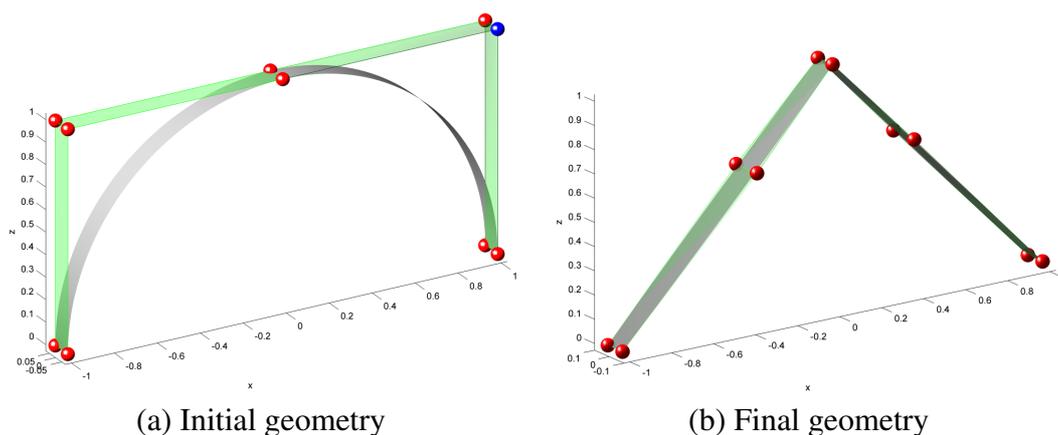


Figure 3: Initial and final geometry.

This example was solved using two different solution techniques for linear systems to check the derivative convergence. These methods are the classical Gauss elimination and an element by element conjugate gradient method (which is an iterative method). As a result it was checked that the derivative convergence rate is equal or faster than the functional convergence rate when the iterative method is used. This iterative method is used in the reverse mode in order to decrease the memory requirement, once this issue is a drawback in the reverse mode of AD.

The order of difference between the optimal values found for the gradient evaluation using the classical Gauss elimination method and an element by element conjugate gradient method was  $10^{-5}$ .

This simple example has a significant role. Beyond to validate the numerical code, it shows that is possible to evaluate the functions using the classical Gauss elimination method and the gradients using a conjugate gradient iterative method in the optimization process. This result allows to optimize complex geometries that needs a refined mesh. It is also a way to use automatic differentiation in complex problems without memory limitations. The classical Gauss elimination method was adopted to evaluate the functions and the conjugate gradient iterative method was adopted to evaluate the gradients in the examples of the next sections.

### 3.4 EXAMPLE 2: COMPARISON BETWEEN FD AND AD

The objective of this example is to compare the results in the optimization process using finite differences (FD) and automatic differentiation (AD).

This example is based on  $10 \times 10$  square plate, simply supported at the corners. The initial volume is 11.27. The load is composed by its own weight and an uniform load ( $q = 5.5 \cdot 10^3$ ) applied over the plate surface. The material properties are:  $E = 3.0 \cdot 10^{10}$ ,  $\nu = 0.2$  e  $\gamma = 2.5 \cdot 10^4$ . The initial mesh is shown in fig. 4a.

This problem is performed with 9 control points with variable positions in the normal direction. The variables are the blue control points shown in fig. 4b. The thickness variables are 3 parameters to define the quadratic thickness variation. The objective is to minimize the relative strain energy. The geometrical constraints are: the geometry must hold the initial symmetry, the planar projection and a constant volume. The mechanical constraints are: the displacement of the center of the plate and the center of the edges must be less than 0.05. The control point are bounded by  $-15$  to  $0$  in  $zw$  direction. The thicknesses are bounded by 0.03 to 0.45.

The control points  $\mathbf{P}^w$  in the initial geometry are given in tab. 2. The basis functions are

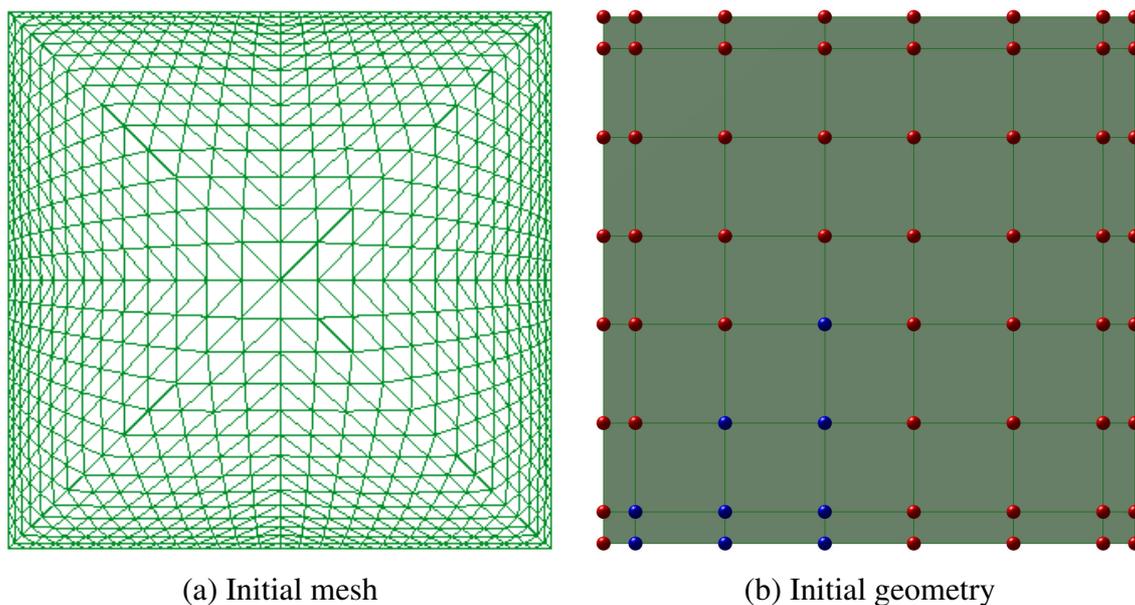


Figure 4: Initial mesh and geometry.

$\mathbf{P}^w$	$xw$	$yw$	$zw$	$w$	$\mathbf{P}^w$	$xw$	$yw$	$zw$	$w$
(0,0)	-5	-5	0	1	(1,1)	-4.4	-4.4	0	1
(1,0)	-4.4	-5	0	1	(2,1)	-2.7125	-4.4	0	1
(2,0)	-2.7125	-5	0	1	(3,1)	-0.8375	-4.4	0	1
(3,0)	-0.8375	-5	0	1	(2,2)	-2.7125	-2.7125	0	1
(3,3)	-0.8375	-0.8375	0	1	(3,2)	-0.8375	-2.7125	0	1

Table 2: Control points  $\mathbf{P}^w$  in the initial geometry.

defined over the knot vectors

$$\mathcal{U} = \{0, 0, 0, 0, 0, 0, 0.25, 0.75, 1, 1, 1, 1, 1, 1\}$$

$$\mathcal{V} = \{0, 0, 0, 0, 0, 0, 0.25, 0.75, 1, 1, 1, 1, 1, 1\}$$

In the optimization process, the initial functional value is  $\mathcal{H}_{ini} = 1$  and, at the end, the functional value is  $\mathcal{H}_{opt.FD} = 5.81 \cdot 10^{-4}$  for the approach using FD and  $\mathcal{H}_{opt.AD} = 5.50 \cdot 10^{-4}$  for the approach using AD.

The optimal control points are given in tab. 3. The optimal thickness variation is shown in

FD	$\mathbf{P}^w$	$zw$	$\mathbf{P}^w$	$zw$	-	AD	$\mathbf{P}^w$	$zw$	$\mathbf{P}^w$	$zw$
	(0,0)	0	(1,1)	-1.30	-		(0,0)	0	(1,1)	-0.97
	(1,0)	-0.74	(2,1)	-3.18	-		(1,0)	-0.64	(2,1)	-3.05
	(2,0)	-1.41	(3,1)	-3.94	-		(2,0)	-2.59	(3,1)	-3.88
	(3,0)	-5.30	(2,2)	-2.92	-		(3,0)	-4.88	(2,2)	-3.35
	(3,3)	-4.94	(3,2)	-5.27	-		(3,3)	-5.02	(3,2)	-5.23

Table 3: Control points  $\mathbf{P}^w$  in the final geometry for FD and AD.

fig. 5, for an approach with FD (fig. 5a) and for an approach with AD (fig. 5b). The optimal

geometry using FD is shown in fig. 6a, while the optimal geometry using AD is presented in fig. 6b. Results of displacements for the optimal shapes are shown in fig. 7.

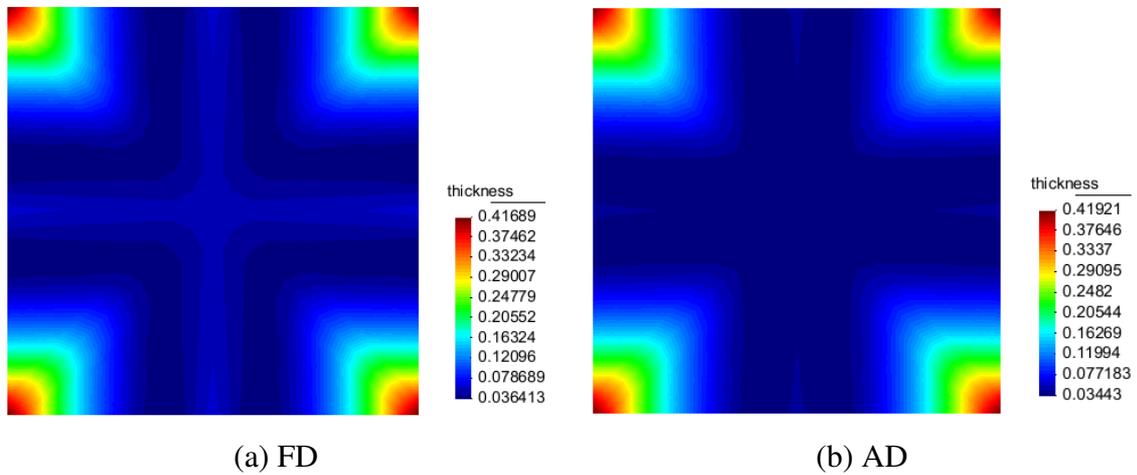


Figure 5: FD vs AD (Thickness).

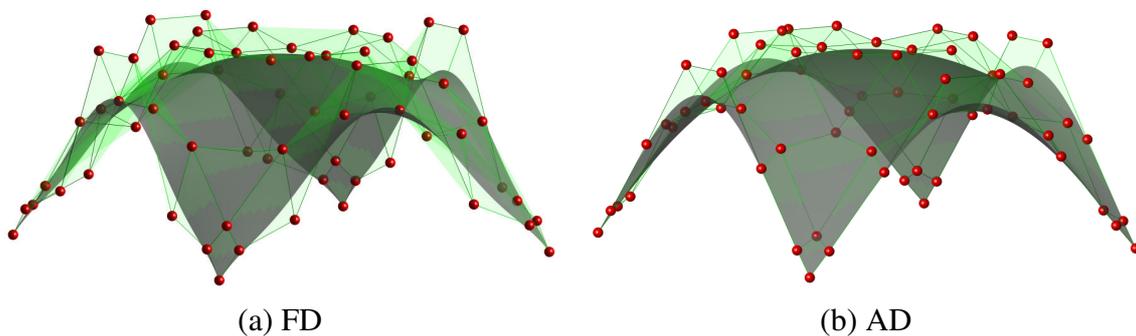


Figure 6: FD vs AD (Geometry).

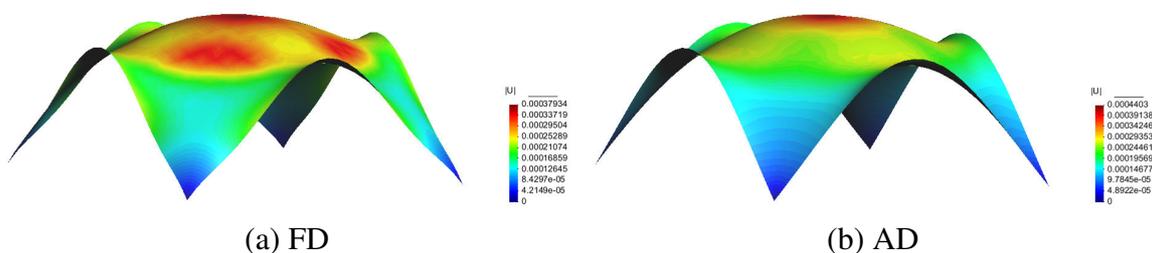


Figure 7: FD vs AD (Displacement analysis).

It can be observed, the difference between the gradients evaluated by FD and AD are small. In this example, the maximum difference observed was 7% and the minimal was 1.5%. However, the maximum difference in the control point position was 83.7%. The errors intrinsic to the gradients do not only affect the search direction; they may also affect the optimal point.

Probably the estimated gradients (obtained with FD) evaluated at the real optimal point do not satisfy the *KKT* conditions, and leading to another point.

In order to achieve better results with the least possible computational cost, only AD will be used in the following examples.

### 3.5 EXAMPLE 3: LOAD INFLUENCE

In this section two problems are analyzed separately. These two cases are the same example presented in section 3.4, but with two different loads. In the first case only the weight load is applied and in second case only uniform load is applied, minimizing the same objective function subjected to the same constraints. The basic idea is to show the load influence on the optimal shape.

In the optimization process, the initial functional value is  $\mathcal{H}_{ini} = 1$  and, at the end, the functional value is  $\mathcal{H}_{opt.w} = 2.88 \cdot 10^{-4}$  when the weight load is considered and  $\mathcal{H}_{opt.u} = 4.84 \cdot 10^{-4}$  when the uniform load is considered.

The optimal control points are given in tab. 4. The optimal thickness variation is shown in

weight	$\mathbf{P}^w$	$zW$		$\mathbf{P}^w$	$zW$	-	uniform	$\mathbf{P}^w$	$zW$		$\mathbf{P}^w$	$zW$
	(0,0)	0		(1,1)	-0.24	-		(0,0)	0		(1,1)	-0.55
	(1,0)	-0.72		(2,1)	-3.42	-		(1,0)	-0.50		(2,1)	-2.59
	(2,0)	-5.06		(3,1)	-2.32	-		(2,0)	-3.36		(3,1)	-4.35
	(3,0)	-4.17		(2,2)	-3.53	-		(3,0)	-3.83		(2,2)	-2.71
	(3,3)	-3.82		(3,2)	-4.57	-		(3,3)	-4.90		(3,2)	-4.74

Table 4: Control points  $\mathbf{P}^w$  in the final geometry weight and uniform load.

fig. 8, for the weight load (fig. 8a) and for the uniform load (fig. 8b). The optimal geometry for the weight load is shown in fig. 9a, while the optimal geometry for the uniform load is presented in fig. 9b. Results of displacements for the optimal shapes are shown in fig. 10.

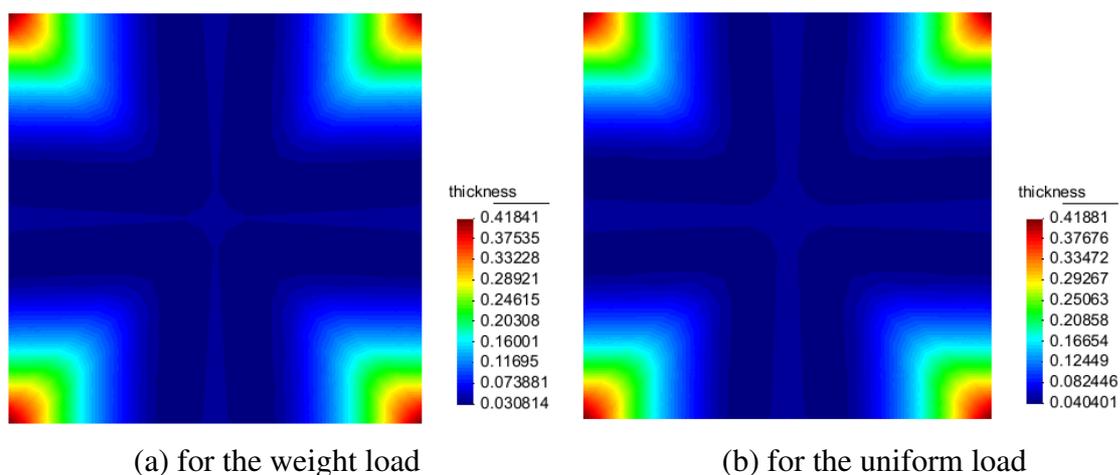


Figure 8: weight load vs uniform load (Thickness).

In this case, the same example shown in section 3.4 was analyzed, but two different loads were considered separately, while in the example of section 3.4 the weight load and an uniform

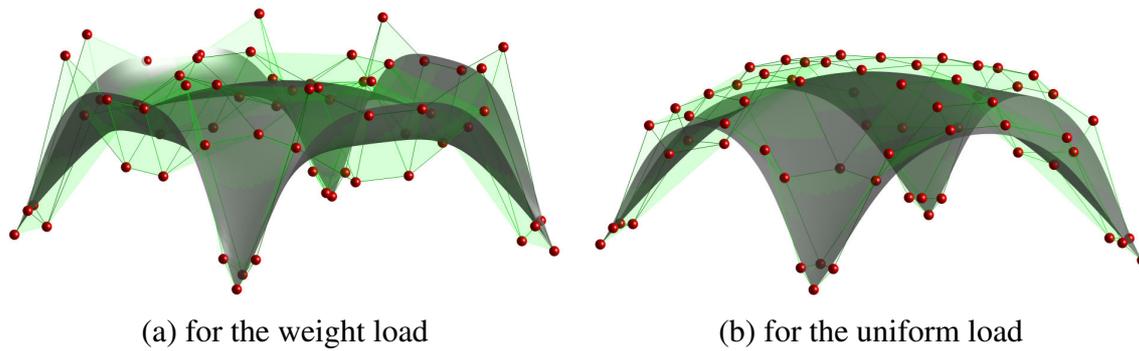


Figure 9: weight load vs uniform load (Geometry).

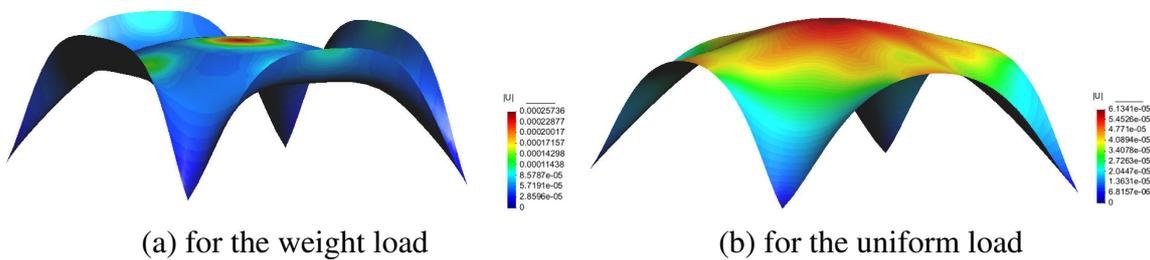


Figure 10: weight load vs uniform load (Displacement analysis).

load were applied simultaneously. Results of the example in section 3.4 and those obtained here for the two cases are very different. Thus, an optimization process must consider all types of load that can occur. For example, if snow accumulation or the wind incident may occur over the structure, they must be properly considered. The problem complexity increases with a variable load. In some cases, when these effects are important, some additional geometric algorithm must be used to handle with the snow accumulation and a fluid-structure interaction algorithm must be used for the wind load.

### 3.6 EXAMPLE 4: WEIGHTS, CONTROL POINTS AND KNOT VECTORS AS VARIABLES OF THE SHAPE OPTIMIZATION PROCEDURE

This example is based on a square plate and is performed for two different cases. The idea is to optimize all parameters describing the surface, i.e., control points, weights and knot vectors. However, the mesh distortion is always a limitant. Thus, this example is divided in two cases. In the first case control points and knot vectors are used as variables. In the second case the control points and the weights are used as variables.

In the first case the problem presented in section 3.4 is solved with an additional variable, i.e. the 6-th knot value (which is the central knot). In the second case the objective is to minimize the relative strain energy. The geometrical constraints are: the geometry must hold the initial symmetry and a constant volume. The material properties are:  $E = 2.1 \cdot 10^{11}$ ,  $\nu = 0.3$ ,  $e = 1 \cdot 10^{-2}$ . The applied load is the weight. The variables are the 9 control points used in section 3.4 and their 9 weights. The initial volume is 4.68. The control points are bounded in  $zw$  direction by  $-15$  to  $0$ . The weights are unbounded. The thicknesses are bounded by  $0.01$  to  $0.12$ .

In the first case the optimal control points are given in tab. 5. In the first case the optimal

$\mathbf{P}^w$	$zw$		$\mathbf{P}^w$	$zw$		$\mathbf{P}^w$	$zw$
(1,0)	-0.71		(1,1)	-0.63		(2,2)	-3.41
(2,0)	-2.28		(2,1)	-3.21		(3,2)	-4.78
(3,0)	-4.60		(3,1)	-3.49		(3,3)	-5.00

Table 5: Control points  $\mathbf{P}^w$  in the final geometry (First case).

knot vectors are

$$\mathcal{U} = \{0, 0, 0, 0, 0, 0, 0.50, 0.50, 1, 1, 1, 1, 1, 1\}$$

$$\mathcal{V} = \{0, 0, 0, 0, 0, 0, 0.50, 0.50, 1, 1, 1, 1, 1, 1\}$$

In the second case the optimal control points are given in tab. 6.

$\mathbf{P}^w$	$zw$	$w$		$\mathbf{P}^w$	$zw$	$w$		$\mathbf{P}^w$	$zw$	$w$
(1,0)	-0.94	0.92		(1,1)	-1.11	1.76		(2,2)	-1.26	1.56
(2,0)	-3.29	0.69		(2,1)	-1.89	0.80		(3,2)	-3.91	2.25
(3,0)	-4.40	0.67		(3,1)	-1.50	0.68		(3,3)	-5.68	1.28

Table 6: Control points  $\mathbf{P}^w$  in the final geometry (Second case).

In the optimization process, the initial functional value is  $\mathcal{H}_{ini} = 1$  and, at the end, the functional value is  $\mathcal{H}_{opt.knot} = 5.36 \cdot 10^{-4}$  for the first case and  $\mathcal{H}_{opt.weight} = 5.71 \cdot 10^{-5}$  for the second case.

In the first case the optimal thickness variation is shown in fig. 11a, while the optimal geometry is presented in fig. 11b. Results of displacements for the optimal shape are shown in fig. 12. In the second case the optimal thickness variation is shown in fig. 13a, while the optimal geometry is presented in fig. 13b. Results of displacements for the optimal shape are shown in fig. 14.

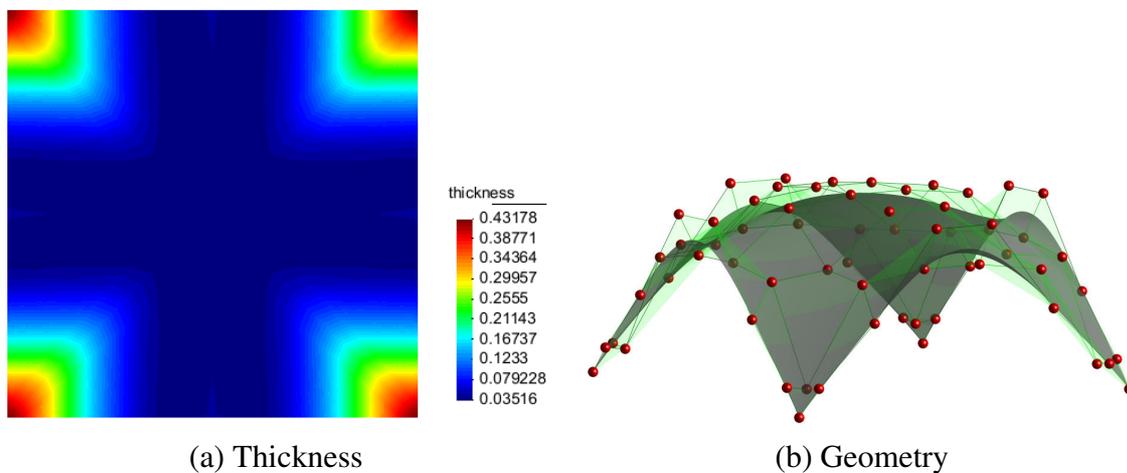


Figure 11: Thickness and geometry (First case).

As the second case a constant volume is required, only a constant total weight load will be applied at each iteration. It is natural that the result does not remain with the same planar

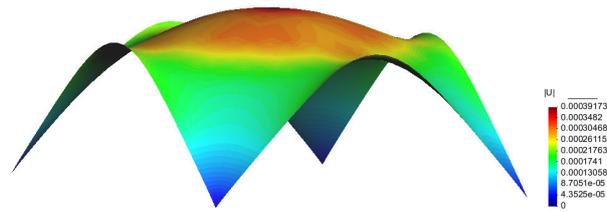
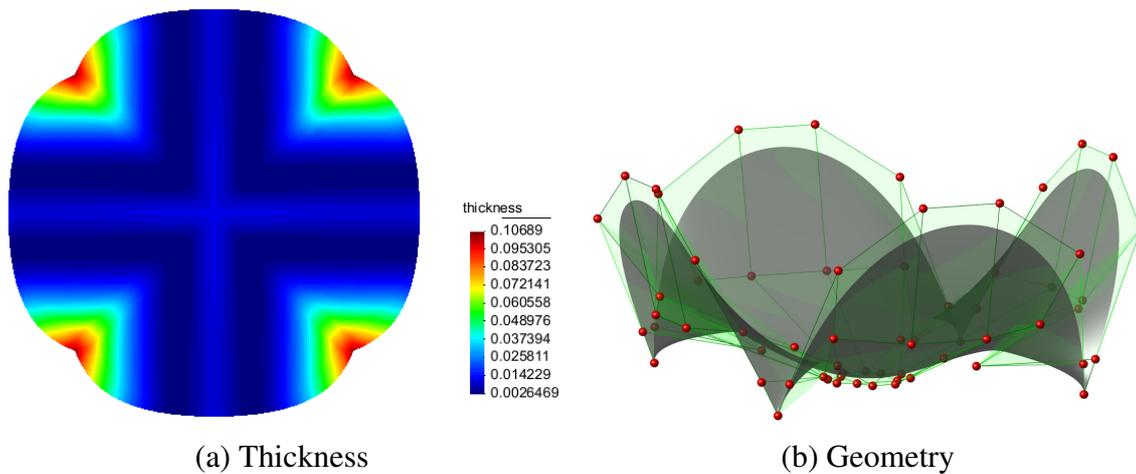


Figure 12: Displacement analysis (First case).

projection at each iteration (because this is not a prescribed constraint), then an uniform load is not used because if the total planar area varies, the total load will also vary and it is not desirable.



(a) Thickness

(b) Geometry

Figure 13: Thickness and geometry (Second case).

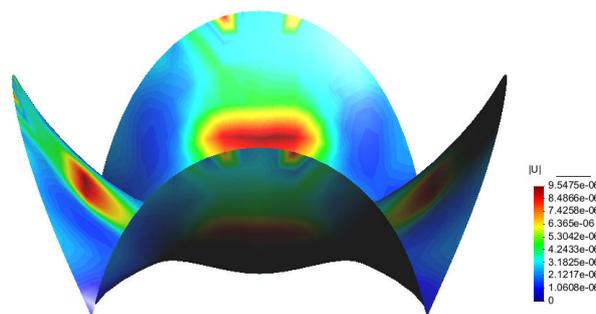


Figure 14: Displacement analysis (Second case).

In these cases the idea is to introduce a new perspective in shape optimization using NURBS. Basically, the aim is to work with all parameters that describe a NURBS surface, i.e., the control points with their weights and the knot position. The number of variables increases, but the structural response will be more adequate.

In the first case of this section, the optimal geometry may be compared with the example presented in section 3.4. The difference is the 6-th knot position (which is the central knot) as an extra variable. The result obtained here is better than that of the example in section 3.4.

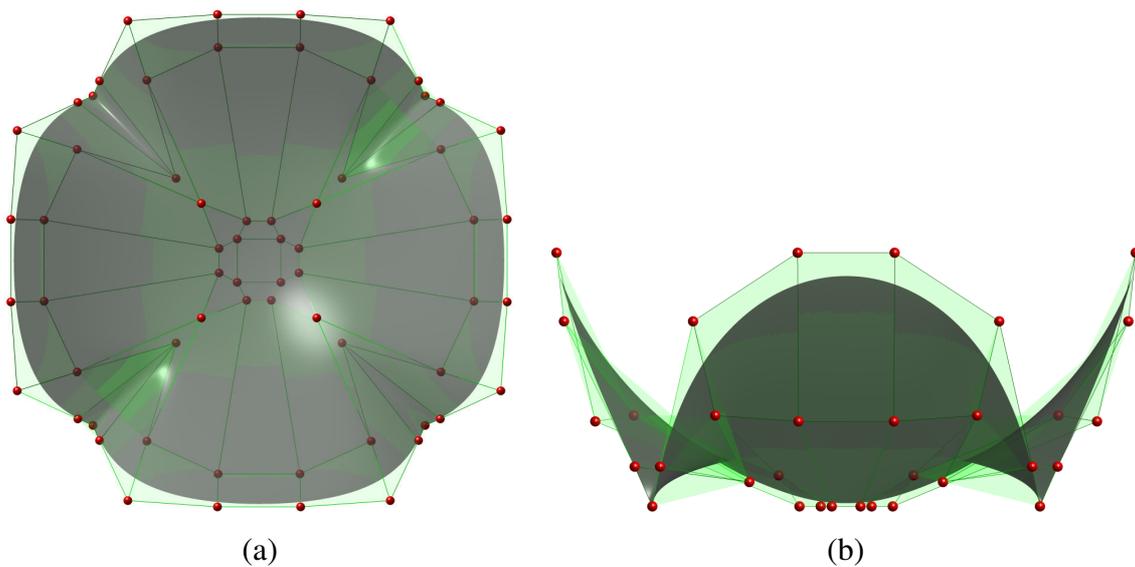


Figure 15: Views of the geometry (Second case).

In the second case of this section, the optimum value of the objective function is the lowest one in this work. The number of variables is twice than those used in section 3.4, but the performance obtained is excellent.

### 3.7 EXAMPLE 5: SHAPE OPTIMIZATION OF SHELLS WITH GEOMETRICALLY NON LINEAR BEHAVIOR

In this example a minimax problem is analyzed. The objective is to minimize the maximum displacement squared as indicated in eq. (24). The material properties used here are the same used in section 3.4. The geometry is the square plate with an initial volume 10. The applied load is increased in 10 times the load used in section 3.4.

$$\min \max \|U\|_2^2 \tag{24}$$

The variables are the 9 control points used in section 3.4. The geometrical constraints are: the geometry must hold the initial symmetry and the planar projection. The control point variables are bounded in  $zw$  direction by  $-15$  to  $0$ .

The optimal control points are given in tab. 7.

$\mathbf{P}^w$	$zw$	$\mathbf{P}^w$	$zw$	$\mathbf{P}^w$	$zw$
(1,0)	-0.61	(1,1)	-1.98	(2,2)	-11.58
(2,0)	-0.76	(2,1)	-5.63	(3,2)	-14.75
(3,0)	-1.00	(3,1)	-5.44	(3,3)	-14.85

Table 7: Control points  $\mathbf{P}^w$  in the final geometry.

In the optimization process, the initial functional value is  $\mathcal{H}_{ini} = 1$  and, at the end, the functional value is  $\mathcal{H}_{opt} = 9.00 \cdot 10^{-4}$ .

The optimal geometry is presented in fig. 16a and results of displacements for the optimal shape are shown in fig. 16b.

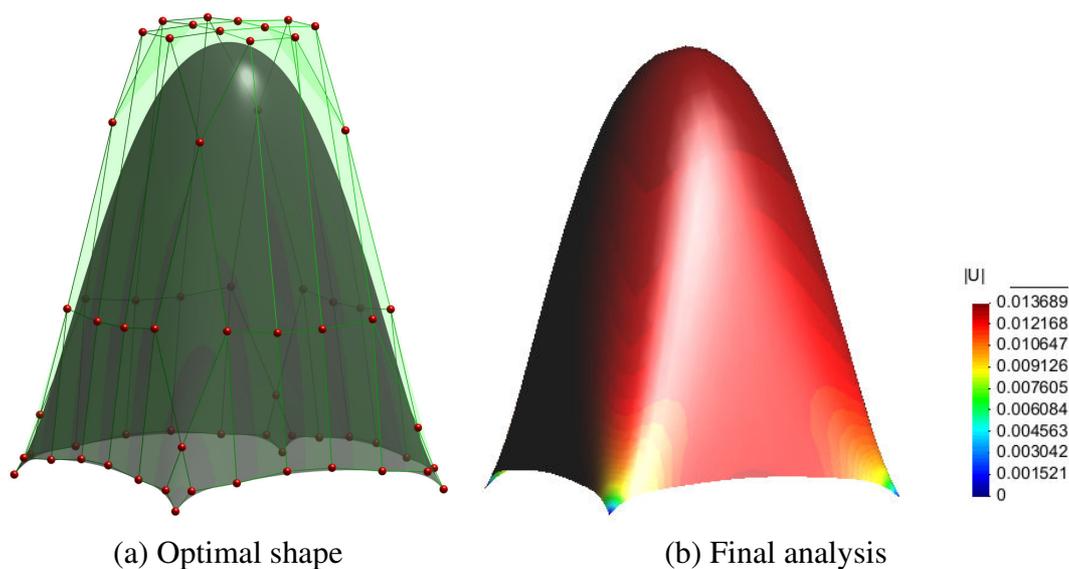


Figure 16: Optimal shape and displacement analysis.

The non linear problem is solved by an incremental iterative method. Then, it is necessary to check the derivative convergence rate with respect to the functional convergence rate. The derivative convergence was checked decreasing the convergence parameters of the non linear analysis, using  $10^{-3}, \dots, 10^{-9}$ , and comparing with FD. For the adopted functional it was checked that the derivative convergence rate is equal or faster than the functional convergence rate.

#### 4 CONCLUSIONS

Although a global minimum is required in optimization problems, such those presented here, results obtained in previous examples may corresponds to a local minima, once the convexity behavior of the Lagrangian function have not been studied here.

Mesh distortion with respect to the initial grid is one of the difficulties to be avoided. An isogeometric structural analysis would be an alternative, but an algorithm to reorganize nodes localization at each iteration may be also effectively used.

Regarding the sensitivity analysis, several options have been successfully used. The simplest tool is Finite Differences (FD), but in terms of numerical accuracy Automatic Differentiation avoids numerical errors, in spite of the fact that some numerical errors may occurs originated in the structural analysis. AD is very fast in terms of time processing (about twice the objective function evaluation), but its main limitation is the required computer memory.

NURBS provides a smooth and accurate parametric representation of surfaces and it is well suitable to shape modifications in the optimization problem. In this work is easy to observe that the most adequate way to optimize a surface shape described by NURBS is to take all its characteristics (knot vectors, weights and control points) as variable to be optimized.

An important limitation of this work is to consider that the thickness remains constant in each element and that its variation is quadratic. To generalize the thickness variation it would be necessary to implement solid NURBS, increasing twice the number of variables.

Finally, aims achieved in this work are listed below:

- Solution in some case is not unique, which means a non uniqueness proof (Example 1);

- Convergence of derivatives in AD are faster than the convergence of the objective and the constraints if an iterative method is used (Example 1);
- Comparison between FD and AD (Example 2);
- The influence of different load in the optimum shape (Example 3);
- Use of control points, weights and knot position optimization as variables (Example 4);
- Convergence of derivatives in AD are faster than the convergence of the objective and the constraints if an incremental-iterative method in non linear analysis (Example 5);
- Shells with an excellent structural performance (Example 1 to 5).

## 5 ACKNOWLEDGMENTS

The authors thank to CAPES and CNPq for their financial support.

## REFERENCES

- Bathe K.J. and Ho L.W. A simple and effective element for analysis of general shell structures. *Computers & Structures*, 13(5-6):673 – 681, 1981.
- Galilei G. Discorsi e dimonstrazioni matematiche intorno à due nuove scienze. *attententi alla Mecanica & i Movimenti Locali*, 0:112p., 1638.
- Giering R. and Kaminski T. Recipes for adjoint code construction. Technical Report, Internal Report 212 from Max-Plank Institute für Meteorologie. Hamburg, Germany., 1996.
- Griewank A. and Walther A. *Evaluating derivatives: principles and techniques of algorithmic differentiation*. Philadelphia, 2<sup>nd</sup> edition, 2008.
- INRIA. *Tapenade AD*. <http://tapenade.inria.fr:8080/tapenade/index.jsp>, 2002.
- Kegl M. and Brank B. Shape optimization of truss-stiffened shell structures with variable thickness. *Computer Methods in Applied Mechanics and Engineering*, 195(19-22):2611 – 1634, 2006.
- Nocedal J. and Wright S.J. *Numerical Optimization*. Springer, 2<sup>nd</sup> edition, 1999.
- Piegl L.A. and Tiller W. *The Nurbs Book*. Springer, 2<sup>nd</sup> edition, 1997.
- Ramm E., Bletzinger K.U., and Reitingner R. Shape optimization of shell structures. *IASS Bulletin of the international association for shell and spatial structures*, 34(112):103–121, 1993.
- Spellucci P. *DONLP2*. <http://www.mathematik.tu-darmstadt.de/fbereiche/numerik/staff/spellucci/DONLP2/index.html>, 2001.
- Yang Y.B. and Shieh M.S. Solution method for nonlinear problems with multiple critical points. *AIAA Journal*, 28:2110–2116, 1990.