OPTIMAL DESIGN OF STRUCTURES (MAP 562)

CHAPTER V

PARAMETRIC (OR SIZING) OPTIMIZATION

S. Amstutz, B. Bogosel

Original version by G. Allaire

Academic year 2019-2020

Ecole Polytechnique
Department of applied mathematics

Optimization of a membrane thickness

Consider a membrane occupying a bounded domain Ω in $\mathbb{R}^N$, with forces $f \in L^2(\Omega)$ and displacement $u \in H_0^1(\Omega)$ solution of

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

The variable is the thickness h. It is called parametric (or sizing) optimization because the computational domain Ω is fixed. The thickness $h(x)$ is just a parameter.

The admissible set is defined by

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

Parametric shape optimization problem:

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

where u depends on h through the state equation, and j is a C^1 function from $\mathbb{R}$ to $\mathbb{R}$ such that $|j(u)| \leq C(u^2 + 1)$ and $|j'(u)| \leq C(|u| + 1)$.

Examples:

- Compliance or work done by the load (a measure of rigidity)

$$j(u) = fu$$

- Least square criterion to reach a target displacement $u_0 \in L^2(\Omega)$

$$j(u) = |u - u_0|^2$$

Continuity of the cost function

Proposition. The application

$$h \to J(h) = \int_\Omega j(u)\, dx$$

is continuous from $\mathcal{U}_{ad}$ (equipped with the L^∞ norm) into $\mathbb{R}$.

Proof. By composition of the 2 continuous functions below.

Lemma. The map $\hat{u} \to \int_\Omega j(\hat{u})\, dx$ is continuous from $L^2(\Omega)$ into $\mathbb{R}$.

Proof. By using the Lebesgue dominated convergence theorem.

Lemma. The map $h \to u$ is continuous from $\mathcal{U}_{ad}$ into $H_0^1(\Omega)$.

Proof

Let $h_n \in \mathcal{U}_{ad}$ be a sequence such that $\|h_n - h_\infty\|_{L^\infty(\Omega)} \to 0$. Let u_n be the unique solution in $H_0^1(\Omega)$ of the membrane equation with the associated thickness h_n

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

$$\Leftrightarrow \quad \int_\Omega h_n \nabla u_n \cdot \nabla \phi \, dx = \int_\Omega f \phi \, dx \quad \forall \phi \in H_0^1(\Omega).$$

We substract the variational formulation for u_m to that for u_n

$$\int_\Omega h_n \nabla(u_n - u_m) \cdot \nabla \phi \, dx = \int_\Omega (h_m - h_n) \nabla u_m \cdot \nabla \phi \, dx \quad \forall \phi \in H_0^1(\Omega).$$

Choosing $\phi = u_n - u_m$ we deduce

$$\|\nabla(u_n - u_m)\|_{L^2(\Omega)} \leq \frac{C}{h_{min}^2} \|f\|_{L^2(\Omega)} \|h_m - h_n\|_{L^\infty(\Omega)},$$

which proves that u_n is a Cauchy sequence in $H_0^1(\Omega)$ and thus converges.

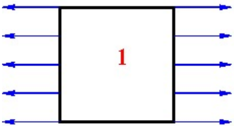
Existence theories

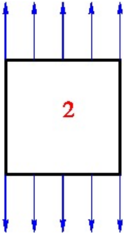
- None of the theorems studied in the chapter on optimization applies in general !
- **Usually there exists no optimal shape !**
- It is an important issue because this non-existence phenomenon has dramatic consequences for the numerical computations.
- Possible remedies: the definition of the set $\mathcal{U}_{ad}$ of admissible designs has to be modified to obtain existence.
    1. Discretization: finite dimensional admissible set.
    2. Regularization: compact admissible set.
    3. A “miracle”: compliance minimization is a convex problem.

Generic non-existence of optimal shapes

- There are precise mathematical counter-examples (a bit complicated).
- It shows up numerically: non convergence, instabilities...

Intuitive counter-example (which can be rigorously justified) with 2 state equations:





One seeks an elastic plate which is

1. strong for the horizontal loading 1,
2. weak for the vertical loading 2.

Remark: the elastic plate problem is more intuitive, but similar arguments apply to the membrane problem.

Definition of the counter-example

Loads:

$$\begin{cases} -\mathrm{div}\sigma_1 = 0 & \text{in } \Omega, \\ \sigma_1 n = (e_1 \cdot n) e_1 & \text{on } \partial\Omega, \end{cases} \qquad \begin{cases} -\mathrm{div}\sigma_2 = 0 & \text{in } \Omega, \\ \sigma_2 n = (e_2 \cdot n) e_2 & \text{on } \partial\Omega, \end{cases}$$

$$\sigma_i = h\left(2\mu e(u_i) + \lambda\, \mathrm{tr} e(u_i)\, \mathrm{Id}\right), \qquad e(u_i) = \frac{1}{2}\left(\nabla u_i + (\nabla u_i)^t\right)$$

Objective:

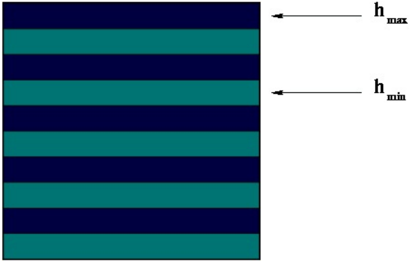
$$\inf_{h \in \mathcal{U}_{ad}} J(h) = \int_{\partial\Omega} (e_1 \cdot n)\,(e_1 \cdot u_1)\, ds - \int_{\partial\Omega} (e_2 \cdot n)\,(e_2 \cdot u_2)\, ds$$

We minimize the compliance in the e_1 direction and we maximize it in the e_2 direction.

The same plate is subjected to the 2 loadings (not simultaneously).

Hand-waving argument

If h is uniform $\Rightarrow$ isotropic material $\Rightarrow$ same mechanical behavior in all directions, thus not optimal.



It is better to build horizontal layers of alternating small and large thicknesses:

$\Rightarrow$ laminated structure which is horizontally strong and vertically weak.

Hand-waving argument (continued)

- Vertically, the efforts of traction are transmitted by crossing the layers of minimal thickness: the structure is thus weak.
- Horizontally, the efforts of traction are transmitted along the layers of maximal thickness: the structure is thus strong.
- **However**, since the boundary conditions are uniform, the plate is horizontally stronger if the layers are **finer** because each loading point is closer to a thick layer.

If h oscillates at a small scale, we obtain an anisotropic composite material.

To reach the minimum the oscillation scale must **go to 0**.

Therefore, there does not exist an optimal design !

Existence for a discretized model

We go back to the membrane problem (but the same holds for the elastic plate).

Let $(\omega_i)_{1\le i\le n}$ be a partition of Ω such that

$$\overline{\Omega} = \bigcup_{i=1}^{n} \overline{\omega}_i, \quad \omega_i \cap \omega_j = \emptyset \text{ for } i \neq j.$$

We introduce the subspace $\mathcal{U}_{ad}^n$ of $\mathcal{U}_{ad}$ defined by

$$\mathcal{U}_{ad}^n = \left\{ h \in \mathcal{U}_{ad} \ , \quad h(x) = h_i = C^{st} \text{ in } \omega_i,\ 1 \le i \le n \right\}.$$

Any function $h(x) \in \mathcal{U}_{ad}^n$ is uniquely characterized by a vector $(h_i)_{1\le i\le n} \in \mathbb{R}^n$: $\mathcal{U}_{ad}^n$ is thus identified to a subspace of $\mathbb{R}^n$.
We are now back to the finite dimensional case. It is much easier !

Theorem (finite dimension). The optimization problem

$$\inf_{h \in \mathcal{U}_{ad}^n} J(h)$$

admits at least one minimizer.
Proof. $\mathcal{U}_{ad}^n$ is a compact subspace of $\mathbb{R}^n$ and $J(h)$ is a continuous function on $\mathcal{U}_{ad}^n$.

Remark. What happens when $n \to \infty$? Numerically, local or global minimizers ? Conclusion: **theorem of limited interest**.

Existence with a regularity constraint

Consider the space $C^1(\overline{\Omega})$ which is a Banach space for the norm

$$\|\phi\|_{C^1(\overline{\Omega})} = \max_{x\in\overline{\Omega}} \left(|\phi(x)| + |\nabla\phi(x)|\right).$$

Take a given constant $R > 0$, and introduce the subspace $\mathcal{U}_{ad}^{reg}$

$$\mathcal{U}_{ad}^{reg} = \left\{ h \in \mathcal{U}_{ad} \cap C^1(\overline{\Omega}) \,, \quad \|h\|_{C^1(\overline{\Omega})} \le R \right\}.$$

Interpretation: "feasibility" constraint because, in practice, the thickness cannot rapidly vary.

Theorem. The optimization problem

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

admits at least one minimizer.

Proof. Consider a minimizing sequence $(h_n)_{n\geq 1}$

$$\lim_{n\to\infty} J(h_n) = \left(\inf_{h\in\mathcal{U}_{ad}^{reg}} J(h) \right).$$

By definition, the sequence h_n is bounded (uniformly in n) in the space $C^1(\overline{\Omega})$. We then apply a variant of Rellich theorem which states that one can extract a subsequence (still denoted by h_n for simplicity) which converges in $C^0(\overline{\Omega})$ towards a limit function h_∞ (furthermore $h_\infty \in C^1(\overline{\Omega})$). We already know that the map $h \to J(h)$ is continuous from $\mathcal{U}_{ad}$ into $\mathbb{R}$, thus

$$\lim_{n\to\infty} J(h_n) = J(h_\infty),$$

which proves that h_∞ is a global minimizer of J in $\mathcal{U}_{ad}^{reg}$.

Theorem of limited practical interest.

- How to choose the upper bound R in the definition of $\mathcal{U}_{ad}^{reg}$?
- Usually, no convergence when R goes to infinity.
- Numerically, global or local minimizers ?
- Numerically, the following regularity constraint is preferred

$$\|h\|_{H^1(\Omega)} \leq R.$$

Computation of a continuous gradient

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

We consider the open set

$$\mathcal{U} = \left\{ h \in L^\infty(\Omega)\,, \quad \exists h_0 > 0 \text{ such that } h(x) \geq h_0 \text{ in } \Omega \right\}.$$

Lemma. The application $h \to u(h)$, which gives the solution $u(h) \in H_0^1(\Omega)$ for $h \in \mathcal{U}$, is differentiable and its directional derivative at h in the direction $k \in L^\infty(\Omega)$ is given by

$$u'(h)(k) = v,$$

where v is the unique solution in $H_0^1(\Omega)$ of

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

Proof. We admit the differentiability (implicit function theorem) and calculate directional derivatives.
Let $\hat{u}(t)$ be the solution for the thickness $h(t) := h + tk$.
We differentiate

$$\int_\Omega h(t)\nabla\hat{u}(t)\cdot\nabla\varphi\,dx = \int_\Omega f\varphi \qquad \forall\varphi\in H_0^1(\Omega),$$

and obtain

$$\int_\Omega h'(t)\nabla\hat{u}(t)\cdot\nabla\varphi\,dx + \int_\Omega h(t)\nabla\hat{u}'(t)\cdot\nabla\varphi\,dx = 0 \qquad \forall\varphi\in H_0^1(\Omega).$$

For $t = 0$,

$$\int_\Omega k\nabla u\cdot\nabla\varphi\,dx + \int_\Omega h\nabla\hat{u}'(0)\cdot\nabla\varphi\,dx = 0 \qquad \forall\varphi\in H_0^1(\Omega).$$

Then $\hat{u}'(0) = u'(h)(k)$ is identified with v.

Lemma. For $h \in \mathcal{U}$, let $u(h)$ be the state in $H_0^1(\Omega)$ and

$$J(h) = \int_\Omega j\big(u(h)\big)\, dx \ ,$$

where j is a C^1 function from $\mathbb{R}$ into $\mathbb{R}$ such that $|j(u)| \leq C(u^2+1)$ and $|j'(u)| \leq C(|u|+1)$ for any $u \in \mathbb{R}$. The application $h \mapsto J(h)$, from $\mathcal{U}$ into $\mathbb{R}$, is differentiable and its derivative at h in the direction $k \in L^\infty(\Omega)$ is given by

$$\langle J'(h), k\rangle = \int_\Omega j'\big(u(h)\big) v\, dx \ ,$$

where $v = u'(h)(k)$ is the unique solution in $H_0^1(\Omega)$ of

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

Proof. From the assumptions the derivative of $u \mapsto \int_\Omega j(u)dx$ is $w \mapsto \int_\Omega j'(u)w dx$. Then we use the chain rule for differentiable functions.

Adjoint state

We introduce an adjoint state p defined as the unique solution in $H_0^1(\Omega)$ of

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

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

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

If $h \in \mathcal{U}_{ad}$ is a local minimizer of J in $\mathcal{U}_{ad}$, then it satisfies the necessary optimality condition

$$\int_\Omega \nabla u \cdot \nabla p \, (k - h) \, dx \geq 0$$

for any $k \in \mathcal{U}_{ad}$.

Proof. To make explicit $J'(h)$ from the lemma, we must eliminate $v = u'(h)(k)$. We use the adjoint state for that: testing the variational formulation for p against v and that for v against p, we obtain

$$\int_\Omega h\nabla p \cdot \nabla v \, dx = -\int_\Omega j'(u) v \, dx$$

$$\int_\Omega h\nabla v \cdot \nabla p \, dx = -\int_\Omega k\nabla u \cdot \nabla p \, dx.$$

Comparing these two equalities we infer

$$\langle J'(h), k\rangle = \int_\Omega j'(u) v \, dx = \int_\Omega k\nabla u \cdot \nabla p \, dx,$$

for any $k \in L^\infty(\Omega)$. Since $\nabla u \cdot \nabla p$ belongs to $L^1(\Omega)$, we check that $J'(h)$ is continuous on $L^\infty(\Omega)$.

How to find the adjoint state

For independent variables $(\hat{h}, \hat{u}, \hat{p}) \in L^\infty(\Omega) \times H^1_0(\Omega) \times H^1_0(\Omega)$, we introduce the Lagrangian

$$\mathcal{L}(\hat{h}, \hat{u}, \hat{p}) = \int_\Omega j(\hat{u})\,dx + \int_\Omega \left(\hat{h}\nabla\hat{p}\cdot\nabla\hat{u} - f\hat{p}\right)\,dx.$$

Note that integration by parts gives formally

$$\mathcal{L}(\hat{h}, \hat{u}, \hat{p}) = \int_\Omega j(\hat{u})\,dx + \int_\Omega \hat{p}\left(-\operatorname{div}\left(\hat{h}\nabla\hat{u}\right) - f\right)\,dx,$$

hence $\hat{p}$ is a Lagrange multiplier (here a function) for the constraint which connects u to h.
The partial derivative of $\mathcal{L}$ with respect to u in the direction $\phi \in H^1_0(\Omega)$ is

$$\langle\frac{\partial\mathcal{L}}{\partial u}(\hat{h}, \hat{u}, \hat{p}), \phi\rangle = \int_\Omega j'(\hat{u})\phi\,dx + \int_\Omega \left(\hat{h}\nabla\hat{p}\cdot\nabla\phi\right)\,dx,$$

which, when it vanishes, is nothing else than the variational formulation of the adjoint equation.

A simple formula for the derivative

The Lagrangian yields the following formula

$$\boxed{J'(h) = \frac{\partial \mathcal{L}}{\partial h}(h, u, p)}$$

with the state u and the adjoint p.
This is not a surprise ! Indeed,

$$J(h) = \mathcal{L}(h, u, \hat{p}) \quad \forall \hat{p}$$

because u is the state. Thus, if $u(h)$ is differentiable, we get

$$\langle J'(h), k\rangle = \langle \frac{\partial \mathcal{L}}{\partial h}(h, u, \hat{p}), k\rangle + \langle \frac{\partial \mathcal{L}}{\partial u}(h, u, \hat{p}), \frac{\partial u}{\partial h}(k)\rangle.$$

Then, taking $\hat{p} = p$, the adjoint, we obtain

$$\langle J'(h), k\rangle = \langle \frac{\partial \mathcal{L}}{\partial h}(h, u, p), k\rangle.$$

The self-adjoint case: 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

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

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 irrelevant.

An existence result

We rewrite the problem under the form

$$\inf_{(h,\tau)\in\mathcal{U}_{ad}\times H} \int_\Omega h^{-1}|\tau|^2 dx \ ,$$

with $H = \{\tau \in L^2(\Omega)^N \, , \ -\mathrm{div}\tau = f \text{ in } \Omega\}$.
Lemma. The function $\phi(a,\sigma) = a^{-1}|\sigma|^2$, defined from $\mathbb{R}^+ \times \mathbb{R}^N$ into $\mathbb{R}$, is convex and satisfies

$$\phi(a,\sigma) = \phi(a_0,\sigma_0) + \phi'(a_0,\sigma_0)\cdot(a-a_0,\sigma-\sigma_0) + \underbrace{\phi(a,\sigma-\frac{a}{a_0}\sigma_0)}_{\geq 0},$$

where the derivative is given by

$$\phi'(a_0,\sigma_0)\cdot(b,\tau) = -\frac{b}{a_0^2}|\sigma_0|^2 + \frac{2}{a_0}\sigma_0\cdot\tau.$$

Theorem. There exists a minimizer to the shape optimization problem.

Sketch of proof. We consider the formulation

$$\inf_{(h,\tau)\in\mathcal{U}_{ad}\times H}\int_\Omega h^{-1}|\tau|^2dx = \inf_{\tau\in H}\left\{\mathcal{I}(\tau) := \inf_{h\in\mathcal{U}_{ad}}\int_\Omega h^{-1}|\tau|^2dx\right\}.$$

$\mathcal{I}$ is convex as a marginal of convex bivariate function.
H is a closed subset of a Hilbert space.
Finally,

$$\mathcal{I}(\tau) \geq h_{max}^{-1}\int_\Omega |\tau|^2dx.$$

Hence there exists an optimal τ.
For each τ there exists a unique optimal h (see next lemma).

Optimality conditions

Lemma. Take $\tau \in L^2(\Omega)^N$. The problem

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

admits a unique 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
$\displaystyle\int_\Omega h(x)\, dx = h_0 |\Omega|$.

Proof. The function $h \to \int_\Omega h^{-1} |\tau|^2 dx$ is strictly convex from $\mathcal{U}_{ad}$ into $\mathbb{R}$ and we easily find the stationary point of the Lagrangian

$$\int_\Omega h^{-1} |\tau|^2 dx + \ell \left(\int_\Omega h(x)\, dx - h_0 |\Omega| \right).$$

Discrete approach

Is the problem simpler after discretization ?

Applying a finite element method, the equation becomes a linear system of order n

$$K(h)y(h) = b$$

where $K(h)$ is the **stiffness matrix** of the membrane (which depends on h), b the right hand side of the forces f, $y(h)$ the vector of the coordinates of the solution u in the finite element basis (of dimension n). We also discretize h

$$\mathcal{U}_{ad}^{disc} = \left\{ h \in \mathbb{R}^n \,, \quad h_{max} \geq h_i \geq h_{min} > 0, \sum_{i=1}^{n} c_i h_i = h_0 |\Omega| \right\},$$

where $\sum_{i=1}^{n} c_i h_i$ is an approximation of $\int_\Omega h(x)dx$.

Approximating the cost function, the discrete problem is

$$\inf_{h \in \mathcal{U}_{ad}^{disc}} \left\{ J^{disc}(h) = j^{disc}(y(h)) \right\},$$

where j^{disc} is a smooth approximation of j from $\mathbb{R}^n$ into $\mathbb{R}$. In the case of the compliance

$$j^{disc}(y(h)) = b \cdot y(h) = K(h)^{-1} b \cdot b.$$

In the case of a least square criteria for a target displacement

$$j^{disc}(y(h)) = B(y(h) - y_0) \cdot (y(h) - y_0).$$

Practical question: how to compute the gradient $J^{disc}(h)$?
Applications: optimality conditions, numerical method of minimization.

A naive idea

Explicit formula: $y(h) = K(h)^{-1}b$, thus

$$\left(J^{disc}\right)'(h) = (y'(h))^t \left(j^{disc}\right)'(y(h))$$

$$\text{with} \quad y'(h) = -K(h)^{-1}K'(h)K(h)^{-1}b.$$

Notations: $f'(h) = (\partial f(h)/\partial h_i)_{1\leq i\leq n}$.
Inoperative because one must solve $n+1$ linear systems with the matrix $K(h)$ to obtain all components of $y'(h)$. Recall that $K(h)$ is a very large matrix (of size n) and its inverse is **never** explicitly computed.
As a consequence, **we do not use** the explicit formula $y(h) = K(h)^{-1}b$. We rather use an adjoint method.

Adjoint state

We define the adjoint state $p \in \mathbb{R}^n$ solution of

$$K(h)p(h) = -\left(j^{disc}\right)'(y(h)).$$

Taking the scalar product of $K(h)y'(h) = -K'(h)y(h)$ with $p(h)$ and that of $K(h)p(h) = -\left(j^{disc}\right)'(y(h))$ with $y'(h)$, we obtain by symmetry of $K(h)$, for each component i,

$$-\left(j^{disc}\right)'(y(h))\cdot\frac{\partial y}{\partial h_i}(h) = K(h)p(h)\cdot\frac{\partial y}{\partial h_i}(h) = -\frac{\partial K}{\partial h_i}(h)y(h)\cdot p(h),$$

from which we obtain

$$\left(J^{disc}\right)'(h) = K'(h)y(h)\cdot p(h) = \left(\frac{\partial K}{\partial h_i}(h)y(h)\cdot p(h)\right)_{1\leq i\leq n}.$$

In practice, this is the very formula that we use for evaluating the gradient $\left(J^{disc}\right)'(h)$ since it requires only two solutions of linear systems.

Conclusion

There is no practical simplification on using a discrete approach rather than a continuous one.

Some authors prefer to discretize first, optimize afterwards.

- It guarantees a perfect consistency between the gradient and the cost function (exact discrete descent direction).
- But it requires a deep knowledge of the numerical solver (almost impossible if one has not written himself the source code !).
- Optimization can be performed accurately, but on a model containing intrinsic discretization errors.

Here, we follow another philosophy: first optimize (at least differentiate) in a continuous framework, then discretize.

- Derivatives are calculated on the continuous model, independently of the discretization scheme.
- The optimization method may be designed and analyzed on the continuous model (robustness against mesh refinement).
- Analysis tools in infinite dimension are required.

Numerical algorithm: projected gradient

We differentiated J on the open subset of $L^\infty(\Omega)$

$$\mathcal{U} = \{h \in L^\infty(\Omega)\,, \quad \exists h_0 > 0 \text{ such that } h(x) \geq h_0 \text{ in } \Omega\}\,.$$

But $L^\infty(\Omega)$ is not a Hilbert space...
We define the gradient of J and the projection on $\mathcal{U}_{ad}$ through the L^2 scalar product. Consistency will be retrieved after discretization.

Projected gradient algorithm

1. Initialization of the thickness $h_0 \in \mathcal{U}_{ad}$ (for 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 with the thickness h_n).

To make the algorithm fully explicit, we have to specify 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 ℓ is the unique Lagrange multiplier such that

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

The determination of the constant ℓ 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 and continuous 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

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

then we proceed by dichotomy to find the root ℓ.

- 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 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 stiffness matrix).
- Convergence is detected when the optimality condition is satisfied with a threshold $\epsilon > 0$

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

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

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

We can rewrite the optimization problem as a double minimization

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

The problem is convex and admits a minimizer.

We recall:

Lemma (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 $n \geq 0$:
   2.1 Computation of the state τ_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.
   2.2 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 τ is equivalent to solving the equation

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

and we recover τ_n by the formula

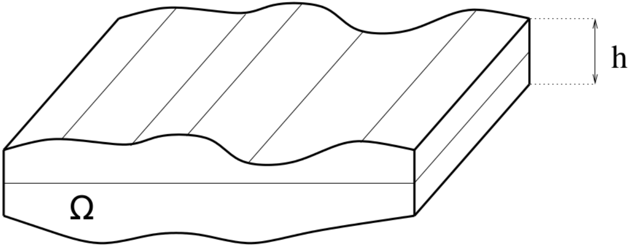
$$\tau_n = h_n \nabla u_n.$$

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

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

This algorithm can also be interpreted as an optimality criteria method (fixed point on the optimality conditions).

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 $e(u) = \frac{1}{2}\left(\nabla u + (\nabla u)^t\right)$.

Set of admissible thickness:

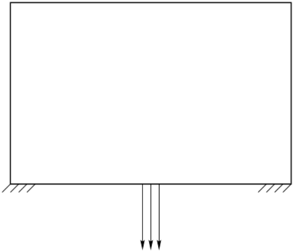
$$\mathcal{U}_{ad} = \left\{ h \in L^{\infty}(\Omega)\ , \quad h_{max} \geq h(x) \geq h_{min} > 0 \text{ in } \Omega, \right.$$
$$\left. \int_{\Omega} h(x)\, dx = h_0 |\Omega| \right\}.$$

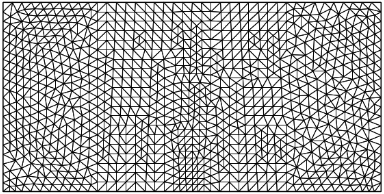
The compliance optimization can be written

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

The theoretical results are the same.
We apply the optimality criteria method.

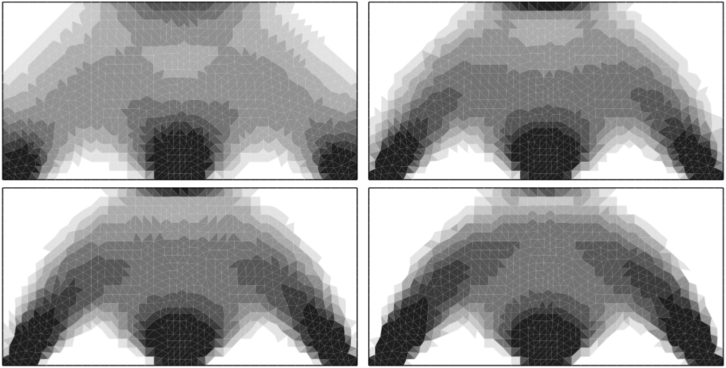
Boundary conditions and mesh for an elastic plate





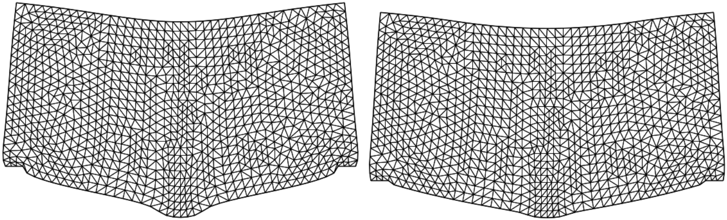
FreeFem++ computations ; scripts available on the web page
http://www.cmap.polytechnique.fr/~allaire/cours_X_annee3.html

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

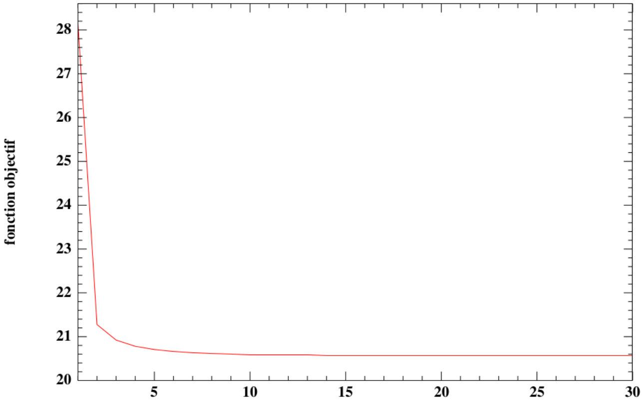


$h_{min} = 0.1$, $h_{max} = 1.0$, $h_0 = 0.5$ (increasing thickness from white to black)

Comparing the initial and final deformed shapes

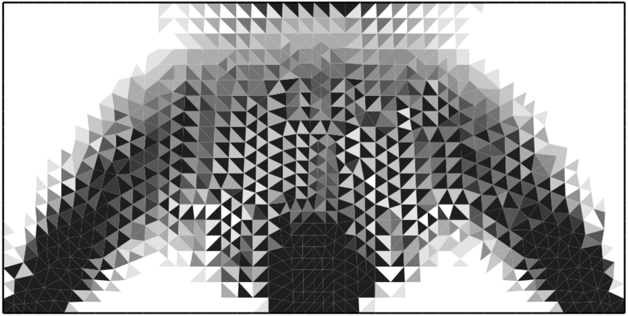


Convergence history



Numerical instabilities (checkerboards)

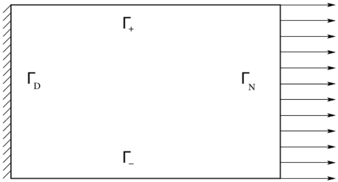
- Finite elements $P2$ for u and $P0$ for $h \Rightarrow$ OK
- Finite elements $P1$ for u and $P0$ for $h \Rightarrow$ unstable !
- Finite elements $P1$ for u and $P1$ for $h \Rightarrow$ OK



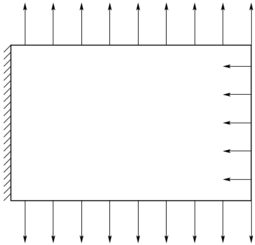
$P1$ / $P0$ finite elements.

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.

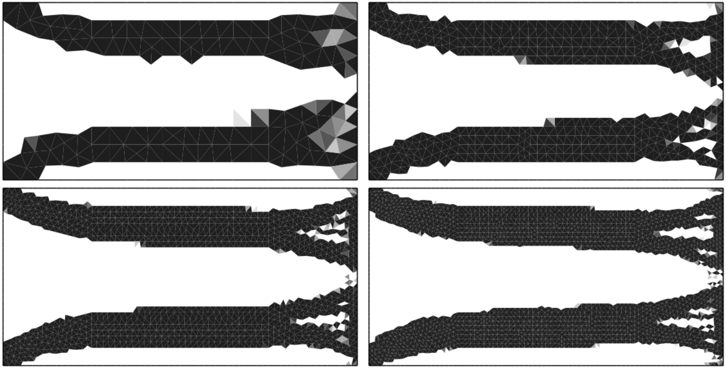


boundary conditions



target displacement

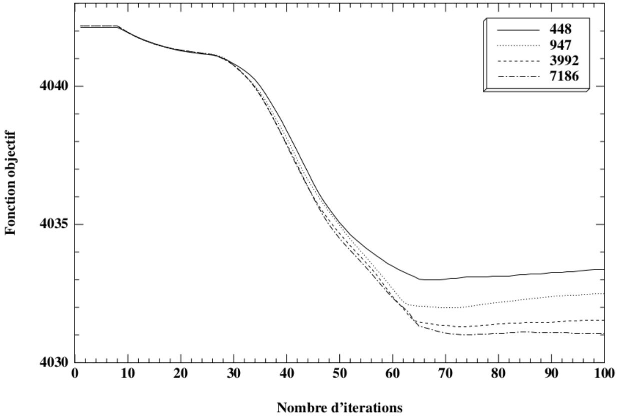
Optimal shapes for meshes with 448, 947, 3992, 7186 triangles



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.



Regularization

In order to avoid numerical instabilities and obtain less mesh-dependent results it is common to introduce filters.

- Filter of the thickness, considering for instance

  $$\tilde{J}(h) = J(\phi * h)$$

  for some convolution kernel ϕ.

- H^1 penalization of the thickness

  $$\tilde{J}(h) = J(h) + \frac{\alpha}{2} \int_\Omega |\nabla h|^2 \, dx$$

  and use of an H^1 gradient.

Consider this latter case. We have the derivative

$$\langle J'(h), k\rangle = \int_\Omega k\nabla u\cdot\nabla p\,dx + \alpha\int_\Omega \nabla h\cdot\nabla k\,dx \quad \forall\, k\in L^\infty(\Omega)\cap H^1(\Omega).$$

The (weighted) H^1 scalar product (with weight $\beta > 0$) is

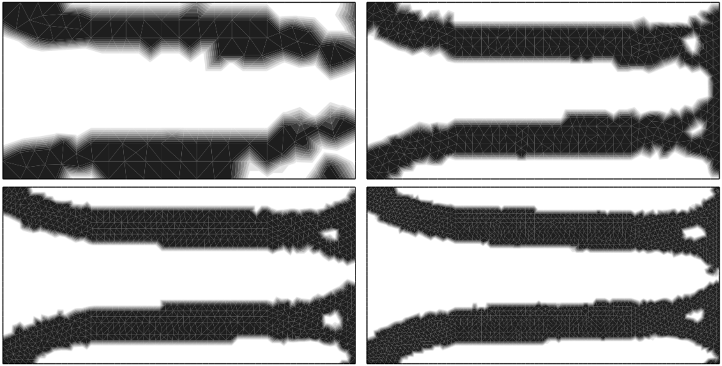
$$\langle J'(h), k\rangle_{H^1} = \int_\Omega \big(\beta\nabla J'(h)\cdot\nabla k + J'(h)k\big)\,dx\ .$$

We obtain a new formula for the gradient

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

Unfortunately the H^1 projection onto $\mathcal{U}_{ad}$ is quite involved (linear complementarity problem). In practice we choose β "small" and use the L^2 projection.

Regularized optimal shapes



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