# OPTIMAL DESIGN OF STRUCTURES (MAP 562)

**G. ALLAIRE**

**January 14th, 2015**

**Department of Applied Mathematics, Ecole Polytechnique**

CHAPTER V

# PARAMETRIC (OR SIZING) OPTIMIZATION (end)

Numerical algorithms

## Reminder: membrane model problem

Fixed membrane $\Omega \subset \mathbb{R}^N$, forces $f \in L^2(\Omega)$, displacement $u \in H_0^1(\Omega)$

$$\begin{cases} -\operatorname{div}(h\nabla u) = f & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases}$$

$$\inf_{h \in \mathcal{U}_{ad}} J(h) = \int_\Omega j(u)\,dx$$

Admissible set of thicknesses

$\mathcal{U}_{ad} = \left\{ h \in L^\infty(\Omega)\ , \quad h_{max} \geq h(x) \geq h_{min} > 0 \text{ in } \Omega, \int_\Omega h(x)\,dx = h_0|\Omega| \right\}.$

**Theorem 5.19.** The cost function $J(h)$ is differentiable on $\mathcal{U}_{ad}$ and

$$J'(h) = \nabla u \cdot \nabla p$$

where $p$ is the adjoint state, unique solution in $H_0^1(\Omega)$ of

$$\begin{cases} -\operatorname{div}(h\nabla p) = -j'(u) & \text{in } \Omega \\ p = 0 & \text{on } \partial\Omega. \end{cases}$$

## 5.3.2 Numerical algorithm

### Projected gradient

1. Initialization of the thickness $h_0 \in \mathcal{U}_{ad}$ (by example, a constant function which satisfies the constraints).

2. Iterations until convergence, for $n \geq 0$:

$$h_{n+1} = P_{\mathcal{U}_{ad}}\Big(h_n - \mu J'(h_n)\Big),$$

where $\mu > 0$ is a descent step, $P_{\mathcal{U}_{ad}}$ is the projection operator on the closed convex set $\mathcal{U}_{ad}$ and the derivative is given by

$$J'(h_n) = \nabla u_n \cdot \nabla p_n$$

with the state $u_n$ and the adjoint $p_n$ (associated to the thickness $h_n$).

To make the algorithm fully explicit, we have to precise what is the projection operator $P_{\mathcal{U}_{ad}}$.

We characterize the projection operator $P_{\mathcal{U}_{ad}}$

$$\Big(P_{\mathcal{U}_{ad}}(h)\Big)(x) = \max\left(h_{min}, \min\left(h_{max}, h(x) + \ell\right)\right)$$

where $\ell$ is the unique Lagrange multiplier such that

$$\int_\Omega P_{\mathcal{U}_{ad}}(h)\, dx = h_0|\Omega|.$$

The determination of the constant $\ell$ is not explicit: we must use an iterative algorithm based on the property of the function

$$\ell \to F(\ell) = \int_\Omega \max\left(h_{min}, \min\left(h_{max}, h(x) + \ell\right)\right)\, dx$$

which is strictly increasing on the interval $[\ell^-, \ell^+]$, reciprocal image of $[h_{min}|\Omega|, h_{max}|\Omega|]$. Thanks to this monotonicity property, we propose a simple iterative algorithm: we first bracket the root by an interval $[\ell^1, \ell^2]$ such that

$$F(\ell^1) \leq h_0|\Omega| \leq F(\ell^2),$$

then we proceed by dichotomy to find the root $\ell$.

- ☞ In practice, we rather use a projected gradient algorithm with a variable step (not optimal) which guarantees the decrease of the functional $J(h_{n+1}) < J(h_n)$.
- ☞ The algorithm is rather slow. A possible acceleration is based on the quasi-Newton algorithm.
- ☞ The overhead generated by the adjoint computation is very modest : one has to build a new right-hand-side (using the state) and solve the corresponding linear system (with the same rigidity matrix).
- ☞ Convergence is detected when the optimality condition is satisfied with a threshold $\epsilon > 0$

$$|h_n - \max\left(h_{min}, \min\left(h_{max}, h_n - \mu_n J'(h_n) + \ell_n\right)\right)| \leq \epsilon \mu_n h_{max}.$$

## 5.4.3 Numerical algorithm for the compliance

When $j(u) = fu$, we find $p = -u$ since $j'(u) = f$. This particular case is said to be **self-adjoint**.

We use the dual or complementary energy

$$J(h) = \int_\Omega f u \, dx = \min_{\substack{\tau \in L^2(\Omega)^N \\ -\operatorname{div}\tau = f \text{ in } \Omega}} \int_\Omega h^{-1} |\tau|^2 dx \, .$$

We can rewrite the optimization problem as a double minimization

$$\inf_{h \in \mathcal{U}_{ad}} \min_{\substack{\tau \in L^2(\Omega)^N \\ -\operatorname{div}\tau = f \text{ in } \Omega}} \int_\Omega h^{-1} |\tau|^2 dx \, ,$$

and the order of minimization is irrelevent.

The problem is convex and admits a minimizer.

**Lemma 5.25 (optimality conditions).** For a given $\tau \in L^2(\Omega)^N$, the problem

$$\min_{h \in \mathcal{U}_{ad}} \int_\Omega h^{-1} |\tau|^2 dx$$

admits a minimizer $h(\tau)$ in $\mathcal{U}_{ad}$ given by

$$h(\tau)(x) = \begin{cases} h^*(x) & \text{if } h_{min} < h^*(x) < h_{max} \\ h_{min} & \text{if } h^*(x) \leq h_{min} \\ h_{max} & \text{if } h^*(x) \geq h_{max} \end{cases} \qquad \text{with } h^*(x) = \frac{|\tau(x)|}{\sqrt{\ell}},$$

where $\ell \in \mathbb{R}^+$ is the Lagrange multiplier such that $\displaystyle\int_\Omega h(x)\, dx = h_0 |\Omega|$.

## Optimality criteria method

1. Initialization of the thickness $h_0 \in \mathcal{U}_{ad}$.

2. Iterations until convergence, for $n \geq 0$:

   (a) Computation of the state $\tau_n$, unique solution of

   $$\min_{\substack{\tau \in L^2(\Omega)^N \\ -\operatorname{div}\tau = f \text{ in } \Omega}} \int_\Omega h_n^{-1} |\tau|^2 dx ,$$

   with the previous thickness $h_n$.

   (b) Update of the thickness :

   $$h_{n+1} = h(\tau_n),$$

   where $h(\tau)$ is the minimizer defined by the optimality condition. The Lagrange multiplier is computed by dichotomy.

Remark that minimizing in $\tau$ is equivalent to solving the equation

$$\left\{\begin{array}{ll} -\operatorname{div}(h_n \nabla u_n) = f & \text{in } \Omega \\ u_n = 0 & \text{on } \partial\Omega, \end{array}\right.$$

and we recover $\tau_n$ by the formula

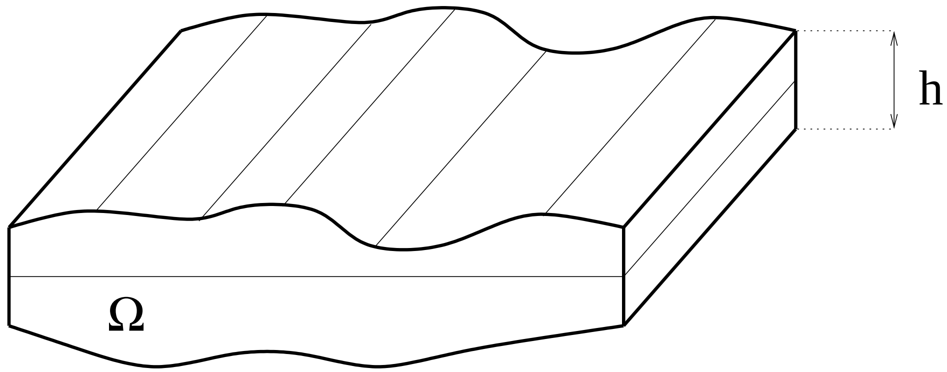
$$\tau_n = h_n \nabla u_n.$$

This algorithm can be interpreted as an alternate minimization in $\tau$ and $h$ of the objective function. In particular, we deduce that the objective function **always decreases** through the iterations

$$J(h_{n+1}) = \int_\Omega h_{n+1}^{-1} |\tau_{n+1}|^2 dx \leq \int_\Omega h_{n+1}^{-1} |\tau_n|^2 dx \leq \int_\Omega h_n^{-1} |\tau_n|^2 dx = J(h_n).$$

This algorithm can also be interpreted as an optimality criteria method.

## 5.6 Thickness optimization of an elastic plate




$$\begin{cases} -\operatorname{div}\sigma = f & \text{in } \Omega \\ \sigma = 2\mu h\, e(u) + \lambda h\, \operatorname{tr}(e(u))\, \mathrm{Id} & \text{in } \Omega \\ u = 0 & \text{on } \Gamma_D \\ \sigma n = g & \text{on } \Gamma_N \end{cases}$$

with the strain tensor $e(u) = \frac{1}{2}\big(\nabla u + (\nabla u)^t\big)$.

Set of admissible thicknesses:

$$\mathcal{U}_{ad} = \left\{ h \in L^\infty(\Omega) \ , \quad h_{max} \geq h(x) \geq h_{min} > 0 \text{ in } \Omega, \int_\Omega h(x)\,dx = h_0|\Omega| \right\}.$$

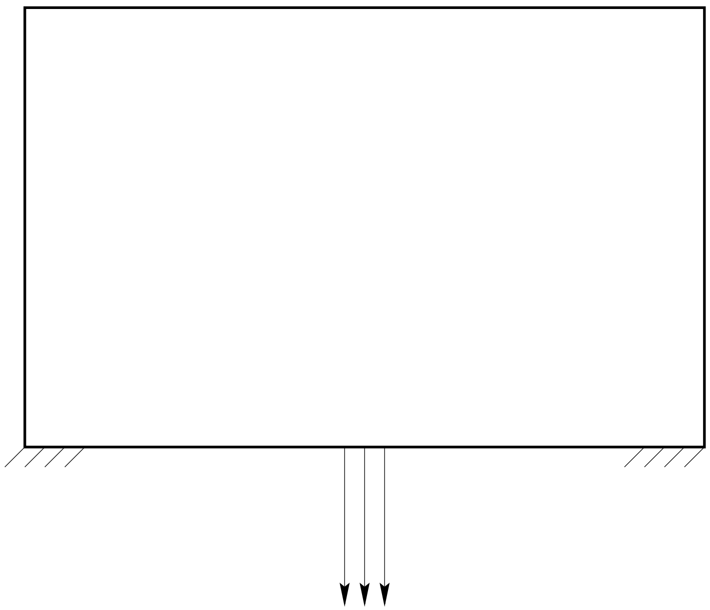
The compliance optimization can be written

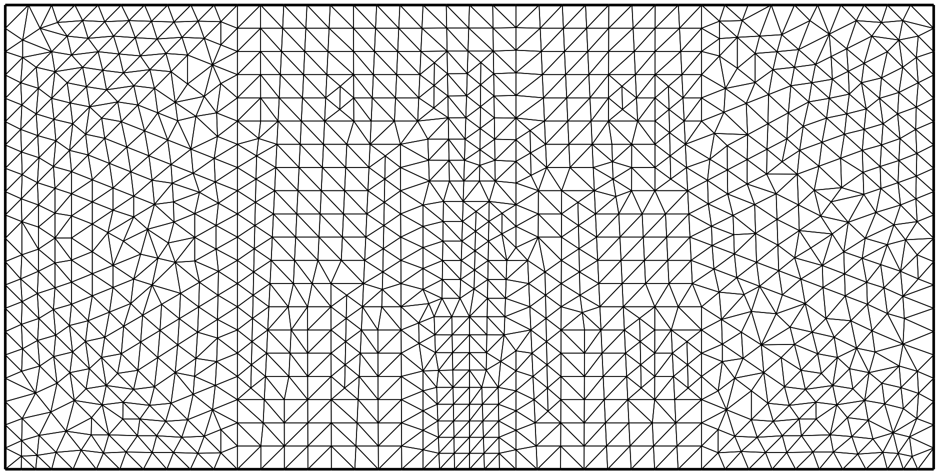
$$\inf_{h \in \mathcal{U}_{ad}} J(h) = \int_\Omega f \cdot u\,dx + \int_{\Gamma_N} g \cdot u\,ds.$$

The theoretical results are the same.

We apply the optimality criteria method.

## Boundary conditions and mesh for an elastic plate

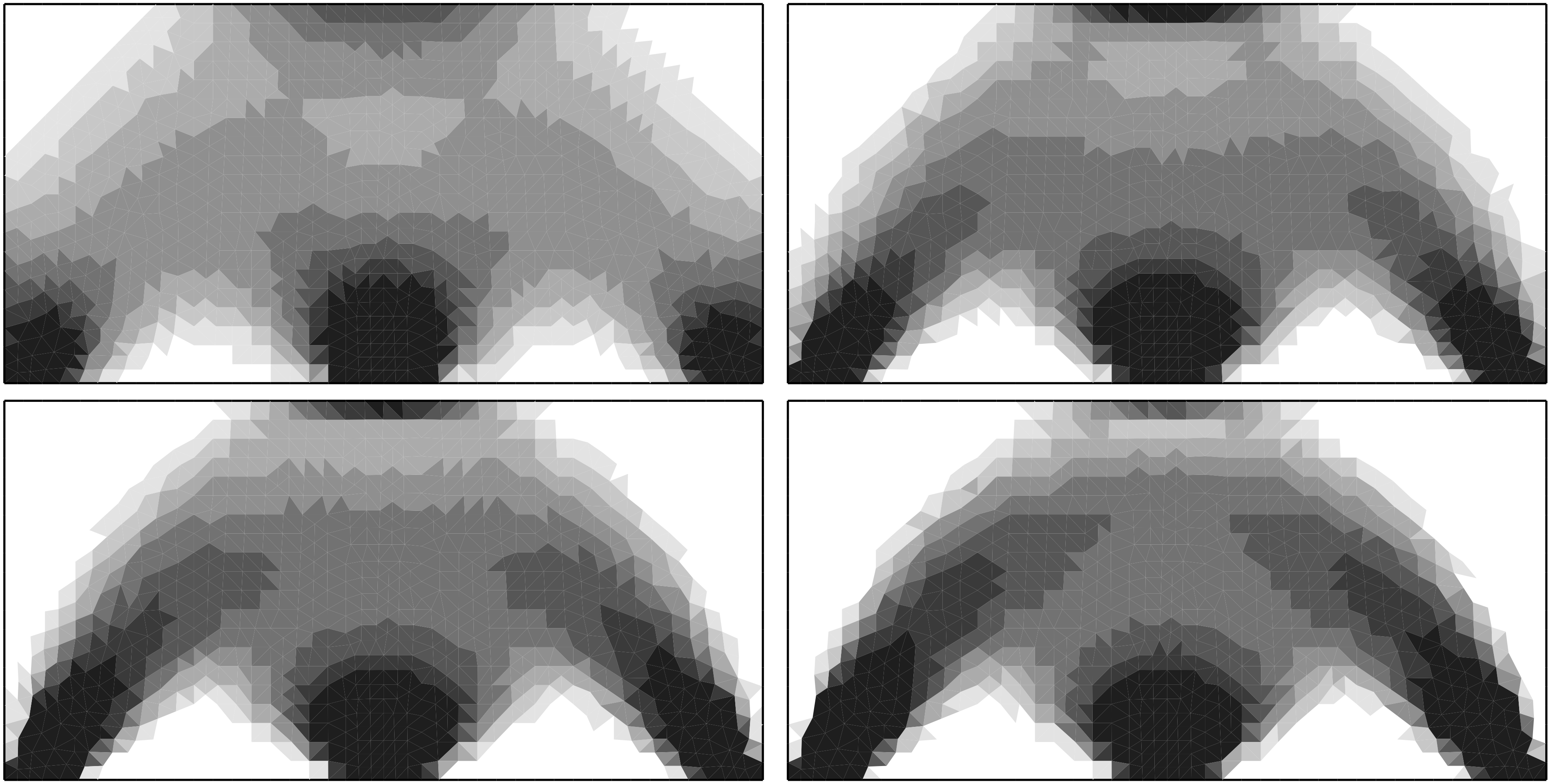




FreeFem++ computations ; scripts available on the web page

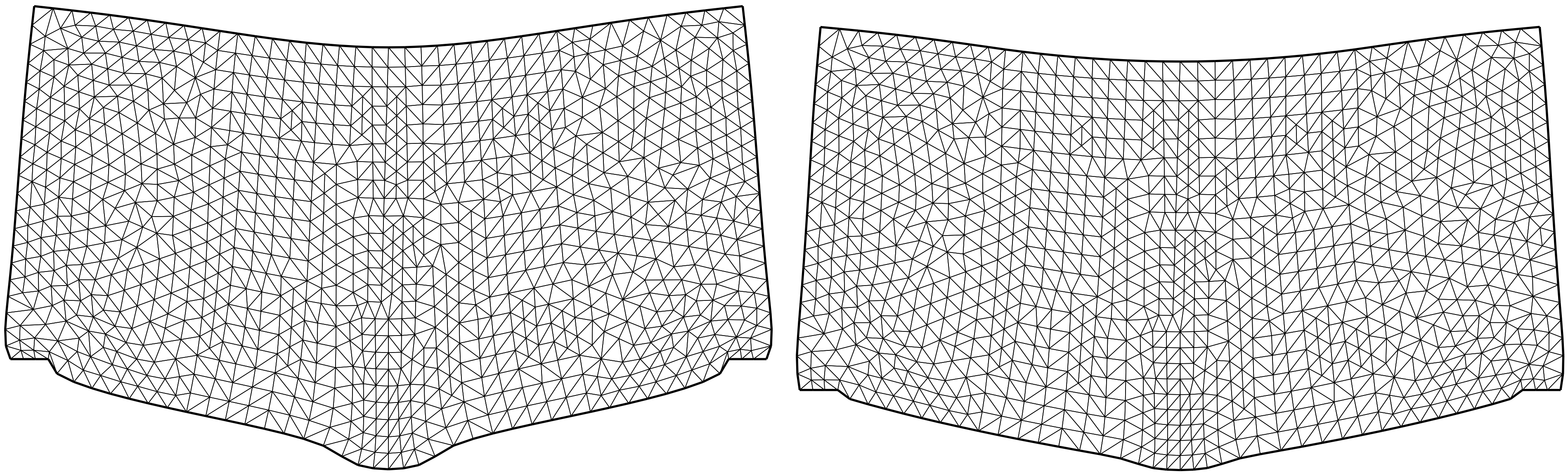
`http://www.cmap.polytechnique.fr/~allaire/cours_X_annee3.html`

# Thickness at iterations 1, 5, 10, 30 (uniform initialization).

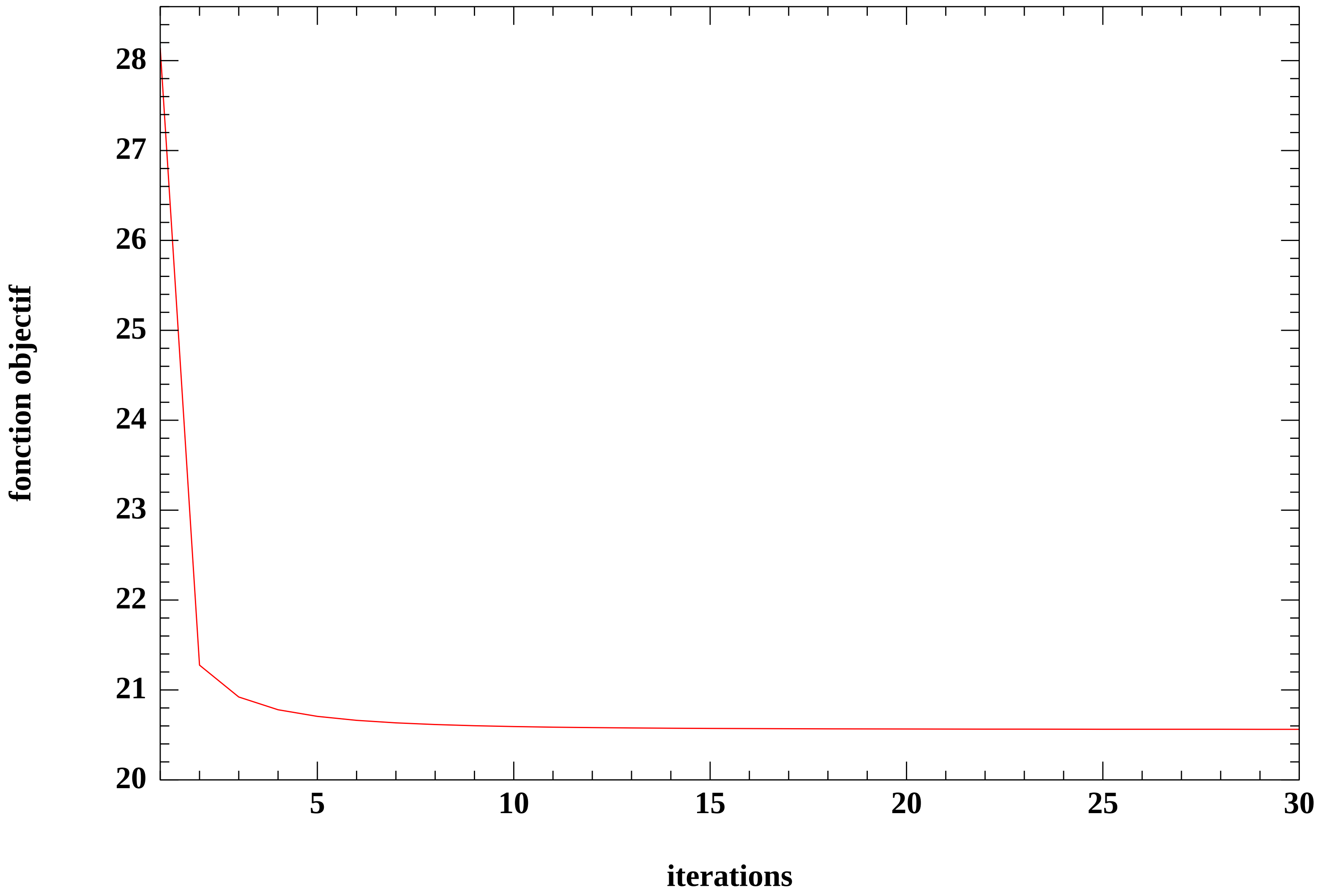


$h_{min} = 0.1$, $h_{max} = 1.0$, $h_0 = 0.5$ (increasing thickness from white to black)

## Comparing the initial and final deformed shapes



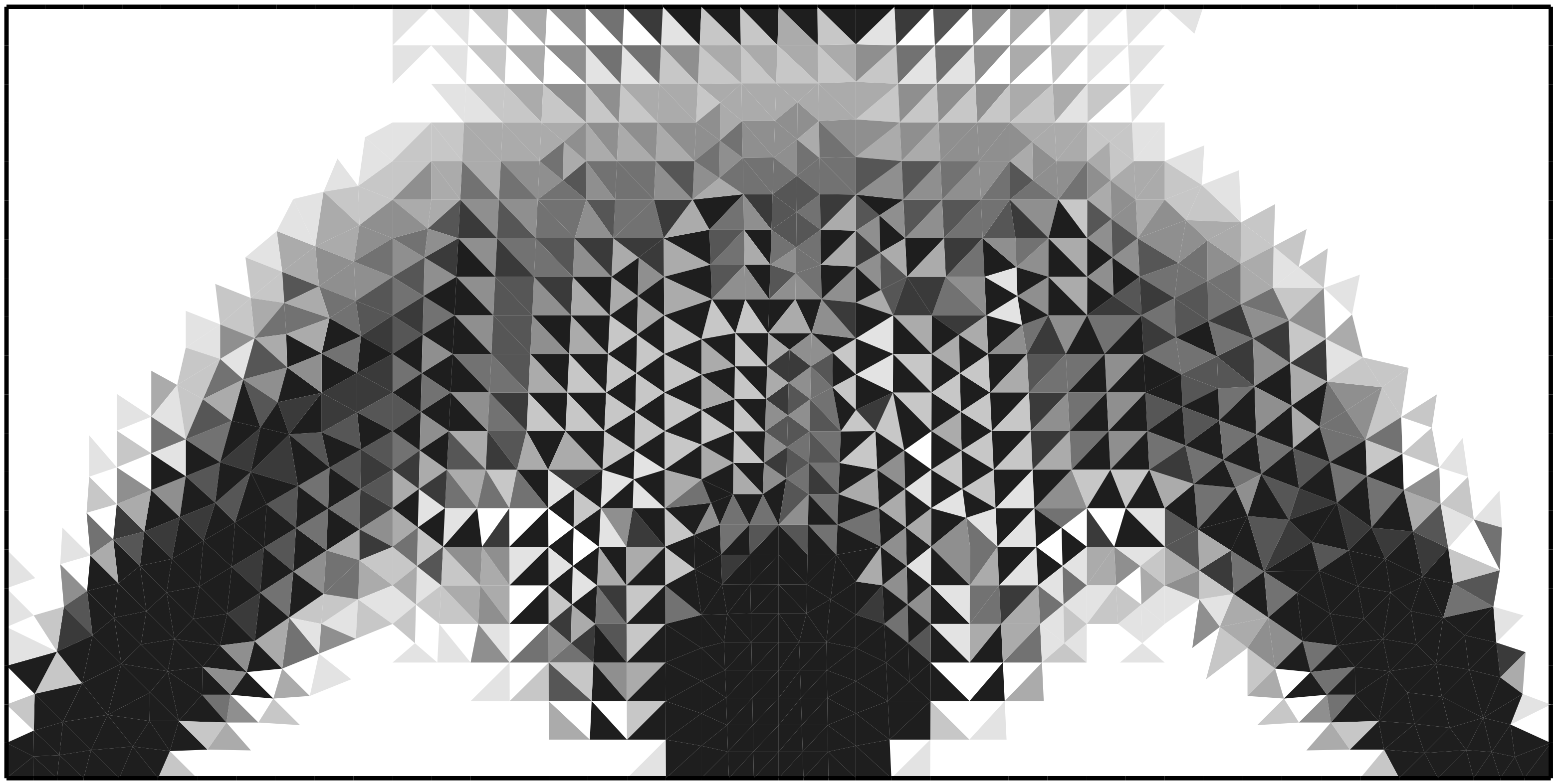
# Convergence history



## Numerical instabilities (checkerboards)

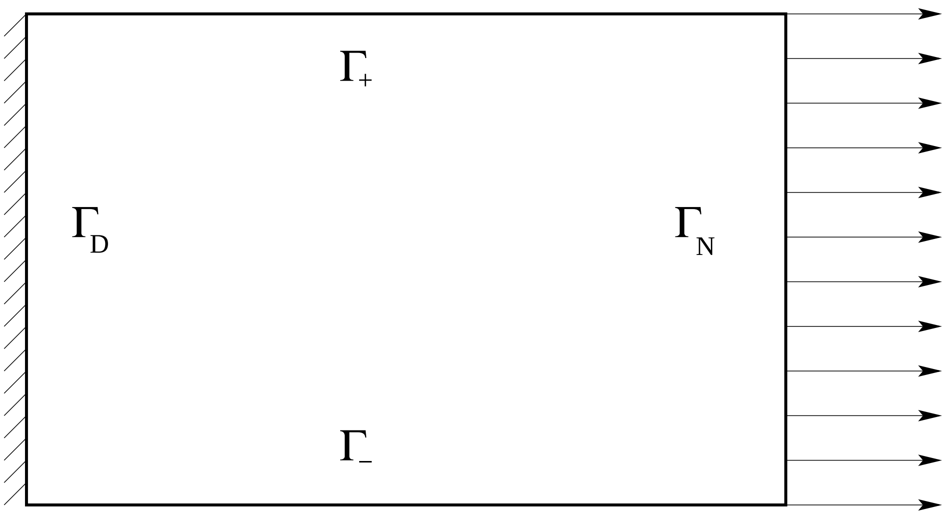
☞ Finite elements $P2$ for $u$ and $P0$ for $h$ $\Rightarrow$ OK

☞ Finite elements $P1$ for $u$ and $P0$ for $h$ $\Rightarrow$ unstable !

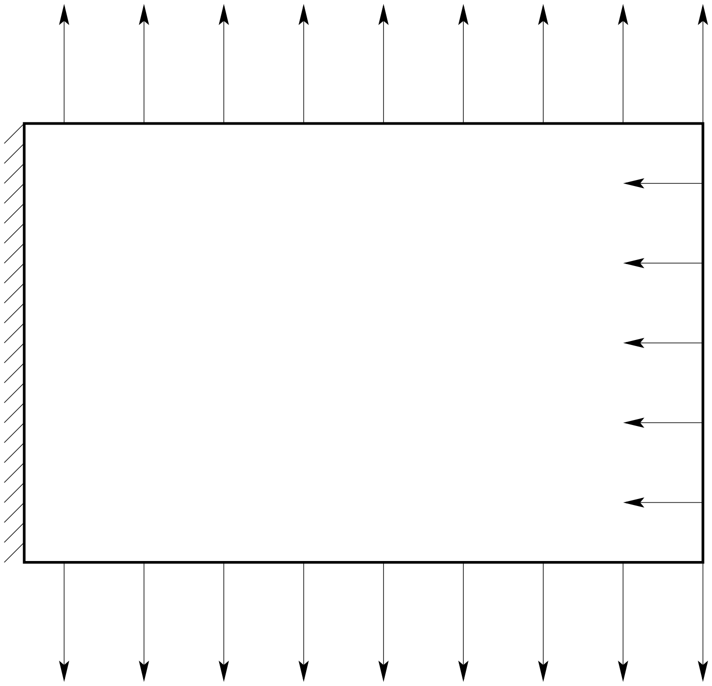


## Numerical counter-example of non-existence of an optimal shape (in elasticity)

We look for the design which horizontally is less deformed and vertically more deformed.

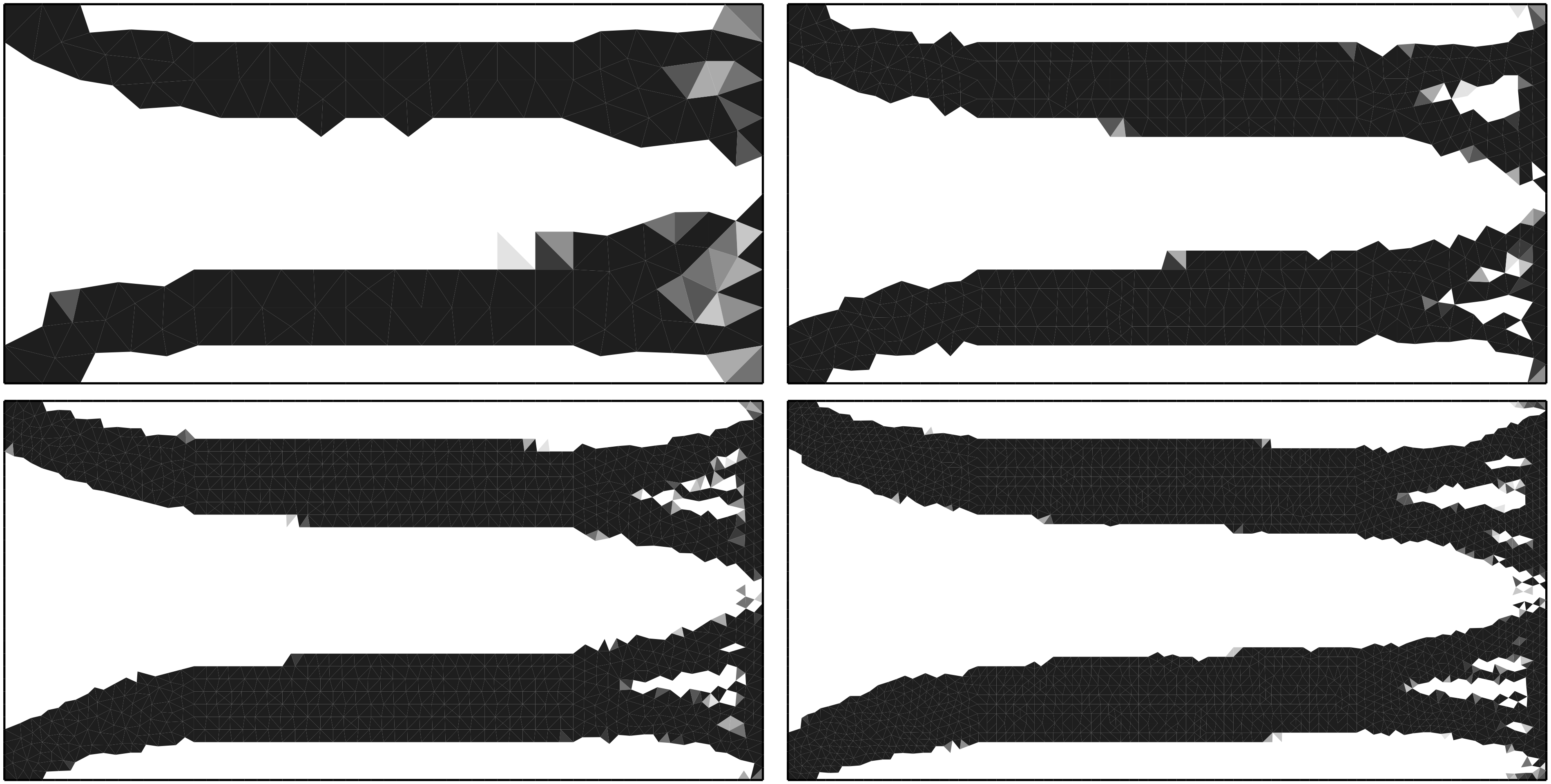



boundary conditions



target displacement

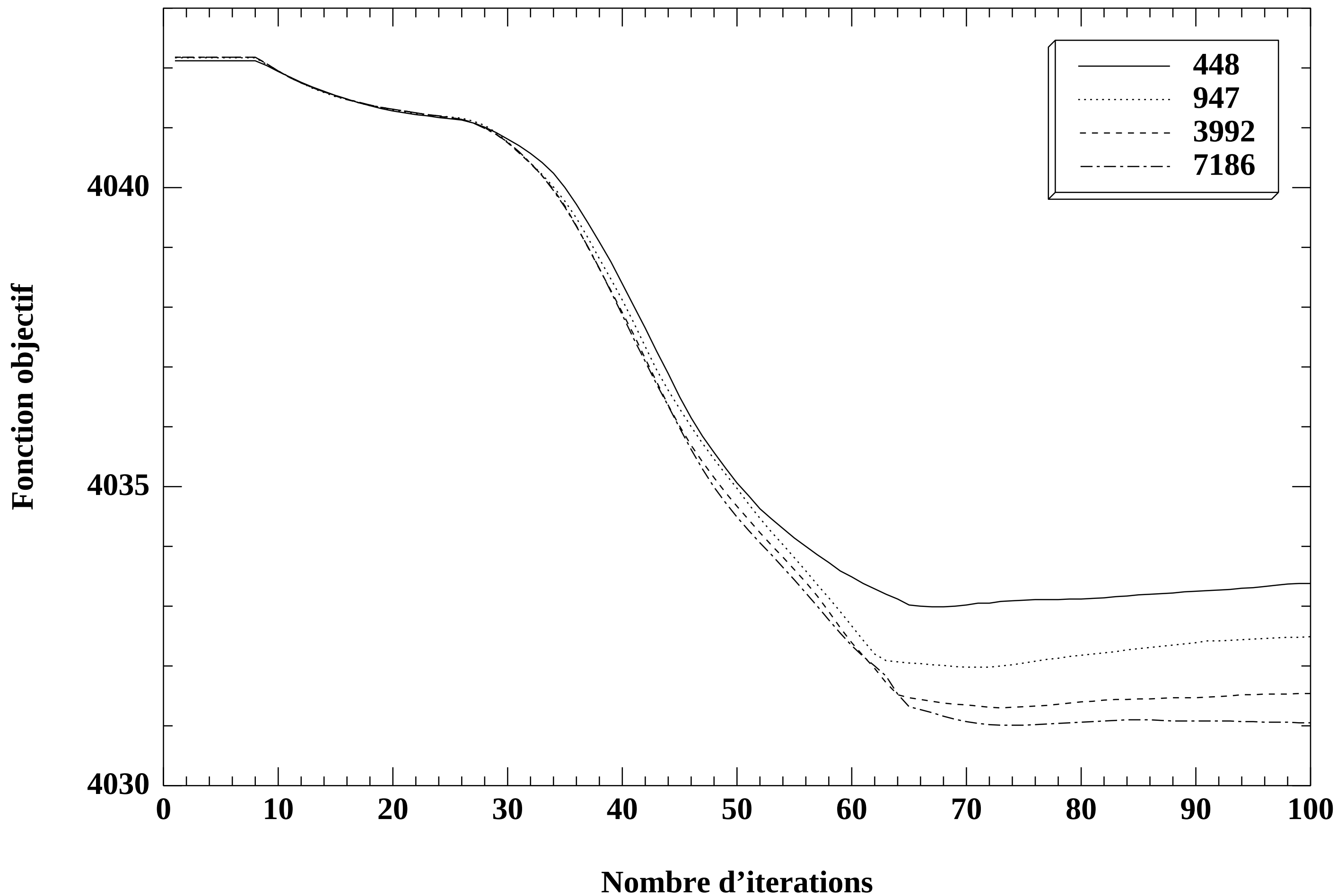
## Optimal shapes for meshes with 448, 947, 3992, 7186 triangles



## No convergence under mesh refinement !

More and more details appear when the mesh size is decreased.

The value of the objective function decreases with the mesh size.



## 5.6.4 Regularization

Triple motivation:

- ☞ To avoid instabilities when using $P1$ finite elements for $u$ and $P0$ for $h$ (less expensive than $P2$-$P0$).
- ☞ To obtain an algorithm which converges by mesh refinement.
- ☞ To adhere to the "regularized" framework of section 5.2.3 (with existence of optimal solutions).

Main idea: we change the scalar product

$$\langle J'(h), k\rangle = \int_\Omega k\nabla u \cdot \nabla p\, dx \quad \forall\, k \in \mathcal{U}_{ad}.$$

Previously we identified $\mathcal{U}_{ad}$ to a subspace of $L^2(\Omega)$, thus

$$\langle J'(h), k\rangle = \int_\Omega J'(h)\, k\, dx \qquad \Rightarrow \qquad J'(h) = \nabla u \cdot \nabla p\ .$$

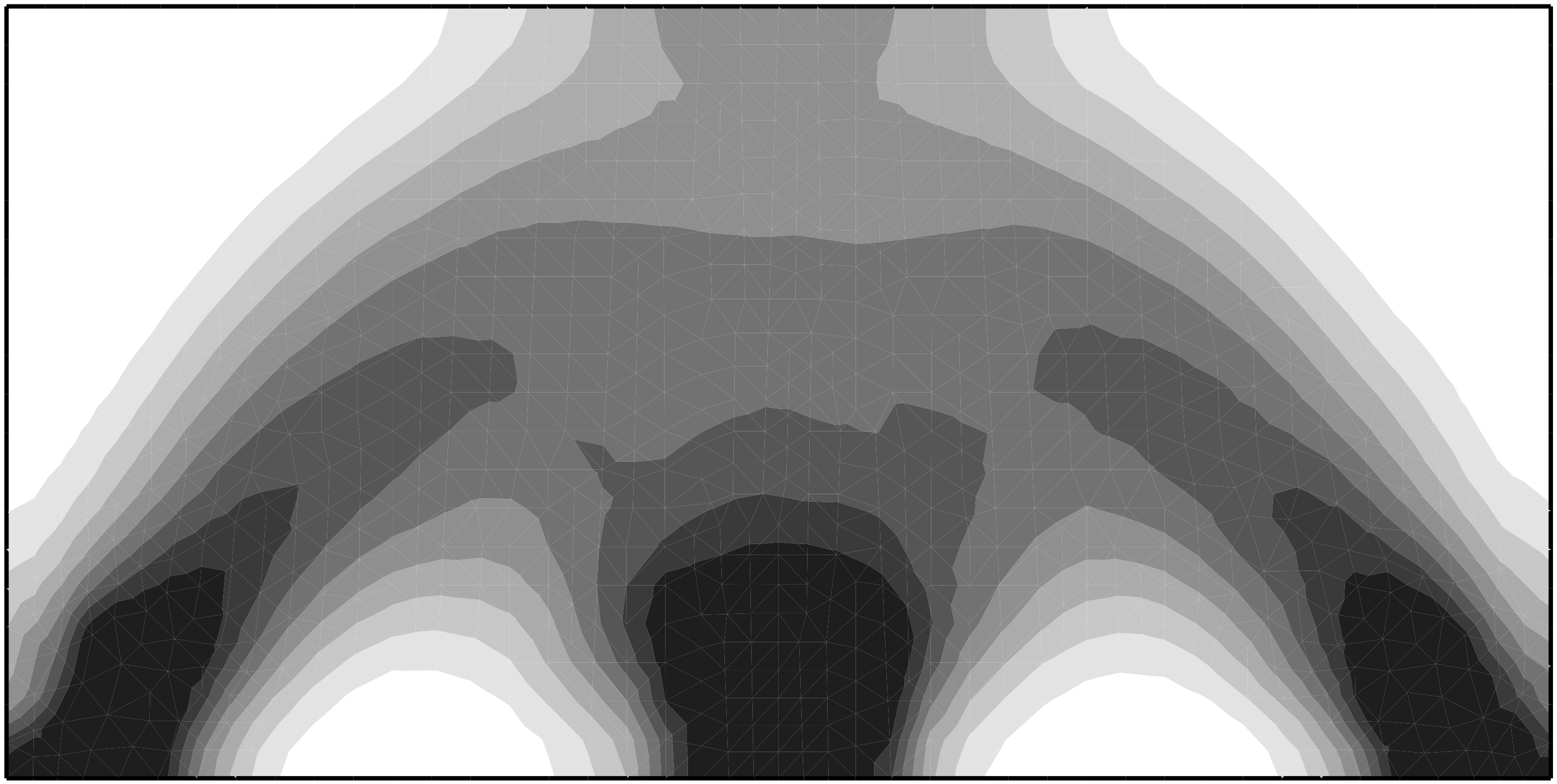
Now, we identify a "regularized" admissible set $\mathcal{U}_{ad}^{reg}$ to a subspace $H^1(\Omega)$, thus

$$\langle J'(h), k\rangle = \int_\Omega (\nabla J'(h) \cdot \nabla k + J'(h)k)\, dx\ ,$$

and we deduce a new formula for the gradient

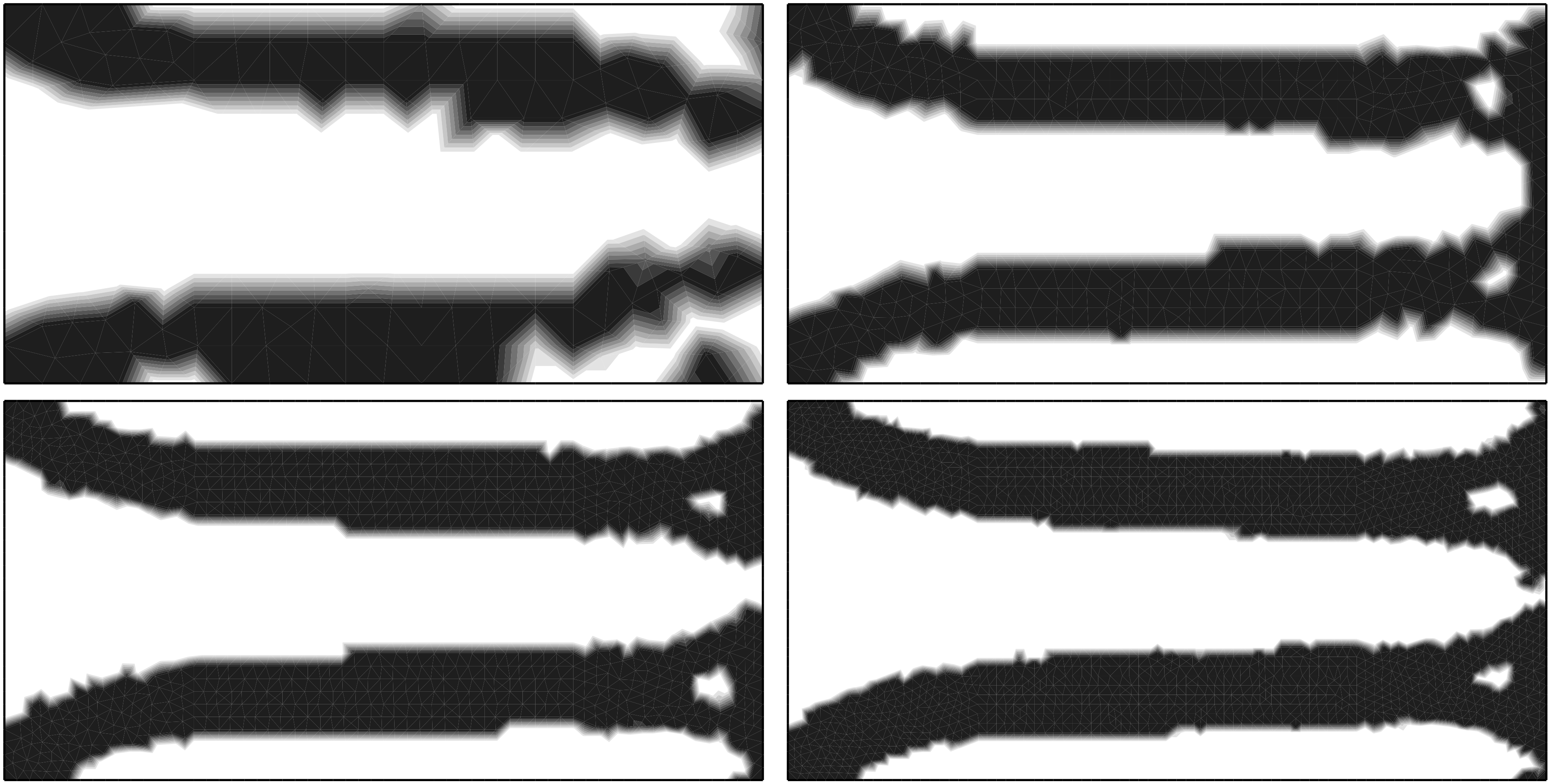
$$\begin{cases} -\Delta J'(h) + J'(h) = \nabla u \cdot \nabla p & \text{in } \Omega, \\ \dfrac{\partial J'(h)}{\partial n} = 0 & \text{on } \partial\Omega. \end{cases}$$

## Regularized optimal shape



Finite elements $P_1$-$P_0$. Compliance minimization. Alternate directions algorithm.

## Convergence by mesh refinement



Same case as the “numerical counter-examples” (meshes 448, 947, 3992, 7186).

## Conclusion

- ☞ Regularization works !
- ☞ It costs a bit more (solving an additional Laplacian to compute the gradient).
- ☞ Difficulty in choosing the regularization parameter $\epsilon > 0$ (which can be interpreted as a lengthscale)

$$-\epsilon^2 \Delta J'(h) + J'(h) = \nabla u \cdot \nabla p \quad \text{in } \Omega$$

- ☞ It has a tendency to smooth the geometric details.

CHAPTER VI

# GEOMETRIC OPTIMIZATION (First Part)

# Geometric optimization of a membrane

A membrane is occupying a variable domain $\Omega$ in $\mathbb{R}^N$ with boundary

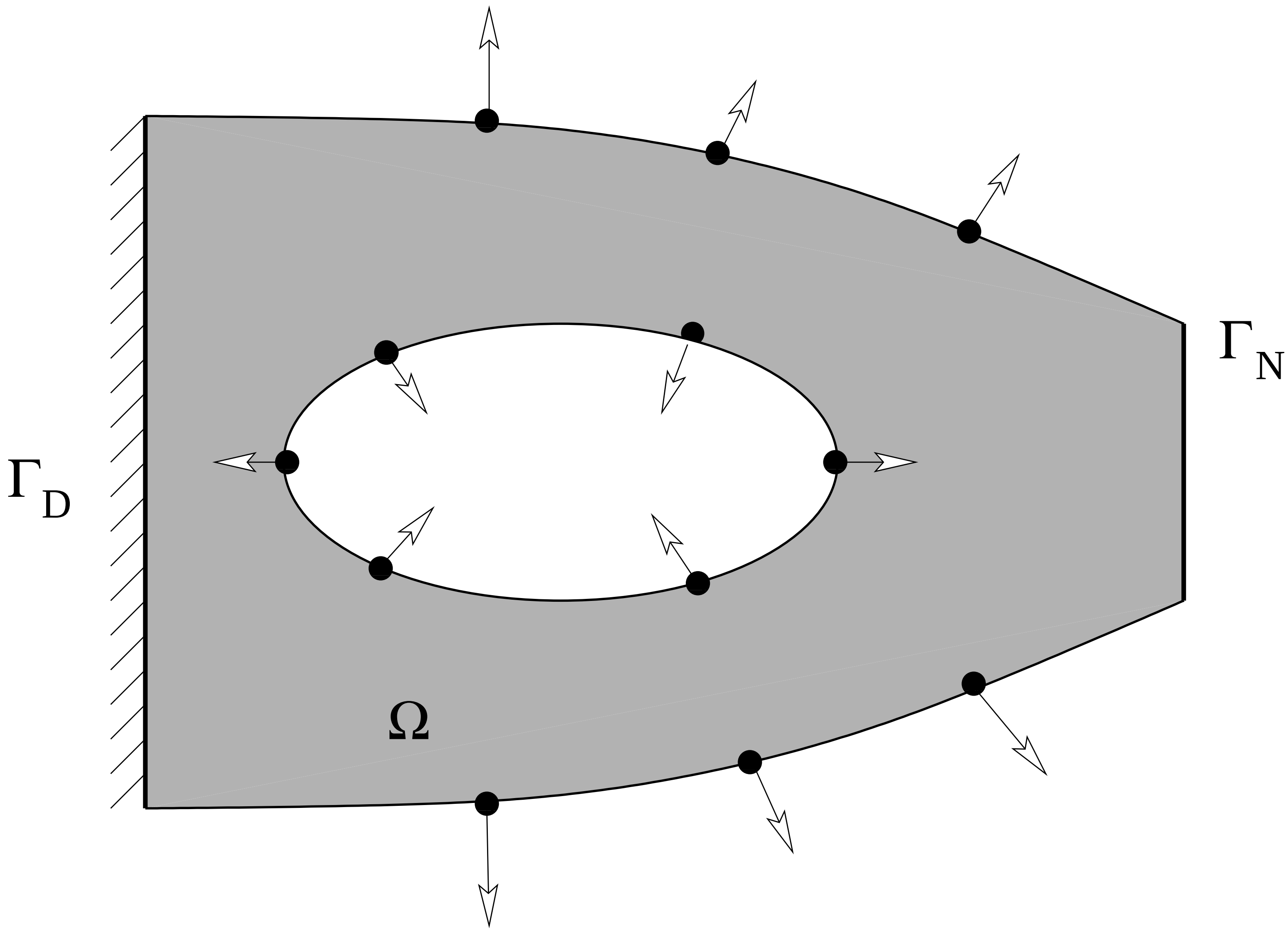
$$\partial\Omega = \Gamma \cup \Gamma_N \cup \Gamma_D,$$

where $\Gamma \neq \emptyset$ is the variable part of the boundary, $\Gamma_D \neq \emptyset$ is a fixed part of the boundary where the membrane is clamped, and $\Gamma_N \neq \emptyset$ is another fixed part of the boundary where the loads $g \in L^2(\Gamma_N)$ are applied.

$$\begin{cases} -\Delta u = 0 & \text{in } \Omega \\ u = 0 & \text{on } \Gamma_D \\ \frac{\partial u}{\partial n} = g & \text{on } \Gamma_N \\ \frac{\partial u}{\partial n} = 0 & \text{on } \Gamma \end{cases}$$

(No bulk forces to simplify)

# Boundary variation in geometric optimization



## Shape optimization of a membrane

**Geometric** shape optimization problem

$$\inf_{\Omega \in \mathcal{U}_{ad}} J(\Omega)$$

We must defined the set of admissible shapes $\mathcal{U}_{ad}$. That is the main difficulty.

**Examples:**

☞ Compliance or work done by the load (rigidity measure)

$$J(\Omega) = \int_{\Gamma_N} g u \, ds$$

☞ Least square criterion for a target displacement $u_0 \in L^2(\Omega)$

$$J(\Omega) = \int_{\Omega} |u - u_0|^2 dx$$

where $u$ depends on $\Omega$ through the state equation.

# 6.2 Existence results

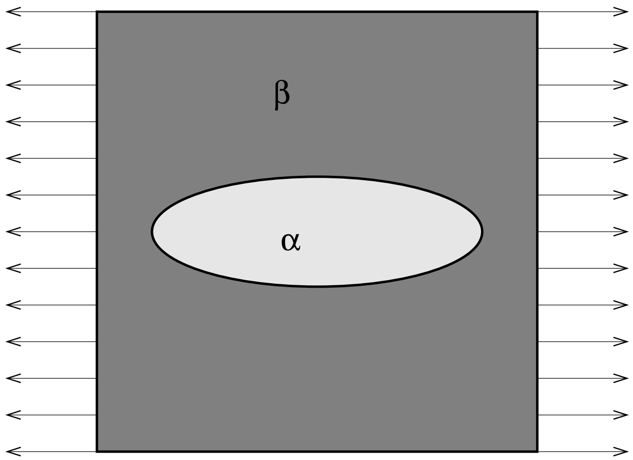
**In full generality, there does not exist any optimal shape !**

- ☞ Existence under a geometric constraint.
- ☞ Existence under a topological constraint.
- ☞ Existence under a regularity constraint.
- ☞ Counter-example in the absence of these conditions.

**related questions:**

- ☞ How to pose the problem ? How to parametrize shapes ?
- ☞ Calculus of variations for shapes.
- ☞ Mathematical framework for establishing numerical algorithms.

## 6.2.1 Counter-example of non-existence




Let $D = ]0;1[\times]0;L[$ be a rectangle in $\mathbb{R}^2$. We fill $D$ with a mixture of two materials, homogeneous isotropic, characterized by an elasticity coefficient $\beta$ for the strong material, and $\alpha$ for the weak material (almost like void) with $\beta >> \alpha > 0$. We denote by $\chi(x) = 0, 1$ the **characteristic function** of the weak phase $\alpha$, and we define

$$a_\chi(x) = \alpha\chi(x) + \beta(1 - \chi(x)).$$

(Other possible interpretation: variable thickness which can take only two values.)

State equation:

$$\begin{cases} -\operatorname{div}(a_\chi \nabla u_\chi) = 0 & \text{in } D \\ a_\chi \nabla u_\chi \cdot n = e_1 \cdot n & \text{on } \partial D \end{cases}$$

Uniform horizontal loading.

Objective function: compliance

$$J(\chi) = \int_{\partial D} (e_1 \cdot n) u_\chi ds$$

Admissible set: **no geometric or smoothness constraint, i.e.** $\chi \in L^\infty(D; \{0,1\})$. There is however a volume constraint

$$\mathcal{U}_{ad} = \left\{ \chi \in L^\infty \left(D; \{0,1\}\right) \text{ such that } \frac{1}{|D|} \int_D \chi(x)\, dx = \theta \right\},$$

otherwise the strong phase would always be prefered !

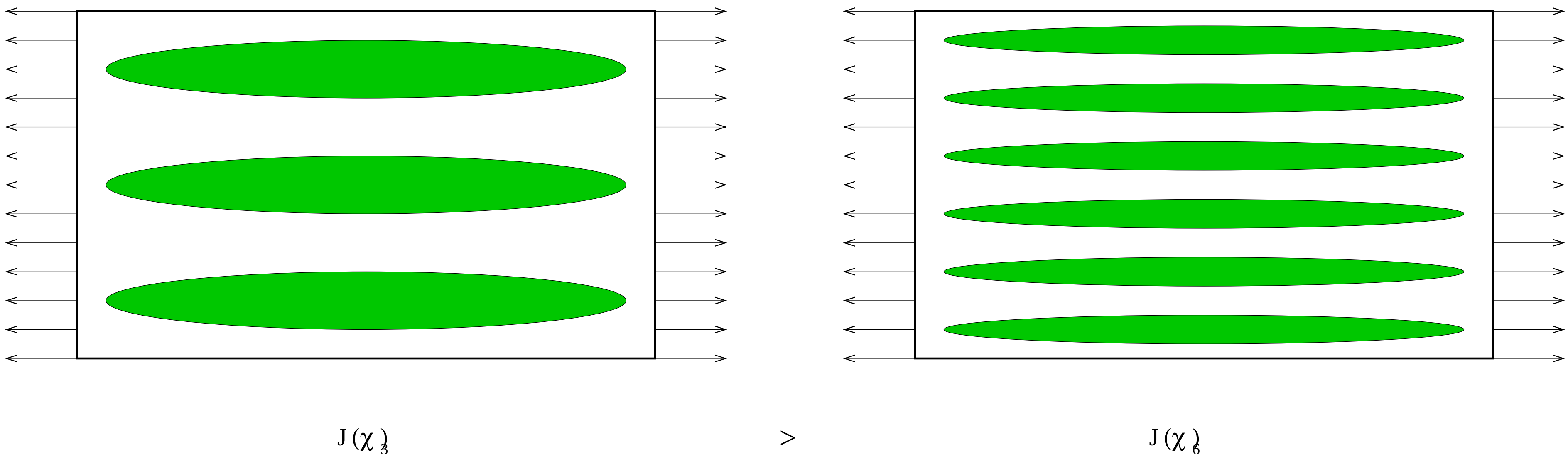
The shape optimization problem is:

$$\inf_{\chi \in \mathcal{U}_{ad}} J(\chi).$$

## Non-existence

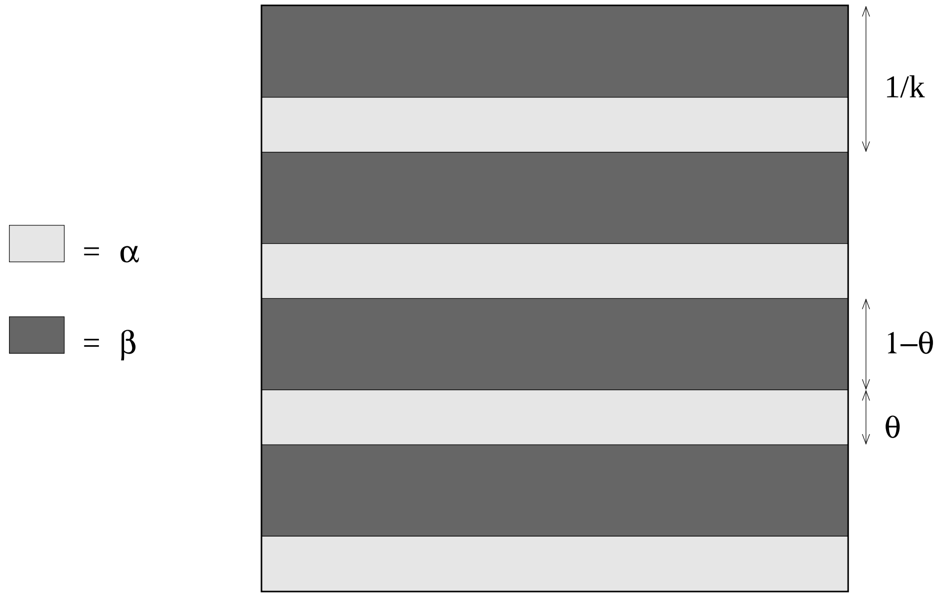
**Proposition 6.2.** If $0 < \theta < 1$, there does not exist an optimal shape in the set $\mathcal{U}_{ad}$.

**Remark.** Cause of non-existence = lack of geometric or smoothness constraint on the shape boundary.




**Many small holes are better than just a few bigger holes !**

## Mechanical intuition




Minimizing sequence $k \to +\infty$: $k$ rigid fibers, aligned in the principal stress $e_1$, and uniformly distributed. To achieve a uniform boundary condition, the fibers must be finer and finer and alternate more and more weak and strong ones.

This is the main idea of a minimizing sequence which never achieves the minimum.

## 6.2.2 Existence under a geometric condition

Let $D$ be a given working domain. We define

$$\mathcal{U}_{ad} = \left\{ \Omega \subset D \text{ such that } \begin{array}{l} (i)\ \Omega \text{ satisfies the uniform cone property} \\ (ii)\ \Gamma_D \bigcup \Gamma_N \subset \partial\Omega \text{ and } |\Omega| = V_0 \end{array} \right\}$$

where $V_0$ is a fixed volume.

**Theorem 6.6 (D. Chenais).** The shape optimization problem

$$\inf_{\Omega \in \mathcal{U}_{ad}} J(\Omega)$$

admits at least one minimizer.

**Remark.** Condition $(i)$ implies a bound on the boundary curvature radius and prevents the creation of small holes.

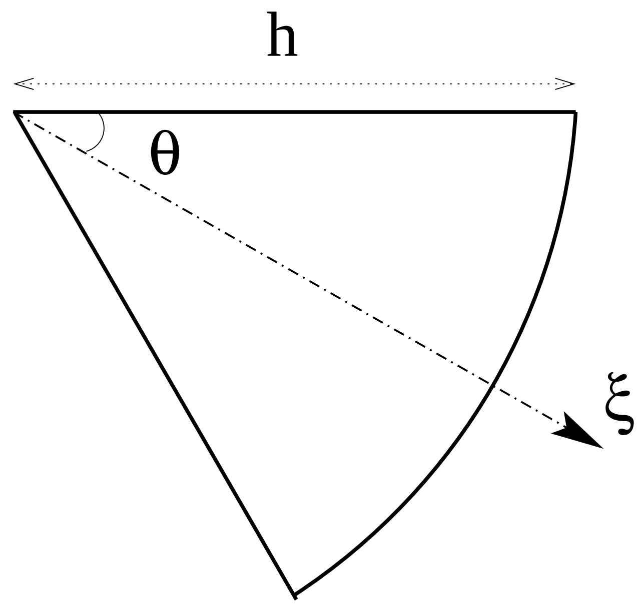
## Definition of a cone

Let $\theta \in ]0, \pi/2[$ be an angle, $h > 0$ a height, and $\xi \in \mathbb{R}^N$ a unit direction. A cone of angle $\theta$, height $h$ and direction $\xi$ is the open set

$$C(\theta, h, \xi) = \left\{ x \in \mathbb{R}^N \text{ such that } x \cdot \xi > \|x\| \cos\theta \text{ et } \|x\| < h \right\}.$$

For $y \in \mathbb{R}^N$, the cone of vertex $y$ is defined by

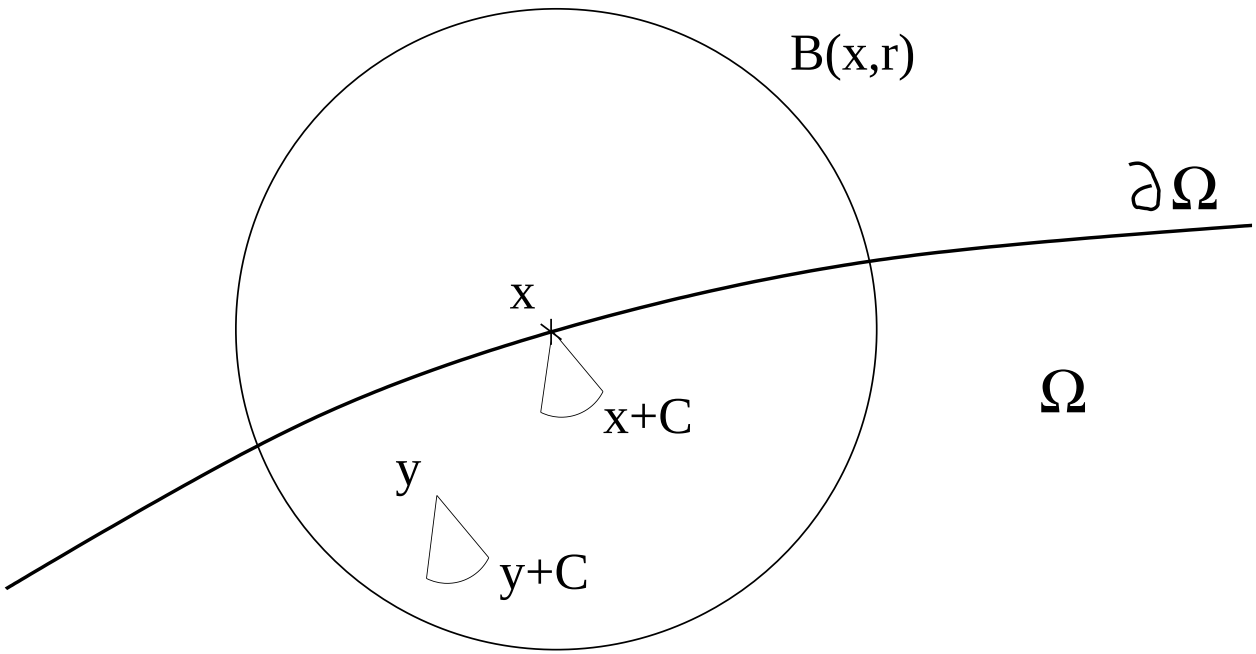
$$y + C(\theta, h, \xi) = \{ y + x \text{ such that } x \in C(\theta, h, \xi) \}.$$



## Uniform cone property

Let $\theta$ be an angle, $h > 0$ a height, and $r > 0$ a radius. An open set $\Omega$ is said to "**satisfy the uniform cone property**" if, for any $x \in \partial\Omega$, there exists a unit vector $\xi_x$ such that

$$\forall\, y \in B(x,r) \cap \Omega \quad y + C(\theta, h, \xi_x) \subset \Omega.$$



### 6.2.3 Existence under a topological condition (in dimension $N = 2$)

A working domain $D \subset \mathbb{R}^2$ is fixed. For any shape $\Omega \subset D$ we define its holes number, or more precisely, the number of connected components of its complementary

$$\#cc(D \setminus \Omega).$$

For a given integer $k$ and a volume $V_0$, we define

$$\mathcal{U}_{ad} = \left\{ \Omega \subset D \text{ such that } \begin{array}{l} (i)\ \#cc(D \setminus \Omega) \leq k \\ (ii)\ \Gamma_D \bigcup \Gamma_N \subset \partial\Omega \text{ and } |\Omega| = V_0 \end{array} \right\}$$

**Theorem 6.9 (V. Sverak, A. Chambolle).** The shape optimization problem

$$\inf_{\Omega \in \mathcal{U}_{ad}} J(\Omega)$$

admits at least one minimizer.

**Remark.** Condition $(i)$ prevents the creation of too many holes.

## 6.2.4 Existence under a regularity condition

Mathematical framework for shape deformation based on diffeomorphisms applied to a reference domain $\Omega_0$ (useful to compute a gradient too).

A space of diffeomorphisms (or smooth one-to-one map) in $\mathbb{R}^N$

$$\mathcal{T} = \left\{ T \text{ such that } (T - \text{ Id}) \text{ and } (T^{-1} - \text{ Id}) \in W^{1,\infty}(\mathbb{R}^N; \mathbb{R}^N) \right\}.$$

(They are perturbations of the identity Id: $x \to x$.)

**Definition of $W^{1,\infty}(\mathbb{R}^N; \mathbb{R}^N)$.** Space of Lipschitzian vectors fields:

$$\phi : \begin{cases} \mathbb{R}^N & \to & \mathbb{R}^N \\ x & \to & \phi(x) \end{cases}$$

$$\|\phi\|_{W^{1,\infty}(\mathbb{R}^N;\mathbb{R}^N)} = \sup_{x\in\mathbb{R}^N} \left( |\phi(x)|_{\mathbb{R}^N} + |\nabla\phi(x)|_{\mathbb{R}^{N\times N}} \right) < \infty$$

**Remark:** $\phi$ is continuous but its gradient is jut bounded. Actually, one can replace $W^{1,\infty}(\mathbb{R}^N; \mathbb{R}^N)$ by $C^1_b(\mathbb{R}^N; \mathbb{R}^N)$.

## Space of admissible shapes

Let $\Omega_0$ be a reference smooth open set.

$$\mathcal{C}(\Omega_0) = \left\{ \Omega \text{ such that there exists } T \in \mathcal{T}, \Omega = T(\Omega_0) \right\}.$$

- ☞ Each shape $\Omega$ is parametrized by a diffeomorphism $T$ (not unique !).
- ☞ All admissible shapes have the same topology.
- ☞ We define a pseudo-distance on $\mathcal{D}(\Omega_0)$

$$d(\Omega_1, \Omega_2) = \inf_{T \in \mathcal{T} \,|\, T(\Omega_1) = \Omega_2} \left( \|T - \mathrm{Id}\| + \|T^{-1} - \mathrm{Id}\| \right)_{W^{1,\infty}(\mathbb{R}^N;\mathbb{R}^N)}.$$

- ☞ If $\Omega_0$ is bounded, it is possible to use $C^1(\mathbb{R}^N;\mathbb{R}^N)$ instead of $W^{1,\infty}(\mathbb{R}^N;\mathbb{R}^N)$.

# Existence theory

Space of admissible shapes

$$\mathcal{U}_{ad} = \left\{ \Omega \in \mathcal{C}(\Omega_0) \text{ such that } \Gamma_D \bigcup \Gamma_N \subset \partial\Omega \text{ and } |\Omega| = V_0 \right\}.$$

For a fixed constant $R > 0$, we introduce the smooth subspace

$$\mathcal{U}_{ad}^{reg} = \{ \Omega \in \mathcal{U}_{ad} \text{ such that } d(\Omega, \Omega_0) \leq R, \ \} .$$

Interpretation: in practice, it is a "feasability" constraint.

**Theorem 6.11.** The shape optimization problem

$$\inf_{\Omega \in \mathcal{U}_{ad}^{reg}} J(\Omega)$$

admits at least one optimal solution.

**Remark.** All shapes share the same topology in $\mathcal{U}_{ad}$. Furthermore, the shape boundaries in $\mathcal{U}_{ad}^{reg}$ cannot oscillate too much.

## 6.3 Shape differentiation

**Goal:** to compute a derivative of $J(\Omega)$ by using the parametrization based on diffeomorphisms $T$.

We restrict ourselves to diffeomorphisms of the type

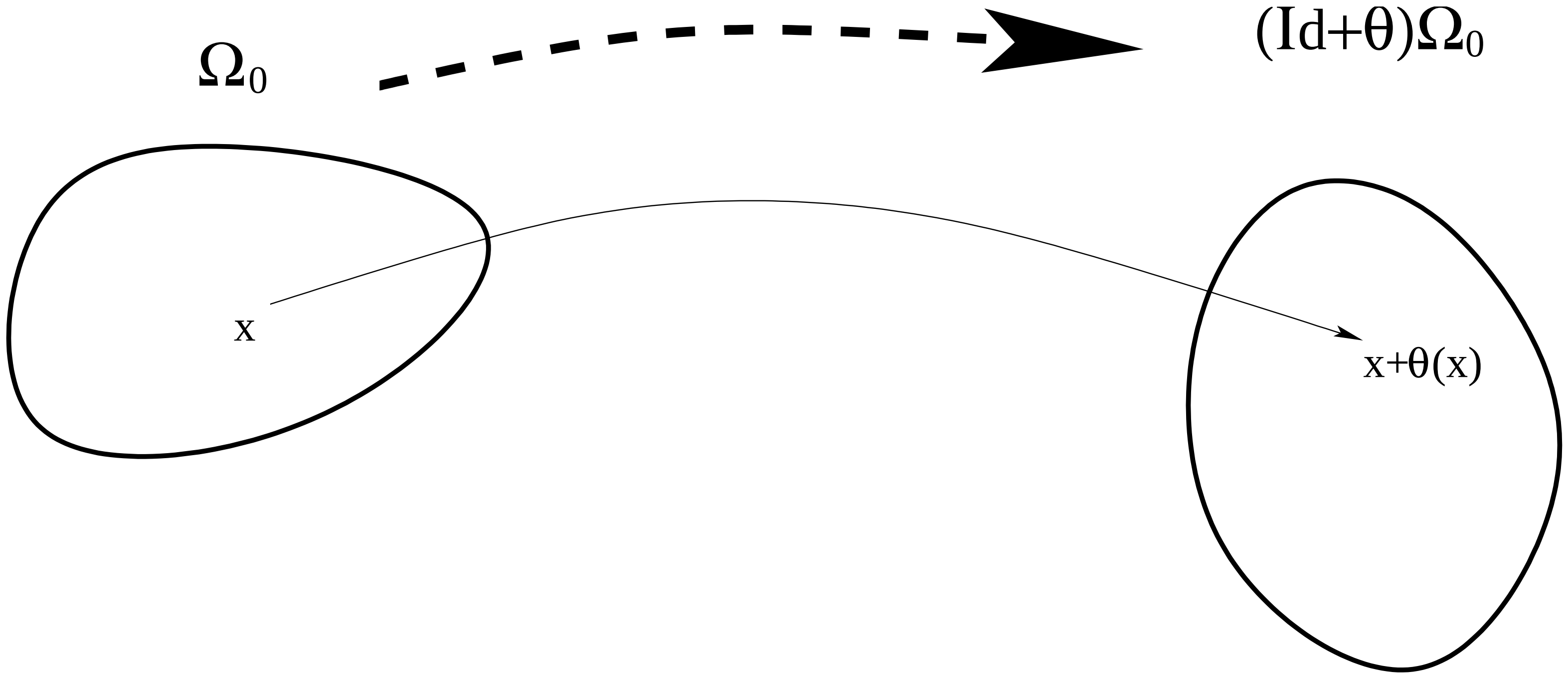
$$T = \mathrm{Id} + \theta \quad \text{with} \quad \theta \in W^{1,\infty}(\mathbb{R}^N; \mathbb{R}^N)$$

**Idea:** we differentiate $\theta \to J\big((\mathrm{Id} + \theta)\Omega_0\big)$ at 0.

**Remark.** This approach generalizes the Hadamard method of boundary shape variations along the normal: $\Omega_0 \to \Omega_t$ for $t \geq 0$

$$\partial\Omega_t = \left\{ x_t \in \mathbb{R}^N \mid \exists x_0 \in \partial\Omega_0 \mid x_t = x_0 + t\, g(x_0)\, n(x_0) \right\}$$

with a given incremental function $g$.




The shape $\Omega = (\,\mathrm{Id} + \theta)(\Omega_0)$ is defined by

$$\Omega = \{x + \theta(x) \mid x \in \Omega_0\}\,.$$

Thus $\theta(x)$ is a vector field which plays the role of the **displacement** of the reference domain $\Omega_0$.

**Lemma 6.13.** For any $\theta \in W^{1,\infty}(\mathbb{R}^N;\mathbb{R}^N)$ satisfying $\|\theta\|_{W^{1,\infty}(\mathbb{R}^N;\mathbb{R}^N)} < 1$, the map $T = \mathrm{Id} + \theta$ is one-to-one into $\mathbb{R}^N$ and belongs to the set $\mathcal{T}$.

**Proof.** Based on the formula

$$\theta(x) - \theta(y) = \int_0^1 (x-y) \cdot \nabla\theta\big(y + t(x-y)\big)\, dt\,,$$

we deduce that $|\ \theta(x) - \theta(y)\ | \leq \|\theta\|_{W^{1,\infty}(\mathbb{R}^N;\mathbb{R}^N)}\ |\ x - y\ |$ and $\theta$ is a strict contraction. Thus, $T = \mathrm{Id} + \theta$ is one-to-one into $\mathbb{R}^N$.

Indeed, $\forall b \in \mathbb{R}^N$ the map $K(x) = b - \theta(x)$ is a contraction and thus admits a unique fixed point $y$, i.e., $b = T(y)$ and $T$ is therefore one-to-one into $\mathbb{R}^N$.

Since $\nabla T = I + \nabla\theta$ (with $I = \nabla\,\mathrm{Id}$) and the norm of the matrix $\nabla\theta$ is strictly smaller than 1 ($\|I\| = 1$), the map $\nabla T$ is invertible. We then check that $(T^{-1} - \mathrm{Id}) \in W^{1,\infty}(\mathbb{R}^N;\mathbb{R}^N)$.

## Definition of the shape derivative

**Definition 6.15.** Let $J(\Omega)$ be a map from the set of admissible shapes $\mathcal{C}(\Omega_0)$ into $\mathbb{R}$. We say that $J$ is shape differentiable at $\Omega_0$ if the function

$$\theta \to J\Big((\,\mathrm{Id}+\theta)(\Omega_0)\Big)$$

is Fréchet differentiable at 0 in the Banach space $W^{1,\infty}(\mathbb{R}^N;\mathbb{R}^N)$, i.e., there exists a linear continuous form $L = J'(\Omega_0)$ on $W^{1,\infty}(\mathbb{R}^N;\mathbb{R}^N)$ such that

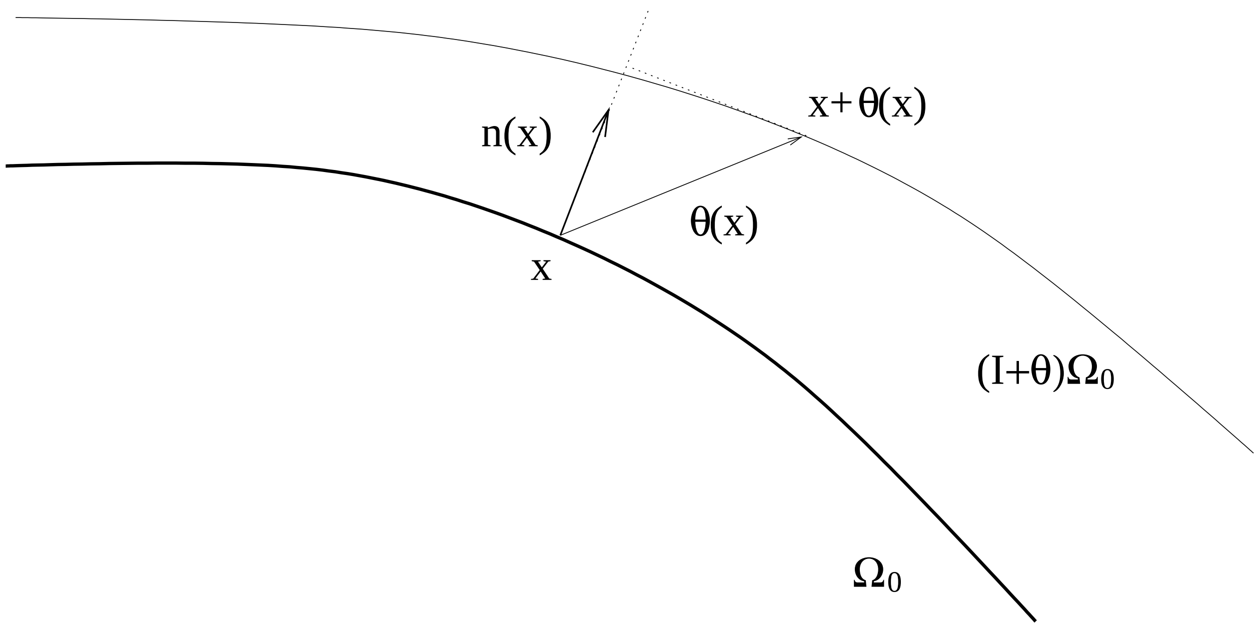
$$J\Big((\,\mathrm{Id}+\theta)(\Omega_0)\Big) = J(\Omega_0) + L(\theta) + o(\theta) \quad , \quad \text{with} \quad \lim_{\theta\to 0}\frac{|o(\theta)|}{\|\theta\|} = 0\,.$$

$J'(\Omega_0)$ is called the shape derivative and $J'(\Omega_0)(\theta)$ is a directional derivative.

The directional derivative $J'(\Omega_0)(\theta)$ depends only on the **normal component of $\theta$ on the boundary** of $\Omega_0$.

This surprising property is linked to the fact that the internal variations of the field $\theta$ does not change the shape $\Omega$, i.e.,

$$\theta \in C_c^1(\Omega)^N \text{ and } \|\theta\| << 1 \;\Rightarrow (\,\mathrm{Id} + \theta)\Omega = \Omega.$$



**Proposition 6.15.** Let $\Omega_0$ be a smooth bounded open set of $\mathbb{R}^N$. Let $J$ be a differentiable map at $\Omega_0$ from $\mathcal{C}(\Omega_0)$ into $\mathbb{R}$. Its directional derivative $J'(\Omega_0)(\theta)$ depends only on the normal trace on the boundary of $\theta$, i.e.

$$J'(\Omega_0)(\theta_1) = J'(\Omega_0)(\theta_2)$$

if $\theta_1, \theta_2 \in C^1(\mathbb{R}^N; \mathbb{R}^N)$ satisfy

$$\theta_1 \cdot n = \theta_2 \cdot n \quad \text{on } \partial\Omega_0.$$

**Proof.** Take $\theta = \theta_2 - \theta_1$ and introduce the solution of

$$\begin{cases} \frac{dy}{dt}(t) = \theta\big(y(t)\big) \\ y(0) = x \end{cases}$$

which satisfies

$$y(t + t', x, \theta) = y(t, y(t', x, \theta), \theta) \quad \text{for any } t, t' \in \mathbb{R}$$
$$y(\lambda t, x, \theta) = y(t, x, \lambda\theta) \quad \text{for any } \lambda \in \mathbb{R}$$

The we define the one-to-one map from $\mathbb{R}^N$ into $\mathbb{R}^N$, $x \ \rightarrow \ e^{\theta}(x) = y(1, x, \theta)$, the inverse of which is $e^{-\theta}$, $e^0 = \text{Id}$, and $t \rightarrow e^{t\theta}(x)$ is the solution of the o.d.e.

**Lemma 6.20.** Let $\theta \in W^{1,\infty}(\mathbb{R}^N;\mathbb{R}^N)$ be such that $\theta \cdot n = 0$ on $\partial\Omega_0$. Then $e^{t\theta}(\Omega_0) = \Omega_0$ for all $t \in \mathbb{R}$.

**Proof (by contradiction).** Assume $\exists x \in \Omega_0$ such that the trajectory $y(t,x)$ escapes from $\Omega_0$ (or conversely). Thus $\exists t_0 > 0$ such that $x_0 = y(t_0,x) \in \partial\Omega_0$.

Locally the boundary $\partial\Omega_0$ is parametrized by an equation $\phi(x) = 0$ and the normal is $n = n_0/|n_0|$ with $n_0 = \nabla\phi$ (defined around $\partial\Omega_0$).

We modify the vector field as $\tilde\theta = \theta - (\theta \cdot n)n$ to obtain a modified trajectory $\tilde y(t,x_0)$ such that, for any $t \geq t_0$,

$$\frac{d}{dt}\Big(\phi(\tilde y(t,x))\Big) = \frac{d\tilde y}{dt} \cdot \nabla\phi(\tilde y) = \tilde\theta(\tilde y) \cdot n|n_0| = 0$$

Since $\phi(\tilde y(t_0,x_0)) = 0$, we deduce $\phi(\tilde y(t,x_0)) = 0$, i.e., the trajectory $\tilde y$ stays on $\partial\Omega_0$. Since $\theta \cdot n = 0$ on $\partial\Omega_0$, $\tilde y$ is **also** a trajectory for the vector field $\theta$. Uniqueness of the o.d.e.'s solution yields $\tilde y(t) = y(t) \in \partial\Omega_0$ for any $t$ which is a contradiction with $x \in \Omega_0$.

**Remark.** The crucial point is that $\theta$ is tangent to the boundary $\partial\Omega_0$.

## Proof of Proposition 6.15 (Ctd.)

Since $e^{t\theta}(\Omega_0) = \Omega_0$ for any $t \in \mathbb{R}$, the function $J$ is constant along this path and

$$\frac{dJ\big(e^{t\theta}(\Omega_0)\big)}{dt}(0) = 0.$$

By the chain rule lemma we deduce

$$\frac{dJ\big(e^{t\theta}(\Omega_0)\big)}{dt}(0) = J'(\Omega_0)\left(\frac{de^{t\theta}}{dt}\right)(0) = J'(\Omega_0)\,(\theta) = 0,$$

because the path $e^{t\theta}(x)$ satisfies

$$\frac{de^{t\theta}(x)}{dt}(0) = \theta(x),$$

which yields the result by linearity in $\theta$.

## Review of known formulas

To compute shape derivatives we need to recall how to change variables in integrals.

**Lemma 6.21.** Let $\Omega_0$ be an open set of $\mathbb{R}^N$. Let $T \in \mathcal{T}$ be a diffeomorphism and $1 \leq p \leq +\infty$. Then $f \in L^p\big(T(\Omega_0)\big)$ if and only if $f \circ T \in L^p(\Omega_0)$, and

$$\int_{T(\Omega_0)} f\,dx = \int_{\Omega_0} f \circ T\,|\det \nabla T\,|\,dx$$

$$\int_{T(\Omega_0)} f\,|\det(\nabla T)^{-1}\,|\,dx = \int_{\Omega_0} f \circ T\,dx.$$

On the other hand, $f \in W^{1,p}\big(T(\Omega_0)\big)$ if and only if $f \circ T \in W^{1,p}(\Omega_0)$, and

$$\big(\nabla f\big) \circ T = \big((\nabla T)^{-1}\big)^t\,\nabla(f \circ T).$$

($^t$ = adjoint or transposed matrix)

Change of variables in a boundary integral.

**Lemma 6.23.** Let $\Omega_0$ be a smooth bounded open set of $\mathbb{R}^N$. Let $T \in \mathcal{T} \cap C^1(\mathbb{R}^N;\mathbb{R}^N)$ be a diffeomorphism and $f \in L^1\big(\partial T(\Omega_0)\big)$. Then $f \circ T \in L^1(\partial\Omega_0)$, and we have

$$\int_{\partial T(\Omega_0)} f\,ds = \int_{\partial\Omega_0} f \circ T \mid \det \nabla T \mid \ \left|\Big((\nabla T)^{-1}\Big)^t n\right|_{\mathbb{R}^N} ds,$$

where $n$ is the exterior unit normal to $\partial\Omega_0$.

## Examples of shape derivatives

**Proposition 6.22.** Let $\Omega_0$ be a smooth bounded open set of $\mathbb{R}^N$, $f(x) \in W^{1,1}(\mathbb{R}^N)$ and $J$ the map from $\mathcal{C}(\Omega_0)$ into $\mathbb{R}$ defined by

$$J(\Omega) = \int_\Omega f(x)\, dx.$$

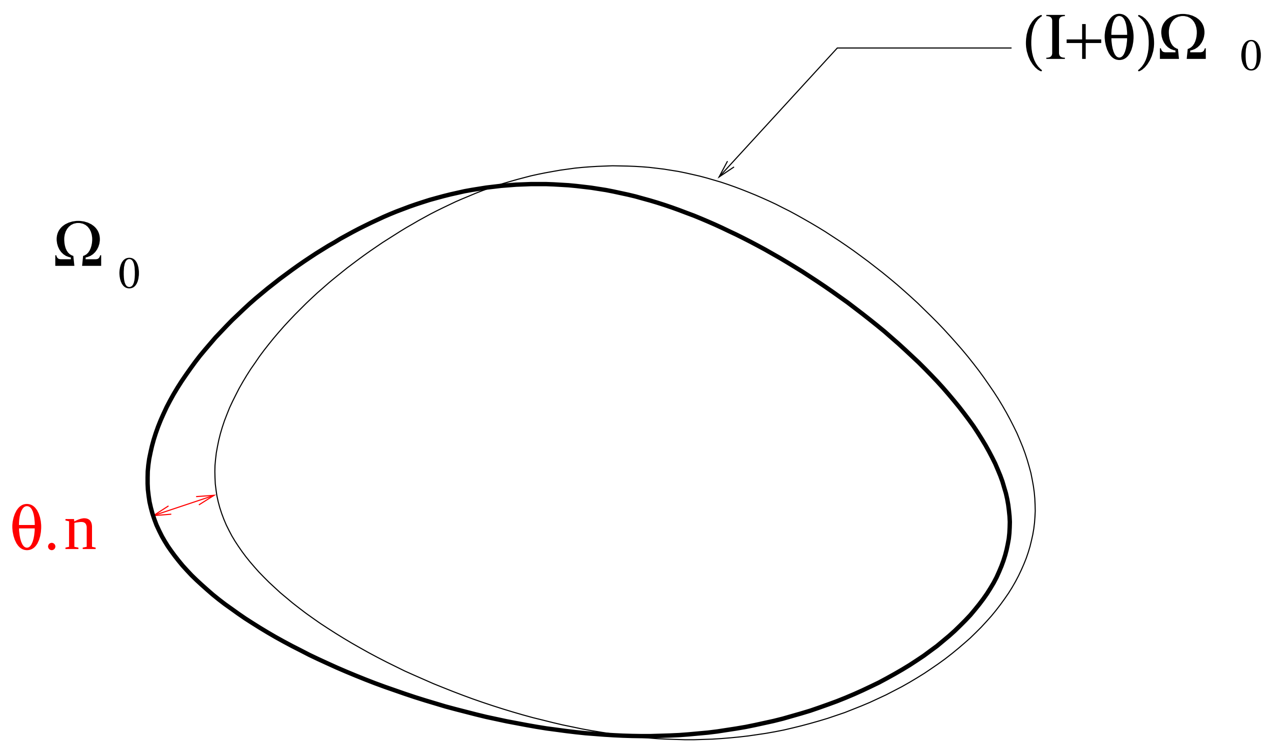
Then $J$ is shape differentiable at $\Omega_0$ and

$$J'(\Omega_0)(\theta) = \int_{\Omega_0} \operatorname{div}\big(\theta(x)\, f(x)\big)\, dx = \int_{\partial\Omega_0} \theta(x) \cdot n(x)\, f(x)\, ds$$

for any $\theta \in W^{1,\infty}(\mathbb{R}^N; \mathbb{R}^N)$.

**Remark.** To make sure the result is right, the safest way (but not the easiest) is to make a change of variables to get back to the reference domain $\Omega_0$.

## Intuitive proof




Surface swept by the transformation: difference between $(\,\mathrm{Id}+\theta)\Omega_0$ and $\Omega_0$ $\approx \partial\Omega_0 \times \Big(\theta\cdot n\Big)$. Thus

$$\int_{(\,\mathrm{Id}+\theta)\Omega_0} f(x)\,dx = \int_{\Omega_0} f(x)\,dx + \int_{\partial\Omega_0} f(x)\theta\cdot n\,ds + o(\theta).$$

**Proof.** We rewrite $J(\Omega)$ as an integral on the reference domain $\Omega_0$

$$J\big((\,\mathrm{Id}+\theta)\Omega_0\big)=\int_{\Omega_0} f\circ(\,\mathrm{Id}+\theta)\,|\det(\,\mathrm{Id}+\nabla\theta)\,|\,dx.$$

The functional $\theta\to\det(\,\mathrm{Id}+\nabla\theta)$ is differentiable from $W^{1,\infty}(\mathbb{R}^N;\mathbb{R}^N)$ into $L^\infty(\mathbb{R}^N)$ because

$$\det(\,\mathrm{Id}+\nabla\theta)=\det\,\mathrm{Id}+\,\mathrm{div}\theta+o(\theta)\quad\text{with}\quad\lim_{\theta\to 0}\frac{\|o(\theta)\|_{L^\infty(\mathbb{R}^N;\mathbb{R}^N)}}{\|\theta\|_{W^{1,\infty}(\mathbb{R}^N;\mathbb{R}^N)}}=0.$$

On the other hand, if $f(x)\in W^{1,1}(\mathbb{R}^N)$, the functional $\theta\to f\circ(\,\mathrm{Id}+\theta)$ is differentiable from $W^{1,\infty}(\mathbb{R}^N;\mathbb{R}^N)$ into $L^1(\mathbb{R}^N)$ because

$$f\circ(\,\mathrm{Id}+\theta)(x)=f(x)+\nabla f(x)\cdot\theta(x)+o(\theta)\quad\text{with}\quad\lim_{\theta\to 0}\frac{\|o(\theta)\|_{L^1(\mathbb{R}^N)}}{\|\theta\|_{W^{1,\infty}(\mathbb{R}^N;\mathbb{R}^N)}}=0.$$

By composition of these two derivatives we obtain the result.

**Proposition 6.24.** Let $\Omega_0$ be a smooth bounded open set of $\mathbb{R}^N$, $f(x) \in W^{2,1}(\mathbb{R}^N)$ and $J$ the map from $\mathcal{C}(\Omega_0)$ into $\mathbb{R}$ defined by

$$J(\Omega) = \int_{\partial\Omega} f(x)\, ds.$$

Then $J$ is shape differentiable at $\Omega_0$ and

$$J'(\Omega_0)(\theta) = \int_{\partial\Omega_0} \left(\nabla f \cdot \theta + f\big(\operatorname{div}\theta - \nabla\theta n \cdot n\big)\right) ds$$

for any $\theta \in W^{1,\infty}(\mathbb{R}^N;\mathbb{R}^N)$. By a (boundary) integration by parts this formula is equivalent to

$$J'(\Omega_0)(\theta) = \int_{\partial\Omega_0} \theta \cdot n \left(\frac{\partial f}{\partial n} + Hf\right) ds,$$

where $H$ is the mean curvature of $\partial\Omega_0$ defined by $H = \operatorname{div} n$.

## Interpretation

Two simple examples:

- ☞ If $\partial\Omega_0$ is an hyperplane, then $H = 0$ and the variation of the boundary integral is proportional to the normal derivative of $f$.
- ☞ If $f \equiv 1$, then $J(\Omega)$ is the perimeter (in 2-D) or the surface (in 3-D) of the domain $\Omega$ and its variation is proportional to the mean curvature.

**Proof.** A change of variable yields

$$J\big((\,\mathrm{Id}+\theta)\Omega_0\big) = \int_{\partial\Omega_0} f\circ(\,\mathrm{Id}+\theta)|\det(\,\mathrm{Id}+\nabla\theta)|\ \left|\,\big((\,\mathrm{Id}+\nabla\theta)^{-1}\big)^t n\,\right|_{\mathbb{R}^N} ds.$$

We already proved that $\theta\to\det(\,\mathrm{Id}+\nabla\theta)$ and $\theta\to f\circ(\,\mathrm{Id}+\theta)$ are differentiables.

On the other hand, $\theta\to\big((\,\mathrm{Id}+\nabla\theta)^{-1}\big)^t n$ is differentiable from $W^{1,\infty}(\mathbb{R}^N;\mathbb{R}^N)$ into $L^\infty(\partial\Omega_0;\mathbb{R}^N)$ because

$$\big((\,\mathrm{Id}+\nabla\theta)^{-1}\big)^t n = n-(\nabla\theta)^t n+o(\theta)\quad\text{with}\quad \lim_{\theta\to 0}\frac{\|o(\theta)\|_{L^\infty(\partial\Omega_0;\mathbb{R}^N)}}{\|\theta\|_{W^{1,\infty}(\mathbb{R}^N;\mathbb{R}^N)}}=0.$$

By composition with the derivative of $g\to|\,g\,|_{\mathbb{R}^N}$, we deduce

$$\left|\,\big((\,\mathrm{Id}+\nabla\theta)^{-1}\big)^t n\,\right|_{\mathbb{R}^N} = 1-(\nabla\theta)^t n\cdot n+o(\theta)\quad\text{with}\quad \lim_{\theta\to 0}\frac{\|o(\theta)\|_{L^\infty(\partial\Omega_0)}}{\|\theta\|_{W^{1,\infty}(\mathbb{R}^N;\mathbb{R}^N)}}=0.$$

Composing these three derivatives leads to the result. The formula, including the mean curvature, is obtained by an integration by parts on the surface $\partial\Omega_0$.