OPTIMAL DESIGN OF STRUCTURES (MAP 562)

G. ALLAIRE

January 7th, 2015

Department of Applied Mathematics, Ecole Polytechnique

CHAPTER IV

OPTIMAL CONTROL

Optimization of distributed systems:
Computing a gradient by the adjoint method

Control of an elastic membrane

For $f \in L^2(\Omega)$, the vertical displacement u of the membrane is solution of

$$\begin{cases} -\Delta u = f + v & \text{in } \Omega \\ u = 0 & \text{on } \partial\Omega, \end{cases}$$

where v is a **control force** which is our optimization variable (for example, a piezzo-electric actuator). We define the set of admissible controls

$$K = \left\{ v \in L^2(\omega) \mid v_{min}(x) \leq v(x) \leq v_{max}(x) \text{ in } \omega \text{ and } v = 0 \text{ in } \Omega \setminus \omega \right\}.$$

We want to control the membrane in order to reach a prescribed displacement $u_0 \in L^2(\Omega)$ by minimizing $(c > 0)$

$$\inf_{v \in K} \left\{ J(v) = \frac{1}{2} \int_\Omega \left(|u - u_0|^2 + c|v|^2 \right) dx \right\}.$$

Existence of an optimal control

Proposition.

There exists a unique optimal control $\overline{v} \in K$.

Proof. $v \to u$ is an affine function from K into $H_0^1(\Omega)$.

The integrand of J is a positive "polynomial" of degree two in v.

$v \to J(v)$ is strongly convex on K which is convex.

Remark. The existence is often more delicate to prove, but the important thing here is to compute a gradient $J'(v)$ for numerical purposes.

Important notice: the solution u of the p.d.e. depends on the control v.

Gradient and optimality condition

The safest and simplest way of computing a gradient is to evaluate the directional derivative

$$j(t) = J(v + tw) \quad \Rightarrow \quad j'(0) = \langle J'(v), w \rangle = \int_\Omega J'(v) w \, dx \, .$$

By linearity, we have $u(v + tw) = u(v) + t\tilde{u}(w)$ with

$$\begin{cases} -\Delta \tilde{u}(w) = w & \text{in } \Omega \\ \tilde{u}(w) = 0 & \text{on } \partial\Omega. \end{cases}$$

In other words, $\tilde{u}(w) = \langle u'(v), w \rangle$.

Since $J(v)$ is quadratic the computation is very simple and we obtain

$$\int_\Omega J'(v) w \, dx = \int_\Omega \Big((u(v) - u_0)\tilde{u}(w) + cvw \Big) dx,$$

Unfortunately $J'(v)$ is not explicit because we cannot factorize out w in $\tilde{u}(w)$!

Adjoint state

To simplify the gradient formula we use the so-called adjoint state p, defined as the unique solution in $H^1_0(\Omega)$ of

$$\left\{\begin{array}{ll} -\Delta p = u - u_0 & \text{in } \Omega \\ p = 0 & \text{on } \partial\Omega. \end{array}\right.$$

We multiply the equation for $\tilde{u}(w)$ by p and conversely

$$\text{equation for } p \times \tilde{u}(w) \quad \Rightarrow \quad \int_\Omega \nabla p \cdot \nabla \tilde{u}(w)\, dx = \int_\Omega (u - u_0)\tilde{u}(w)\, dx$$

$$\text{equation for } \tilde{u}(w) \times p \quad \Rightarrow \quad \int_\Omega \nabla \tilde{u}(w) \cdot \nabla p\, dx = \int_\Omega wp\, dx$$

Comparing these two equalities we deduce that

$$\int_\Omega (u - u_0)\tilde{u}(w)\, dx = \int_\Omega wp\, dx \quad \Rightarrow \quad \int_\Omega J'(v)w\, dx = \int_\Omega (p + cv)w\, dx.$$

Conclusion on the adjoint state

We found an explicit formula of the gradient

$$J'(v) = p + cv.$$

- ☞ Adjoint method: computation of the gradient by solving 2 boundary value problems (u and p).
- ☞ If one does not use the adjoint: for **each** direction w one must solve 2 boundary value problems (u and $\tilde{u}(w)$) to evaluate $\langle J'(v), w\rangle$.

  For example, if $J'(v)$ is a vector of dimension n, its n components are obtained by solving $(n+1)$ problems !
- ☞ Very efficient in practice: it is the best possible method.
- ☞ Inconvenient: if one uses a black-box software to compute u, it can be very difficult to modify it in order to get the adjoint state p.

Further remarks on the notion of adjoint state

- ☞ If the state equation is not self-adjoint (the bilinear form is not symmetric), the operator of the adjoint equation is the transposed or adjoint of the direct operator.
- ☞ If the state equation is time dependent with an initial condition, then the adjoint equation is time dependent too, but backward with a final condition.
- ☞ If the state equation is non-linear, the adjoint equation is linear.

The adjoint is not just a trick ! It can be deduced from the Lagrangian of the problem.

General method to find the adjoint equation

We consider the state equation as a constraint and, for any $(\hat{v}, \hat{u}, \hat{p}) \in L^2(\Omega) \times H_0^1(\Omega) \times H_0^1(\Omega)$, we introduce the Lagrangian of the minimization problem

$$\mathcal{L}(\hat{v}, \hat{u}, \hat{p}) = \frac{1}{2} \int_\Omega \left(|\hat{u} - u_0|^2 + c|\hat{v}|^2 \right) dx + \int_\Omega \hat{p}(\Delta \hat{u} + f + \hat{v})\, dx,$$

where $\hat{p}$ is the Lagrange multiplier for the constraint which links the two independent variables $\hat{v}$ and $\hat{u}$.

Integrating by parts yields

$$\mathcal{L}(\hat{v}, \hat{u}, \hat{p}) = \frac{1}{2} \int_\Omega \left(|\hat{u} - u_0|^2 + c|\hat{v}|^2 \right) dx + \int_\Omega (-\nabla \hat{p} \cdot \nabla \hat{u} + f\hat{p} + \hat{v}\hat{p})\, dx.$$

Proposition. The optimality conditions are equivalent to the stationnarity of the Lagrangian, i.e.,

$$\frac{\partial \mathcal{L}}{\partial v} = \frac{\partial \mathcal{L}}{\partial u} = \frac{\partial \mathcal{L}}{\partial p} = 0.$$

Proof

- $\frac{\partial \mathcal{L}}{\partial p} = 0 \Rightarrow$ by definition, we recover the equation satisfied by the state u.
- $\frac{\partial \mathcal{L}}{\partial u} = 0 \Rightarrow$ equation satisfied by the adjoint state p. Indeed,

$$\ell_u(t) = \mathcal{L}(\hat{v}, \hat{u} + t\phi, \hat{p}) \quad \Rightarrow \quad \ell'_u(0) = \langle \frac{\partial \mathcal{L}}{\partial u}, \phi \rangle = \int_\Omega ((\hat{u} - u_0)\phi - \nabla \hat{p} \cdot \nabla \phi)\, dx$$

which is the variational formulation of the adjoint equation.

- $\frac{\partial \mathcal{L}}{\partial v} = 0 \Rightarrow$ formula for $J'(v)$. Indeed,

$$\ell_v(t) = \mathcal{L}(\hat{v} + tw, \hat{u}, \hat{p}) \quad \Rightarrow \quad \ell'_v(0) = \langle \frac{\partial \mathcal{L}}{\partial v}, w \rangle = \int_\Omega (c\hat{v} + \hat{p}) w\, dx$$

Simple formula for the derivative

In the preceding proof we obtained

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

with the state u and the adjoint p (both depending on v).

It is not a surprise ! Indeed,

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

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

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

We then take $\hat{p} = p$, the adjoint, to obtain

$$\langle J'(v), w\rangle = \langle \frac{\partial \mathcal{L}}{\partial v}(v, u, p), w\rangle$$

Another interpretation of the adjoint state

The adjoint state p is the Lagrange multiplier for the constraint of the state equation. But it is also a sensitivity function.

Define the Lagrangian

$$\mathcal{L}(\hat{v}, \hat{u}, \hat{p}, f) = \frac{1}{2}\int_\Omega \left(|\hat{u} - u_0|^2 + c|\hat{v}|^2\right) dx + \int_\Omega (-\nabla\hat{p}\cdot\nabla\hat{u} + f\hat{p} + \hat{v}\hat{p})\, dx.$$

We study the sensitivity of the minimum with respect to variations of f.

We denote by $v(f)$, $u(f)$ and $p(f)$ the optimal values, depending on f. We assume that they are differentiable with respect to f. Then

$$\nabla_f \Big(J(v(f)) \Big) = p(f).$$

p gives the derivative (without further computation) of the minimun with respect to f !

Indeed $J(v(f)) = \mathcal{L}(v(f), u(f), p(f), f)$ and $\frac{\partial \mathcal{L}}{\partial v} = \frac{\partial \mathcal{L}}{\partial u} = \frac{\partial \mathcal{L}}{\partial p} = 0.$

CHAPTER V

PARAMETRIC (OR SIZING) OPTIMIZATION

Optimization of a membrane thickness

Membrane occupying a bounded domain Ω in $\mathbb{R}^N$. Forces $f \in L^2(\Omega)$, displacement $u \in H_0^1(\Omega)$ which is solution of

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

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

Admissible set of thicknesses h, defined by

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

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 criteria to reach a target displacement $u_0 \in L^2(\Omega)$

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

Continuity of the cost function

Proposition 5.1. The application

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

is continuous from $\mathcal{U}_{ad}$ into $\mathbb{R}$.

Proof. By composition of the 2 continuous functions below.

Lemma 5.2. 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 5.3. The map $h \to u$ is continuous from $\mathcal{U}_{ad}$ into $H_0^1(\Omega)$.

Proof of Lemma 5.3.

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

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

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

5.2 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.

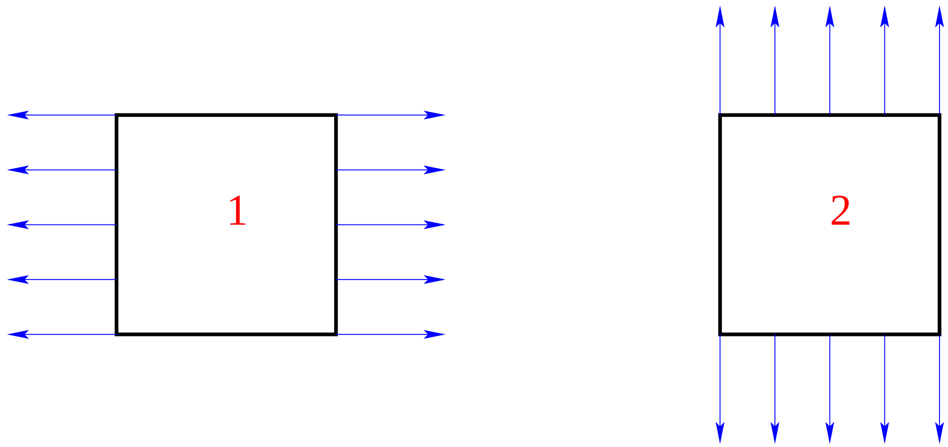
  1. Discretization: finite dimensional admissible set.
  2. Regularization: compact admissible set.
  3. A “miracle”: compliance minimization is a convex problem.

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 a membrane which is

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

Definition of the counter-example

$$\begin{cases} -\operatorname{div}(h\nabla u_1) = 0 & \text{in } \Omega, \\ h\nabla u_1 \cdot n = e_1 \cdot n & \text{on } \partial\Omega, \end{cases} \qquad \begin{cases} -\operatorname{div}(h\nabla u_2) = 0 & \text{in } \Omega, \\ h\nabla u_2 \cdot n = e_2 \cdot n & \text{on } \partial\Omega, \end{cases}$$

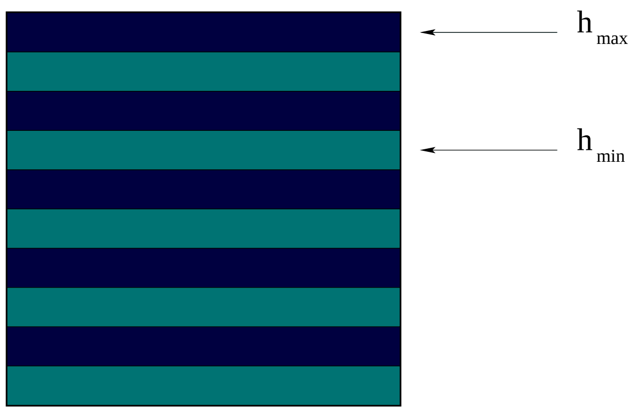
$$\inf_{h\in\mathcal{U}_{ad}} J(h) = \int_{\partial\Omega} e_1 \cdot n\, u_1\, ds - \int_{\partial\Omega} e_2 \cdot n\, u_2\, ds$$

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

The same membrane is subjected to the 2 loadings.

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)

- ✘ Verticaly, the lines of forces must cross the layers of minimal thickness: the structure is thus weak.
- ✘ Horizontaly, the lines of forces follow the layers of maximal thickness: the structure is thus strong.
- ✘ **However**, since the boundary conditions are uniform, the membrane is horizontally stronger if the layers are **finer** because the lines of forces are deviating from the horizontal to a lesser extent.

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 !

5.2.2 Existence for a discretized model

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

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

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

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

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 5.9 (finite dimension). The optimization problem

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

admits at least one minimizer.

Proof. Since $\mathcal{U}_{ad}^n$ is a compact subspace of $\mathbb{R}^n$ and $J(h)$ is a continuous function on $\mathcal{U}_{ad}^n$ (see Proposition 5.1), we can apply Theorem 3.3 which gives the existence of a minimizer of J in $\mathcal{U}_{ad}^n$.

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

5.2.3 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: "feasability" constraint because, in practice, the thickness cannot rapidly vary.

Theorem 5.12. 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 prefered

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

5.3.1 Computation of a continuous gradient

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

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

Lemma 5.15. 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

$$\langle u'(h), k \rangle = v,$$

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. Formaly, one simply differentiates the equation with respect to h. However, to be mathematically rigorous one should rather work at the level of the variational formulation (see the textbook).

To compute the directional derivative, we define $h(t) = h + tk$ for $t > 0$. Let $u(t)$ be the solution for the thickness $h(t)$. Deriving with respect to t leads to

$$\begin{cases} -\operatorname{div}(h(t)\nabla u'(t)) = \operatorname{div}(h'(t)\nabla u(t)) & \text{in } \Omega \\ u'(t) = 0 & \text{on } \partial\Omega, \end{cases}$$

and, since $h'(0) = k$, we deduce $u'(0) = v$.

Lemma 5.17. 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 $|j(u)| \le C(u^2+1)$ and $|j'(u)| \le C(|u|+1)$ for any $u \in \mathbb{R}$. The application $J(h)$, from $\mathcal{U}$ into $\mathbb{R}$, is differentiable and its directional 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 $v = \langle u'(h), k\rangle$ 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. By simple composition of differentiable applications.

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

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

If $h \in \mathcal{U}_{ad}$ is a local minimizer of J in $\mathcal{U}_{ad}$, 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 Lemma 5.17, we must eliminate $v = \langle u'(h), k\rangle$. We use the adjoint state for that: multiplying the equation for v by p and that for p by v, we integrate by parts

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

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

Comparing these two equalities we deduce

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

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 $(\hat{h}, \hat{u}, \hat{p}) \in L^\infty(\Omega) \times H_0^1(\Omega) \times H_0^1(\Omega)$, we introduce the Lagrangian

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

where $\hat{p}$ is a Lagrange multiplier (a function) for the constraint which connects u to h. By integration by parts we get

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

The partial derivative of $\mathcal{L}$ with respect to u in the direction $\phi \in H_0^1(\Omega)$ 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

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

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

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

5.4 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

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

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.

5.4.1 An existence result

We rewrite the problem under the form

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

with $H=\{\tau\in L^2(\Omega)^N\ ,\ -\operatorname{div}\tau=f \text{ in } \Omega\}$.

Lemma 5.8. The function $\phi(a,\sigma)=a^{-1}|\sigma|^2$, defined from $\mathbb{R}^+\times\mathbb{R}^N$ into $\mathbb{R}$, is convex and satisfies

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

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

5.4.2 Optimality conditions

Lemma 5.25. Take $\tau \in L^2(\Omega)^N$. 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 $\displaystyle\int_\Omega h(x)\, dx = h_0 |\Omega|$.

Proof. The function $h \to \int_\Omega h^{-1} |\tau|^2 dx$ is convex from $\mathcal{U}_{ad}$ into $\mathbb{R}$ and we easily compute its derivative.

5.5 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 **rigidity 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

$$\left(J^{disc}\right)'(h) = y'(h)\left(j^{disc}\right)'(y(h)) \quad \text{with} \quad y'(h) = -K(h)^{-1}K(h)'K(h)^{-1}b.$$

Notations: $f'(h) = (\partial f(h)/\partial h_i)_{1\le i\le n}$.

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, for each component i,

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

from which we deduce

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

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 simplification in using a discrete approach rather than a continuous one.

Some authors prefer to discretize first, optimize afterwards. It guarantees a perfect compatibility between the gradient and the cost function, but it requires a deep knowledge of the numerical solver (almost impossible if one has not written himself the source code !).

Here, we follow another philosophy: first optimize in a continuous framework, then discretize. It is much simpler ! No precision is lost if the finite element spaces are adequately chosen.