# An introduction to optimal design

Charles Dapogny

Laboratoire Jacques-Louis Lions, Sorbonne Université, Paris, France

14th April, 2026

# Foreword (I)

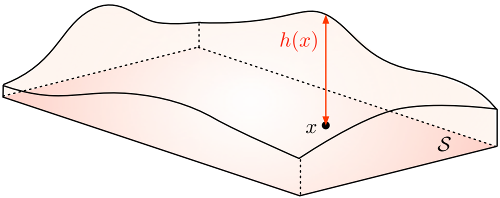
This course focuses on parametric optimal design problems, of the form:

$$\min_{h \in H} J(h) \text{ s.t. } G(h) = 0, \quad (\mathcal{P})$$

where

- The design variable $h$ belongs to a fixed vector space $H$;
- The objective and constraint functions $J(h)$ and $G(h)$ depend on $h$ via a state $u_h \in V$, solution to an $h$-dependent boundary value problem:

$$\text{Search for } u_h \in V \text{ s.t. } \forall v \in V, \quad a_h(u_h, v) = \ell_h(v).$$

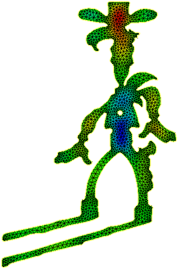



*An elastic plate can be described by its height $h : \mathcal{S} \to \mathbb{R}$ over a fixed cross-section $\mathcal{S}$.*

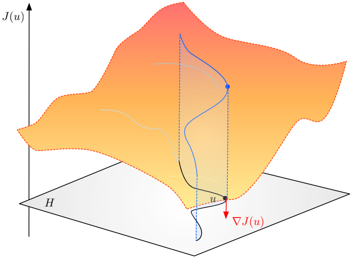
# Foreword (II)

In `Course 1`, we have seen that:

- For a given design $h \in H$, the numerical simulation of $u_h$ can be realized by the Finite Element Method, e.g. using `FreeFem`.
- The solution of $(\mathcal{P})$ can be achieved by a (constrained) gradient-based strategy.



*Solution of a PDE by the Finite Element Method.*

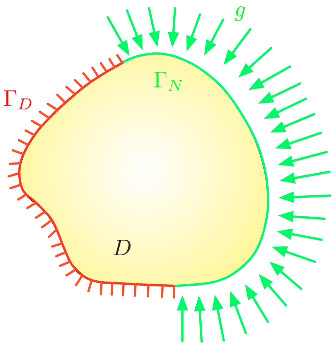



*The gradient algorithm in action*

$\Rightarrow$ This course is devoted to the calculation of the derivative of a function $J(h)$ depending on $h$ via the solution to a PDE depending on $h$.

# Part II

# Optimal control and parametric optimization problems




- We aim to optimize the thermal conductivity $h : D \to \mathbb{R}$ inside a given domain $D \subset \mathbb{R}^d$.
- The temperature $u_h$ is the solution in $H^1(D)$ to the "state", conductivity equation:

$$\begin{cases} -\text{div}(h\nabla u_h) = f & \text{in } D, \\ u_h = 0 & \text{on } \Gamma_D, \\ h\frac{\partial u_h}{\partial n} = g & \text{on } \Gamma_N, \end{cases}$$

where $f \in L^2(D)$ and $g \in L^2(\Gamma_N)$ are sources.

- The set $\mathcal{U}_{\text{ad}}$ of design variables is:

$$\mathcal{U}_{\text{ad}} = \left\{ h \in L^\infty(D), \ \alpha \leq h(x) \leq \beta \text{ a.e. } x \in D \right\} \subset L^\infty(D),$$

where $0 < \alpha < \beta$ are fixed bounds.

# A model problem involving the conductivity equation (II)

- To set ideas, we consider an unconstrained problem of the form:

$$\min_{h \in \mathcal{U}_{\mathrm{ad}}} J(h), \text{ where } J(h) = \int_D j(u_h)\,\mathrm{d}x,$$

  and $j : \mathbb{R} \to \mathbb{R}$ is a smooth function satisfying growth conditions:

$$\forall s \in \mathbb{R}, \ \ |j(s)| \leq C(1 + |s|^2), \text{ and } j'(s) \leq C(1 + |s|).$$

- Many variants are possible, e.g. featuring constraints on $h$ or $u_h$.
- In this setting,
  - For any $h \in \mathcal{U}_{\mathrm{ad}}$, the state $u_h$ is evaluated on the fixed domain $D$;
  - The design variable $h$ acts as a parameter in the coefficients of the state equation.
- Before dealing with the (local) resolution of this problem, let us inquire about the existence of global solutions.

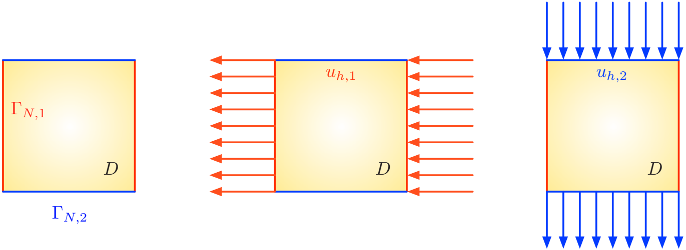
# Part II

# Optimal control and parametric optimization problems

# Non existence of optimal design (I)

- This counter-example is discussed in details in [All] §5.2.
- The defining domain is the unit square $D = (0,1)^2$.
- We consider a variant of the model problem, made of two physical scenarios:

$$\begin{cases} -\text{div}(h\nabla u_{h,1}) = 0 & \text{in D}, \\ h\frac{\partial u_{h,1}}{\partial n} = e_1 \cdot n & \text{in } \Gamma_{N,1}, \\ h\frac{\partial u_{h,1}}{\partial n} = 0 & \text{in } \Gamma_{N,2}, \end{cases} \quad \text{and} \quad \begin{cases} -\text{div}(h\nabla u_{h,2}) = 0 & \text{in D}, \\ h\frac{\partial u_{h,2}}{\partial n} = 0 & \text{in } \Gamma_{N,1}, \\ h\frac{\partial u_{h,2}}{\partial n} = e_2 \cdot n & \text{in } \Gamma_{N,2}. \end{cases}$$




*(Left) Boundary conditions, (middle) boundary data for $u_{h,1}$; (right) boundary data for $u_{h,2}$.*

# Non existence of optimal design (II)

The optimization problem of interest in this context is:

$$\min_{h \in \mathcal{U}_{\mathrm{ad}}} J(h),$$

where the objective function is:

$$J(h) = \underbrace{\int_{\Gamma_{N,1}} e_1 \cdot n \, u_{h,1} \, \mathrm{d}s}_{=\int_{\mathbf{right}} u_{h,1} \, \mathrm{d}s - \int_{\mathbf{left}} u_{h,1} \, \mathrm{d}s > 0} - \underbrace{\int_{\Gamma_{N,2}} e_2 \cdot n \, u_{h,2} \, \mathrm{d}s}_{=\int_{\mathbf{top}} u_{h,1} \, \mathrm{d}s - \int_{\mathbf{bottom}} u_{h,1} \, \mathrm{d}s > 0},$$

and the set $\mathcal{U}_{\mathrm{ad}}$ of admissible designs encompasses a volume constraint:

$$\mathcal{U}_{\mathrm{ad}} = \left\{ h \in L^\infty(D), \begin{array}{c} \alpha < h(x) < \beta \text{ a.e. } x \in D, \\ \int_D h \, \mathrm{d}x = V_T \end{array} \right\}.$$

In other terms, one aims to

- Minimize the temperature difference between the left and right sides in Case 1.
- Maximize the temperature difference between the top and bottom sides in Case 2.

# Non existence of optimal design (III)

## Theorem (Non existence of global solution).

*The parametric optimization problem* $\min\limits_{h \in \mathcal{U}_{\rm ad}} J(h)$ *does not have a global solution.*

*Hint of the proof:* The proof comprises three stages:

**Step 1:** One calculates a lower bound $m$ on the values of $J(h)$ for $h \in \mathcal{U}_{\rm ad}$:

$$\forall h \in \mathcal{U}_{\rm ad}, \;\; J(h) \geq m.$$

**Step 2:** One proves that the value $m$ cannot be attained by an element in $\mathcal{U}_{\rm ad}$:

$$\forall h \in \mathcal{U}_{\rm ad}, \;\; J(h) > m.$$

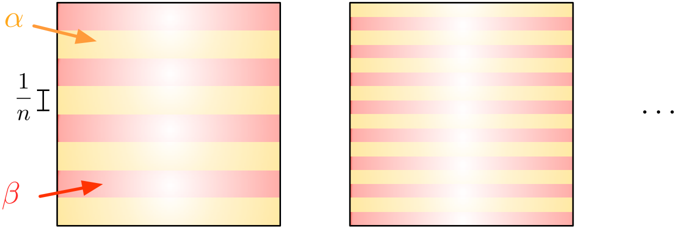
**Step 3:** One constructs a minimizing sequence of designs $h^n \in \mathcal{U}_{\rm ad}$:

$$J(h^n) \xrightarrow{n \to \infty} m.$$

Hence, $m$ is the infimum of $J(h)$ over $\mathcal{U}_{\rm ad}$ but it is not attained by any $h \in \mathcal{U}_{\rm ad}$.

# Non existence of optimal design (IV)

- The minimizing sequence $h^n$ is constructed as a laminate, i.e. a succession of layers with maximum and minimum conductivities.




*Two elements in a minimizing sequence $h^n$ of conductivities: the horizontal stripes allow to reduce the temperature difference between the left and right sides, while insulating the top side from the bottom one.*

- This is an expression of the homogenization effect:

*To get more optimized, designs tend to create very thin structures, at the microscopic level.*

# Non existence of optimal design (V)

- In general, optimal design problems do not have global solutions, for a deep physical reason:

  *It is generally possible to achieve better performance by using smaller patterns.*

- See [Mu] for many such examples of non existence of optimal design.

- To ensure existence of an optimal design, two techniques are usually employed:
  - Relaxation: the set $\mathcal{U}_{\text{ad}}$ of admissible designs is enlarged to enclose "microscopic designs". This is the essence of the Homogenization method [All2].
  - Restriction: the set $\mathcal{U}_{\text{ad}}$ is restricted to, e.g. more regular designs, with fixed scale patterns.

- In practice, one is often interested in the search of local minimizers, which are e.g. "close" to an initial design inspired by intuition.

# Part II

# Optimal control and parametric optimization problems

# Derivative of the objective function (I)

Let us return to our (further simplified) toy problem:

$$\min_{h \in \mathcal{U}_{\text{ad}}} J(h),$$

where

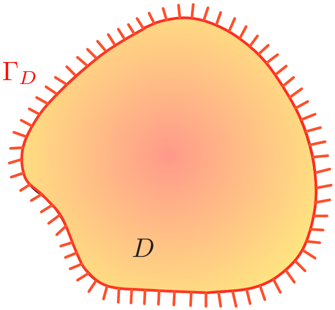
$$J(h) = \int_D j(u_h)\,\mathrm{d}x,$$

the set of admissible designs is:

$$\mathcal{U}_{\text{ad}} = \left\{ h \in L^\infty(D),\ \alpha \leq h(x) \leq \beta \text{ a.e. } x \in D \right\},$$

and the temperature $u_h$ is the solution in $H^1_0(D)$ to:

$$\left\{ \begin{array}{rcll} -\text{div}(h \nabla u_h) & = & f & \text{in } D, \\ u_h & = & 0 & \text{on } \partial D. \end{array} \right.$$




We aim to prove that $J(h)$ is differentiable, and to calculate its derivative.

# Derivative of the objective function (II)

For a given design $h \in \mathcal{U}_{\rm ad}$,

- One variational formulation characterizing $u_h$ is:

$$\text{Search for } u_h \in H^1_0(D) \text{ s.t. } \forall v \in H^1_0(D), \quad \int_D h \nabla u_h \cdot \nabla v \, \mathrm{d}x = \int_D fv \, \mathrm{d}x.$$

- This problem has a unique solution $u_h \in H^1_0(D)$, which satisfies:

$$||u_h||_{H^1_0(D)} \leq C ||f||_{L^2(D)},$$

for some constant $C > 0$, owing to the Lax-Milgram theorem.

# Derivative of the objective function (III)

**Theorem.**

*The objective function*

$$J(h) = \int_D j(u_h)\,\mathrm{d}x$$

*is Fréchet differentiable at any* $h \in \mathcal{U}_{\mathrm{ad}}$*, and its derivative reads*

$$\forall \widehat{h} \in L^\infty(D),\;\; J'(h)(\widehat{h}) = \int_D (\nabla u_h \cdot \nabla p_h)\widehat{h}\,\mathrm{d}x,$$

*where the adjoint state* $p_h \in H^1_0(D)$ *is the unique solution to the problem:*

$$\begin{cases} -\mathrm{div}(h\nabla p_h) = -j'(u_h) & \text{in } D, \\ p_h = 0 & \text{on } \partial D. \end{cases}$$

# The adjoint method (I)

*Proof:* The proof is divided into three steps:

1. Using the implicit function theorem, we prove that the design-to-state mapping

$$\mathcal{U}_{\mathrm{ad}} \ni h \longmapsto u_h \in H^1_0(D)$$

is Fréchet differentiable, with derivative $\widehat{h} \mapsto u_h'(\widehat{h})$.

2. We characterize $u_h'(\widehat{h})$ as the solution to an $h$, $\widehat{h}$ dependent variational problem.

3. We calculate the derivative of $J(h)$ in terms of $u_h'(\widehat{h})$, thanks to the chain rule.

4. We give a more convenient structure to $J'(h)(\widehat{h})$, introducing an adjoint state $p_h$ to eliminate the occurrence of $u_h'(\widehat{h})$.

# Reminder: the implicit function theorem

The implicit function theorem is a key result, ensuring the existence and smoothness of a solution $u = u_\theta$ to a parametrized, non linear equation of the form:

$$\mathcal{F}(\theta, u) = 0,$$

where $u$ is the unknown and $\theta$ is a "parameter"; see [La], Chap. $I$, Th. 5.9.

## Theorem (Implicit function theorem).

*Let $\Theta, E, F$ be Banach spaces, $\mathcal{V} \subset \Theta$, $U \subset E$ be open sets. and $\mathcal{F} : \mathcal{V} \times U \to F$ be a function of class $\mathcal{C}^p$ for $p \geq 1$. Let $(\theta_0, u_0) \in \mathcal{V} \times U$ be such that $\mathcal{F}(\theta_0, u_0) = 0$ and assume that:*

*The derivative $\dfrac{\partial \mathcal{F}}{\partial u}(\theta_0, u_0) : E \to F$ is a linear isomorphism.*

*Then there exist open subsets $\mathcal{V}' \subset \mathcal{V}$ of $\theta_0$ in $\Theta$ and $U' \subset U$ of $u_0$ in $E$, and a mapping $g : \mathcal{V}' \to U'$ of class $\mathcal{C}^p$ satisfying the properties:*

1. $g(\theta_0) = u_0$,
2. *For all $\theta \in \mathcal{V}'$, the equation $\mathcal{F}(\theta, u) = 0$ has a unique solution $u \in U'$, given by $u = g(\theta)$.*

**Step 1:** *Differentiability of $h \mapsto u_h$:*

- For any $h \in \mathcal{U}_{\text{ad}}$, $u_h$ is the unique solution in $H^1_0(D)$ to the variational problem:

$$\forall v \in H^1_0(D), \ \int_D h \nabla u_h \cdot \nabla v \, \mathrm{d}x = \int_D f v \, \mathrm{d}x.$$

- Let

$$\mathcal{F} : \mathcal{U}_{\text{ad}} \times H^1_0(D) \to H^{-1}(D)$$

be the mapping defined by:

For all $h \in \mathcal{U}_{\text{ad}}, u \in H^1_0(D)$, the linear form $\mathcal{F}(u,h) \in H^{-1}(D)$ on $H^1_0(D)$ is:

$$\mathcal{F}(h,u) : v \mapsto \int_D h \nabla u \cdot \nabla v \, \mathrm{d}x - \int_D f v \, \mathrm{d}x.$$

# The adjoint method (III)

One verifies that

- $\mathcal{F}$ is a mapping of class $\mathcal{C}^1$;
- For given $h \in \mathcal{U}_{\rm ad}$, $u_h$ is the unique solution $u$ to the equation

$$\mathcal{F}(h, u) = 0.$$

- The differential of the partial mapping $u \mapsto \mathcal{F}(h, u)$ reads:

$$H^1_0(D) \ni \widehat{u} \longmapsto \left[ v \mapsto \int_D h \nabla \widehat{u} \cdot \nabla v \, \mathrm{d}x \right] \in H^{-1}(D).$$

  It is an isomorphism, owing to the Lax-Milgram theorem:

  For all $g \in H^{-1}(D)$, there exists a unique $u \in H^1_0(D)$ s.t.

$$\forall v \in H^1_0(D), \quad \int_D h \nabla u \cdot \nabla v \, \mathrm{d}x = \langle g, v \rangle_{H^{-1}(D), H^1_0(D)}.$$

$\Rightarrow$ The implicit function theorem guarantees that the mapping $h \mapsto u_h$ is of class $\mathcal{C}^1$.

# The adjoint method (IV)

**Step 2:** *Characterization of the derivative $u_h'(\widehat{h})$ by a variational problem:*

- Let us return to the variational formulation for $u_h$:

$$\forall v \in H^1_0(D), \ \int_D h \nabla u_h \cdot \nabla v \, \mathrm{d}x = \int_D fv \, \mathrm{d}x.$$

- Thanks to **Step 1**, we may take derivatives with respect to $h$ in a direction $\widehat{h} \in L^\infty(D)$, for an arbitrary test function $v \in h^1_0(D)$:

$$\int_D \widehat{h} \nabla u_h \cdot \nabla v \, \mathrm{d}x + \int_D h \nabla u_h'(\widehat{h}) \cdot \nabla v \, \mathrm{d}x = 0.$$

- Hence, for all $\widehat{h} \in L^\infty(D)$, $u_h'(\widehat{h})$ is the unique solution in $H^1_0(D)$ to:

Search for $u_h'(\widehat{h}) \in H^1_0(D)$ s.t. $\forall v \in H^1_0(D)$,

$$\int_D h \nabla u_h'(\widehat{h}) \cdot \nabla v \, \mathrm{d}x = - \int_D \widehat{h} \nabla u_h \cdot \nabla v \, \mathrm{d}x.$$

# Reminder: the "chain rule"

The chain rule is a *fundamental result*, which supplies the Fréchet derivative of the composite $G \circ F$ of two functions

$$F : U \to V \text{ and } G : V \to W$$

between three normed vector spaces $(U, ||\cdot||_U)$, $(V, ||\cdot||_V)$ and $(W, ||\cdot||_W)$.

**Theorem (Chain rule).**

*Let $x \in U$ be a point such that:*

- *$F$ is Fréchet differentiable at $x$;*
- *$G$ is Fréchet differentiable at $F(x) \in V$.*

*Then, the composite function $G \circ F : U \to W$ is Fréchet differentiable at $x$, and its Fréchet derivative $v \mapsto (G \circ F)'(x)(v)$ is the linear mapping defined by:*

$$\forall v \in U, \ (G \circ F)'(x)(v) = G'(F(x))(F'(x)(v)).$$

# The adjoint method (V)

**Step 3:** *Calculation of the derivative of $J(h)$:*

Since $h \mapsto u_h$ is of class $\mathcal{C}^1$, the chain rule yields immediately:

$$\forall \widehat{h} \in L^\infty(D), \;\; J'(h)(\widehat{h}) = \int_D j'(u_h) u_h'(\widehat{h}) \, \mathrm{d}x.$$

- This expression is awkward: the dependence $\widehat{h} \mapsto J'(h)(\widehat{h})$ is not explicit and it is difficult to find a descent direction, i.e. a vector $\widehat{h} \in L^\infty(D)$ such that:

$$J'(h)(\widehat{h}) < 0.$$

$\Rightarrow$ One would have to try multiple $\widehat{h} \in L^\infty(D)$, calculate $u_h'(\widehat{h})$, ... until finding one s.t. $J'(h)(\widehat{h}) < 0$.

## The adjoint method (VI)

**Step 4:** *Reformulation of $J'(h)(\widehat{h})$ using an adjoint state:*

The adjoint state $p_h$ is the unique solution in $H^1_0(D)$ to the variational problem:

$$\forall v \in H^1_0(D), \ \int_D h \nabla p_h \cdot \nabla v \, \mathrm{d}x = - \int_D j'(u_h) v \, \mathrm{d}x,$$

to be compared with the variational formulation for $u_h'(\widehat{h}) \in H^1_0(D)$:

$$\forall v \in H^1_0(D), \ \int_D h \nabla u_h'(\widehat{h}) \cdot \nabla v \, \mathrm{d}x = - \int_D \widehat{h} \nabla u_h \cdot \nabla v \, \mathrm{d}x.$$

Then, we calculate:

$$\begin{aligned} J'(h)(\widehat{h}) &= \int_D j'(u_h) u_h'(\widehat{h}) \, \mathrm{d}x, \\ &= - \int_D h \nabla p_h \cdot \nabla u_h'(\widehat{h}) \, \mathrm{d}x, \\ &= - \int_D h \nabla u_h'(\widehat{h}) \cdot \nabla p_h \, \mathrm{d}x, \\ &= \int_D \widehat{h} \nabla u_h \cdot \nabla p_h \, \mathrm{d}x. \end{aligned}$$

where the last line uses the variational formulation of $u_h'(\widehat{h})$ with $p_h$ as test function.

# About the adjoint state (I)

- The above introduction of the adjoint state $p_h$, as a solution to the problem:

$$\forall v \in H^1_0(D), \quad \int_D h \nabla p_h \cdot \nabla v \, \mathrm{d}x = - \int_D j'(u_h) v \, \mathrm{d}x,$$

and the allowed simplifications may look "miraculous" at first glance.

- After closer inspection, its definition is exactly what is needed to simplify $J'(\widehat{h})$:
  - The right-hand side is what is needed to write:

$$J'(h)(\widehat{h}) = \int_D j'(u_h) u_h'(\widehat{h}) \, \mathrm{d}x = - \int_D h \nabla u_h'(\widehat{h}) \cdot \nabla p_h \, \mathrm{d}x.$$

  - This last expression, featuring the action of $v \mapsto \int_D h \nabla u_h'(\widehat{h}) \cdot \nabla v \, \mathrm{d}x$ is the only available information about $u_h'(\widehat{h})$:

$$- \int_D h \nabla u_h'(\widehat{h}) \cdot \nabla p_h \, \mathrm{d}x = \int_D \widehat{h} \nabla u_h \cdot \nabla p_h \, \mathrm{d}x.$$

# About the adjoint state (II)

- The adjoint state $p_h$ satisfies

$$\begin{cases} -\mathrm{div}(h\nabla p_h) = -j'(u_h) & \text{in } D, \\ p_h = 0 & \text{on } \partial D. \end{cases}$$

  Loosely speaking, it is a "virtual temperature" driven by a source (or sink) equal to the rate of change of the integrand of $J(h)$ at the state described by $u_h$.

- From the last expression, one obviously obtains a descent direction:

$$\widehat{h} = -\nabla u_h \cdot \nabla p_h \;\Rightarrow\; J'(h)(\widehat{h}) < 0,$$

  which can be interpreted as the power induced by the "virtual temperature" $p_h$.

- We shall see soon a second interpretation of $p_h$ as the Lagrange multiplier associated to the PDE constraint if we formulate our optimization problem as:

$$\min_{(h,u)} \int_D j(u)\,\mathrm{d}x \text{ s.t. } \begin{cases} -\mathrm{div}(h\nabla u) = f & \text{in } D, \\ u = 0 & \text{on } \partial D. \end{cases}$$

# Part II

# Optimal control and parametric optimization problems

# The adjoint method in an abstract framework (I)

- Let us consider an abstract optimal design functional, of the form:

$$J(h) = j(u_h),$$

where

- The design $h$ belongs to a Hilbert space $(H, \langle \cdot, \cdot \rangle_H)$.
- The state $u_h$ belongs to another Hilbert space $(V, \langle \cdot, \cdot \rangle_V)$.
- It is the solution to a boundary value problem

$$\text{Search for } u_h \in V \text{ s.t. } \mathcal{F}(h, u_h) = 0, \qquad (BVP)$$

where $\mathcal{F} : H \times V \to V$ is a suitable operator.

- The observable $j : V \to \mathbb{R}$ is smooth enough.

- We aim to calculate the derivative $J'(h)(\widehat{h})$ and to find a descent direction for $J(h)$.

---

**R.-E. Plessix**, *A review of the adjoint-state method for computing the gradient of a functional with geophysical applications*, Geophysical Journal International, 167 (2006), pp. 495–503.

# The adjoint method in an abstract framework (II)

The first 3 steps of the previous derivation unfold exactly as in there:

1. The implicit function theorem ensures that the mapping $h \mapsto u_h$ is differentiable.

2. Differentiation in (*BVP*) yields a characterization of the derivative $u_h'(\widehat{h}) \in V$:

$$\left[\frac{\partial \mathcal{F}}{\partial h}(h, u_h)\right](\widehat{h}) + \left[\frac{\partial \mathcal{F}}{\partial u}(h, u_h)\right] u_h'(\widehat{h}) = 0.$$

3. Likewise, by applying the chain rule, we calculate the derivative of $J(h)$:

$$J'(h)(\widehat{h}) = \left\langle j'(u_h), u_h'(\widehat{h}) \right\rangle_V .$$

Unfortunately, this formula does not easily provide a descent direction.

$\Rightarrow$ One would have to try multiple $\widehat{h} \in H$, calculate $u_h'(\widehat{h})$, ... until finding one such that $J'(h)(\widehat{h}) < 0$.

# The adjoint method in an abstract framework (III)

- To overcome this issue, we introduce the adjoint state $p_h \in V$ as the solution to:

$$\left[\frac{\partial \mathcal{F}}{\partial u}(h, u_h)\right]^* p_h = -j'(u_h),$$

where $\left[\frac{\partial \mathcal{F}}{\partial u}(h, u_h)\right]^* : V \to V$ is the adjoint operator of $\left[\frac{\partial \mathcal{F}}{\partial u}(h, u_h)\right] : V \to V$:

$$\forall v, w \in V, \quad \left\langle \left[\frac{\partial \mathcal{F}}{\partial u}(h, u_h)\right]^* v, w \right\rangle_V = \left\langle v, \left[\frac{\partial \mathcal{F}}{\partial u}(h, u_h)\right] w \right\rangle_V .$$

- An elementary calculation yields:

$$\begin{aligned} J'(h)(\widehat{h}) &= \left\langle j'(u_h), u_h'(\widehat{h}) \right\rangle_V \\ &= -\left\langle \left[\tfrac{\partial \mathcal{F}}{\partial u}(h, u_h)\right]^* p_h, u_h'(\widehat{h}) \right\rangle_V && \text{(Def. of } p_h\text{)} \\ &= -\left\langle \left[\tfrac{\partial \mathcal{F}}{\partial u}(h, u_h)\right] u_h'(\widehat{h}), p_h \right\rangle_V . && \text{(Def. of adjoint operator)} \end{aligned}$$

- Invoking now the problem satisfied by $u_h'(\widehat{h})$, we obtain:

$$J'(h)(\widehat{h}) = \left\langle \left[\frac{\partial \mathcal{F}}{\partial h}(h, u(h))\right] \widehat{h}, p_h \right\rangle_V .$$

## The adjoint method in an abstract framework (IV)

- Let us introduce the adjoint $\left[\frac{\partial \mathcal{F}}{\partial h}(h, u_h)\right]^* : V \to H$ of $\left[\frac{\partial \mathcal{F}}{\partial h}(h, u_h)\right] H \to V$:

$$\forall v \in V,\ h \in H, \quad \left\langle \left[\frac{\partial \mathcal{F}}{\partial h}(h, u_h)\right]^* v, h \right\rangle_H = \left\langle v, \left[\frac{\partial \mathcal{F}}{\partial h}(h, u_h)\right] h \right\rangle_V .$$

  We end up with:

$$J'(h)(\widehat{h}) = \left\langle \left[\frac{\partial \mathcal{F}}{\partial h}(h, u_h)\right]^* p_h, \widehat{h} \right\rangle_H .$$

- Now, a descent direction $\widehat{h}$ for $J(h)$ is immediately revealed:

$$\widehat{h} = -\left[\frac{\partial \mathcal{F}}{\partial h}(h, u_h)\right]^* p_h \ \Rightarrow J'(h)(\widehat{h}) = -\left|\left|\left[\frac{\partial \mathcal{F}}{\partial h}(h, u_h)\right]^* p_h\right|\right|_H^2 .$$

- The evaluation of this descent direction $\widehat{h}$ demands:
  - The calculation of $u_h$ (Finite Element solution);
  - The calculation of $p_h$ (Finite Element solution);
  - The calculation of the operator $\left[\frac{\partial \mathcal{F}}{\partial h}(h, u_h)\right]^*$ (derivative of the explicit dependence of the boundary value problem w.r.t the design).

# The adjoint method: final comments

- The adjoint state $p_h$ can be interpreted as the Lagrange multiplier associated to the PDE constraint if we formulate the minimization problem of $J(h)$ as:

$$\min_{(h,u)} \int_D j(u)\,\mathrm{d}x \text{ s.t. } \left\{ \begin{array}{cl} -\mathrm{div}(h\nabla u) = f & \text{in } D, \\ u = 0 & \text{on } \partial D. \end{array} \right.$$

- This general methodology allows to deal with
  - Different physical situations,
  - Other optimal design frameworks, including "true" geometric optimization.

- Another (formal and sometimes dangerous) method allows to calculate shape derivatives: Céa's method.

# Part II

# Optimal control and parametric optimization problems

# The formal method of Céa (I)

- The method of Céa is a formal way, inspired from constrained optimization theory, to calculate the derivative of $J(h)$.
- It assumes that the mapping $h \mapsto u_h$ is differentiable.
- The original, unconstrained problem

$$\min_{h \in \mathcal{U}_{\text{ad}}} J(h), \text{ where } J(h) = \int_D j(u_h)\, \mathrm{d}x,$$

and $u_h$ is the solution to the conductivity equation:

$$\begin{cases} -\mathrm{div}(h \nabla u_h) = f & \text{in } D, \\ u_h = 0 & \text{on } \partial D, \end{cases}$$

can be rewritten as a constrained optimization problem:

$$\min_{\substack{u \in H^1_0(D), \\ h \in \mathcal{U}_{\text{ad}}}} \int_D j(u)\, \mathrm{d}x, \text{ s.t. } \begin{cases} -\mathrm{div}(h \nabla u) = f & \text{in } D, \\ u = 0 & \text{on } \partial D, \end{cases}$$

where the PDE relating $h$ to $u$ is accounted for by a constraint.

**J. Céa**, *Conception optimale ou identification de formes, calcul rapide de la dérivée directionnelle de la fonction coût*, ESAIM: M2AN, 20(3), (1986), pp. 371–402.

# The formal method of Céa (II)

- This alternative viewpoint suggests the definition of the Lagrangian:

$$\mathcal{L} : \mathcal{U}_{\text{ad}} \times H^1_0(D) \times H^1_0(D) \to \mathbb{R},$$

by the formula:

$$\mathcal{L}(h,u,p) = \underbrace{\int_D j(u)\,\mathrm{d}x}_{\text{Objective function at stake}} + \underbrace{\int_D h\nabla u \cdot \nabla p \,\mathrm{d}x - \int_D fp\,\mathrm{d}x}_{\substack{\textbf{Enforcement of the PDE constraint } -\mathrm{div}(h\nabla u)=f \\ \textbf{with a Lagrange multiplier } p}}.$$

- By construction, it holds, for any $\widehat{p} \in H^1_0(D)$:

$$\boxed{J(h) = \mathcal{L}(h, u_h, \widehat{p}).}$$

# The formal method of Céa (III)

In order to solve the constrained optimal design problem, we search for the saddle points $(u, p)$ of $\mathcal{L}(h, \cdot, \cdot)$ for each, given $h \in \mathcal{U}_{\rm ad}$.

- Imposing the partial derivative of $\mathcal{L}$ with respect to $p$ to vanish amounts to

$$\forall \widehat{p} \in H^1_0(D), \;\; \int_D h \nabla u \cdot \nabla \widehat{p} \, \mathrm{d}x - \int_D f \widehat{p} \, \mathrm{d}x = 0;$$

  this is the variational formulation for $u = u_h$.

- Imposing the partial derivative of $\mathcal{L}$ with respect to $u$ to vanish amounts to

$$\forall \widehat{u} \in H^1_0(D), \;\; \int_D h \nabla p \cdot \nabla \widehat{u} \, \mathrm{d}x = - \int_D j'(u) \widehat{u} \, \mathrm{d}x;$$

  since $u = u_h$, we recognize the variational formulation for $p = p_h$.

# The formal method of Céa (IV)

- Recall that, for arbitrary $\widehat{p} \in H_0^1(D)$,

$$J(h) = \mathcal{L}(h, u_h, \widehat{p}).$$

- Since we have assumed that $h \mapsto u_h$ is differentiable, the chain rule yields:

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

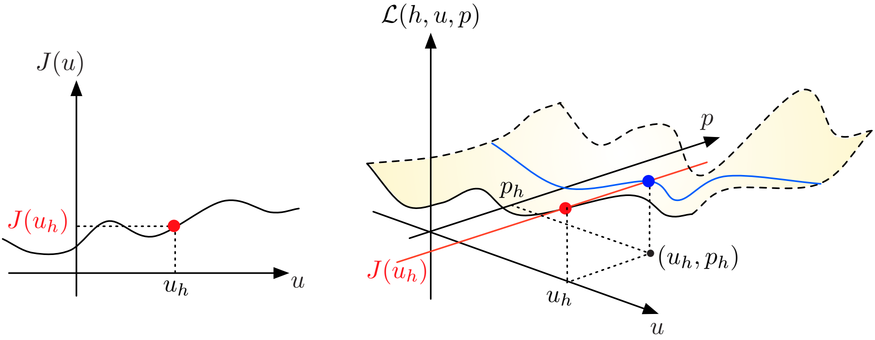
- Now taking $\widehat{p} = p_h$, the last term in the above right-hand side vanishes:

$$J'(h)(\widehat{h}) = \frac{\partial \mathcal{L}}{\partial h}(h, u_h, p_h)(\widehat{h}).$$

- The above partial derivative is the derivative of the mapping $h \mapsto \int_D h \nabla u \cdot \nabla p \, \mathrm{d}x$ evaluated at $u = u_h$ and $p = p_h$:

$$J'(h)(\widehat{h}) = \int_D \widehat{h} \nabla u_h \cdot \nabla p_h \, \mathrm{d}x.$$

# The formal method of Céa: intuition




**intuition:** *The function $J(h)$ is "twisted" into the value $\mathcal{L}(h, u_h, p_h)$ at the parametrized saddle point $(u_h, p_h)$, which is "easy" to differentiate with respect to $h$.*

# Overview of the next lectures

- Course 3 Numerical methods for parametric optimization.
  - Numerical implementation recipes;
  - Density-based topology optimization.

- Practical session A FreeFem implementation of the SIMP method.

# Thank you for your attention!

# Bibliography

# References I


[All2] G. Allaire, *Shape optimization by the homogenization method*, Springer Verlag, (2012).

[All] G. Allaire, *Conception optimale de structures*, Mathématiques & Applications, **58**, Springer Verlag, Heidelberg (2006).

[Ce] J. Céa, *Conception optimale ou identification de formes, calcul rapide de la dérivée directionnelle de la fonction coût*, ESAIM: Modélisation mathématique et analyse numérique, 20(3), (1986), pp. 371–402.

[La] S. Lang, *Fundamentals of differential geometry*, Springer, (1991).

[Li] J.-L. Lions, *Optimal control of systems governed by partial differential equations*, Springer, (1971).

[Mu] F. Murat, *Contre-exemples pour divers problèmes où le contrôle intervient dans les coefficients*, Annali di Matematica Pura ed Applicata, 112, 1, (1977), pp. 49–68.

## References II

[Mu] R. E. Plessix, *A review of the adjoint-state method for computing the gradient of a functional with geophysical applications*, Geophysical Journal International, 167(2), (2006), pp. 495–503.