OPTIMAL DESIGN OF STRUCTURES (MAP 562)

CHAPTER IV

OPTIMAL CONTROL OF DISTRIBUTED SYSTEMS

Computing a gradient by the adjoint method

S. Amstutz, B. Bogosel

Original version by G. Allaire

Academic year 2019-2020

Ecole Polytechnique
Department of applied mathematics

Control of an elastic membrane

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

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

where ξ is a **control force** which is our optimization variable.
Given $\omega \subset \Omega$ we define the set of admissible controls

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

We want to control the membrane in order to reach a prescribed displacement $\bar{u} \in L^2(\Omega)$ by minimizing ($\alpha > 0$)

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

Existence of an optimal control

Proposition.
There exists a unique optimal control $\xi \in K$.
Proof. Write $u = L(f + \xi) = u_0 + L\xi$ the solution of the BVP, $L \in \mathcal{L}(L^2(\Omega), H_0^1(\Omega))$. Then

$$J(\xi) = \frac{1}{2} \int_\Omega \left(|u_0 + L\xi - \bar{u}|^2 + \alpha |\xi|^2 \right) dx.$$

The map $\xi \mapsto J(\xi)$ is a quadratic functional associated with a positive definite quadratic form. Hence it is strictly convex, continuous, and it goes to infinity at infinity. Moreover $L^2(\Omega)$ is reflexive, as a Hilbert space, and K is closed and convex.

Remark. The existence is sometimes more delicate to prove, but the most important thing for us will be to compute a gradient $J'(v)$ for numerical purposes.

Gradient and optimality condition

To calculate the gradient we can proceed by the explicit calculation

$$J(\xi+\tilde{\xi}) = J(\xi)+\int_\Omega \left((L\tilde{\xi})(u-\bar{u})+\alpha\xi\tilde{\xi}\right)dx+\frac{1}{2}\int_\Omega \left(|L\tilde{\xi}|^2+\alpha|\tilde{\xi}|^2\right)dx.$$

We infer

$$\langle J'(\xi),\tilde{\xi}\rangle = \int_\Omega \left((L\tilde{\xi})(u-\bar{u})+\alpha\xi\tilde{\xi}\right)dx.$$

By definition $\tilde{u} := L\tilde{\xi}$ is the solution of

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

Unfortunately $J'(\xi)$ is not explicit because we cannot straightforwardly factor out $\tilde{\xi}$ in the expression.
For that we need the adjoint L^* of L, more specifically $L^*(u-\bar{u})$: the adjoint state...

Adjoint state

Define the adjoint state $p \in H_0^1(\Omega)$ such that

$$\int_\Omega \nabla p \cdot \nabla v \, dx = \int_\Omega (u - \bar{u}) v \, dx \qquad \forall v \in H_0^1(\Omega),$$

i.e.,

$$\boxed{\begin{cases} -\Delta p = u - \bar{u} & \text{in } \Omega \\ p = 0 & \text{on } \partial\Omega. \end{cases}}$$

Now we test the variational formulation for $\tilde{u}$ against p:

$$\int_\Omega \nabla \tilde{u} \cdot \nabla p \, dx = \int_\Omega \tilde{\xi} p \, dx.$$

Comparing the two variational formulations yields

$$\int_\Omega (u - \bar{u}) \tilde{u} \, dx = \int_\Omega \tilde{\xi} p \, dx \quad \Rightarrow \quad \boxed{\langle J'(\xi), \tilde{\xi} \rangle = \int_\Omega (p + \alpha \xi) \tilde{\xi} \, dx.}$$

Remark. We have obtained $\int_\Omega (u - \bar{u})(L\tilde{\xi}) \, dx = \int_\Omega \tilde{\xi} L(u - \bar{u}) \, dx$: in fact L is self-adjoint.

Conclusion on the adjoint state

We found an explicit formula of the gradient

$$J'(v) = p + \alpha\xi.$$

- 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 $\tilde{\xi}$ one must solve 2 boundary value problems (u and $\tilde{u}$) to evaluate $\langle J'(\xi), \tilde{\xi}\rangle$.
  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 bilinear form in the state equation is not symmetric then L is no longer self-adjoint: the operator of the adjoint equation is the 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 $(\hat{\xi}, \hat{u}, \hat{p}) \in L^2(\Omega) \times H_0^1(\Omega) \times H_0^1(\Omega)$, we introduce the Lagrangian of the minimization problem

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

where $\hat{p}$ is the Lagrange multiplier for the constraint which links the two independent variables $\hat{\xi}$ and $\hat{u}$.
Integrating by parts yields

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

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

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

Proof

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

$$\langle \frac{\partial \mathcal{L}}{\partial u}, \phi \rangle = \int_\Omega ((\hat{u} - \bar{u})\phi - \nabla \hat{p} \cdot \nabla \phi) \, dx$$

which is the variational formulation of the adjoint equation.
- $\frac{\partial \mathcal{L}}{\partial \xi} = 0 \Leftrightarrow J'(\xi) = 0$. Indeed,

$$\langle \frac{\partial \mathcal{L}}{\partial \xi}, \tilde{\xi} \rangle = \int_\Omega (\alpha \hat{\xi} + \hat{p}) \tilde{\xi} \, dx = \langle J'(\hat{\xi}), \tilde{\xi} \rangle.$$

Simple formula for the derivative

In the preceding proof we obtained

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

with the state u and the adjoint p (both depending on ξ).
It is not a surprise ! Indeed,

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

because u is the state. Thus, if $u(\xi)$ is differentiable (it stems from $u(\xi) = u_0 + L\xi$), we get

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

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

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

Another interpretation of the adjoint state

The adjoint state p is the Lagrange multiplier for the constraint of the state equation. But, as such, it is also a sensitivity function. Define the parametric Lagrangian (consider f as a variable)

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

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

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

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

p is the derivative (without further computation) of the minimum value with respect to f !

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