

OPTIMAL DESIGN OF STRUCTURES (MAP 562)

G. ALLAIRE

December 17th, 2014

Department of Applied Mathematics, Ecole Polytechnique

CHAPTER III

A REVIEW OF OPTIMIZATION

DEFINITIONS

Let V be a Banach space, i.e., a normed vector space which is complete (any Cauchy sequence is converging in V).

Let $K \subset V$ be a non-empty subset. Let $J : V \rightarrow \mathbb{R}$. We consider

$$\inf_{v \in K \subset V} J(v).$$

Definition. An element u is called a **local minimizer** of J on K if

$$u \in K \quad \text{and} \quad \exists \delta > 0, \forall v \in K, \|v - u\| < \delta \implies J(v) \geq J(u).$$

An element u is called a **global minimizer** of J on K if

$$u \in K \quad \text{and} \quad J(v) \geq J(u) \quad \forall v \in K.$$

(difference: theory \leftrightarrow global / numerics \leftrightarrow local)

Definition. A **minimizing sequence** of a function J on the set K is a sequence $(u^n)_{n \in \mathbb{N}}$ such that

$$u^n \in K \quad \forall n \quad \text{and} \quad \lim_{n \rightarrow +\infty} J(u^n) = \inf_{v \in K} J(v).$$

By definition of the infimum value of J on K **there always exist minimizing sequences !**

Optimization in finite dimension $V = \mathbb{R}^N$

Theorem. Let K be a non-empty closed subset of \mathbb{R}^N and J a continuous function from K to \mathbb{R} satisfying the so-called “infinite at infinity” property, i.e.,

$$\forall (u^n)_{n \geq 0} \text{ sequence in } K, \quad \lim_{n \rightarrow +\infty} \|u^n\| = +\infty \implies \lim_{n \rightarrow +\infty} J(u^n) = +\infty .$$

Then there exists at least one minimizer of J on K . Furthermore, from each minimizing sequence of J on K one can extract a subsequence which converges to a minimum of J on K .

(Main idea: the closed bounded sets are compact in finite dimension.)

Optimization in infinite dimension

Difficulty: a continuous function on a closed bounded set does not necessarily attained its minimum !

Counter-example of non-existence: let $H^1(0, 1)$ be the usual Sobolev space with its norm $\|v\| = \left(\int_0^1 (v'(x)^2 + v(x)^2) dx \right)^{1/2}$. Let

$$J(v) = \int_0^1 \left((|v'(x)| - 1)^2 + v(x)^2 \right) dx .$$

One can check that J is continuous and “infinite at infinity”. Nevertheless the minimization problem

$$\inf_{v \in H^1(0,1)} J(v)$$

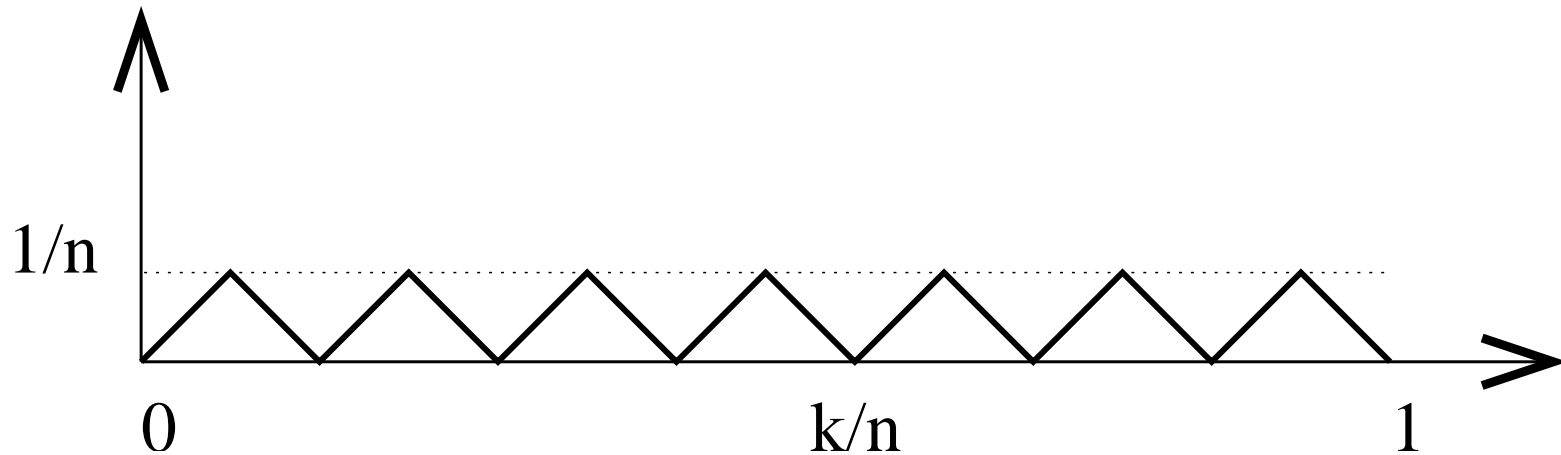
does not admit a minimizer. (Difficulty independent on the choice of the functional space.)

Proof

There exists no $v \in H^1(0, 1)$ such that $J(v) = 0$ but, still,

$$\left(\inf_{v \in H^1(0,1)} J(v) \right) = 0,$$

since, upon defining the sequence u^n such that $(u^n)' = \pm 1$,



we check that $J(u^n) = \int_0^1 u^n(x)^2 dx = \frac{1}{4n} \rightarrow 0$.

We clearly see in this example that the minimizing sequence u^n is “oscillating” more and more and is not compact in $H^1(0, 1)$ (although it is bounded in the same space).

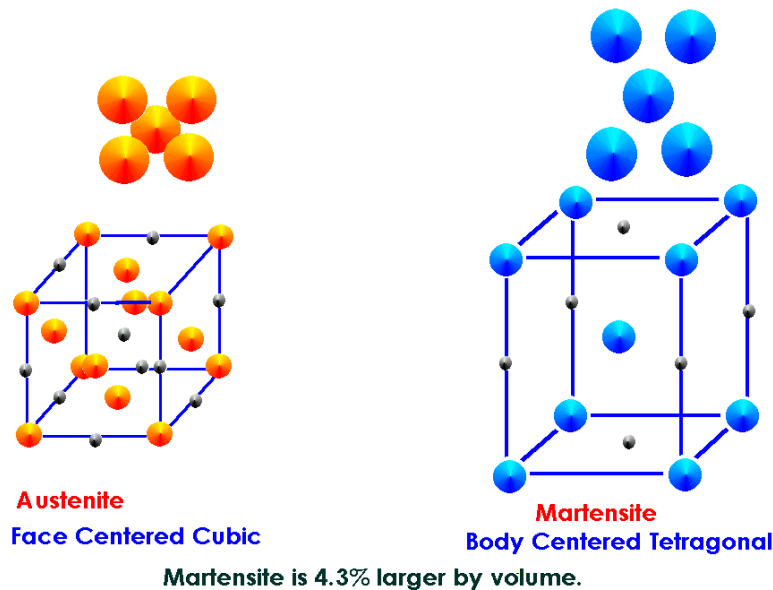
A parenthesis in material sciences

The non-existence of minimizers for minimization problems is useful in material sciences !

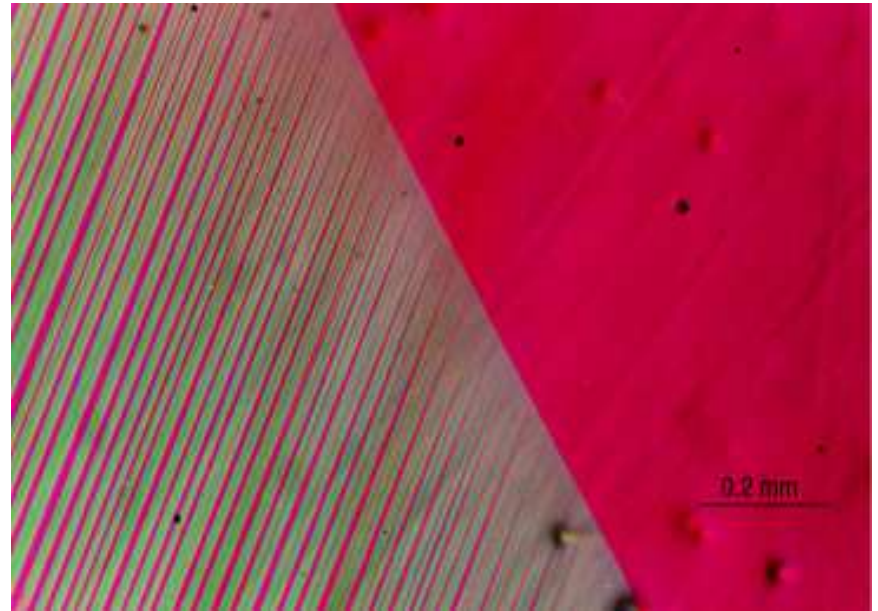
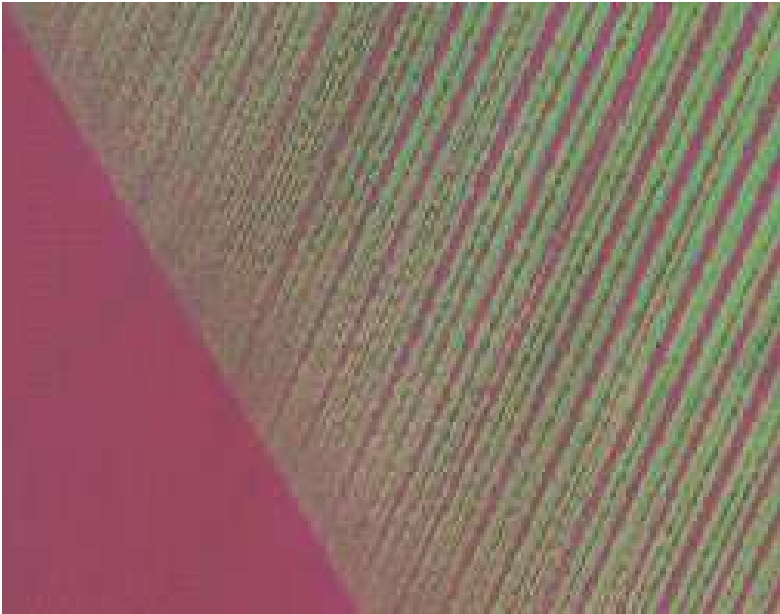
The Ball-James theory (1987).

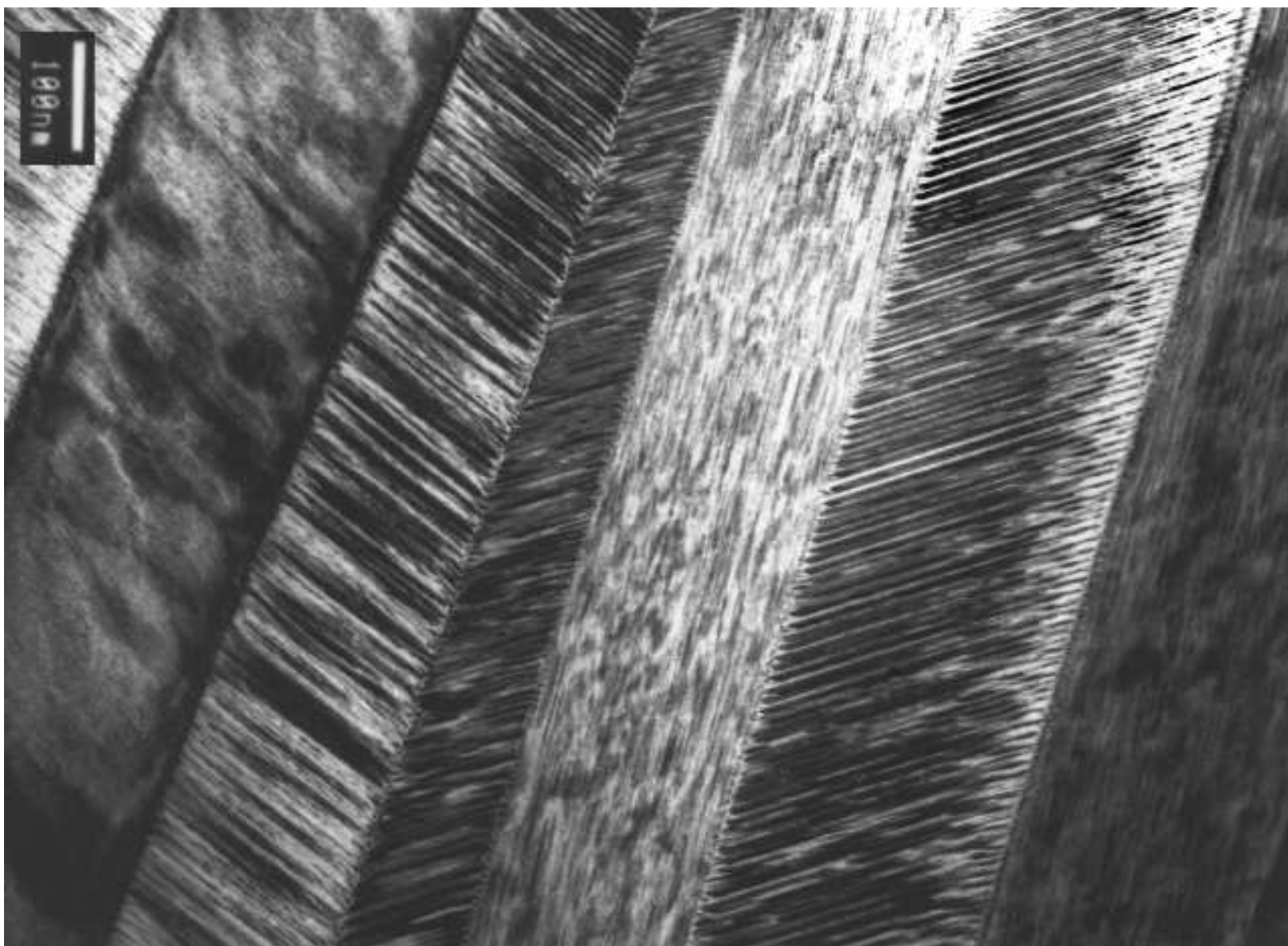
Shape memory materials = alloys with phase transition.

Co-existence of several crystalline phases: austenite and martensite.

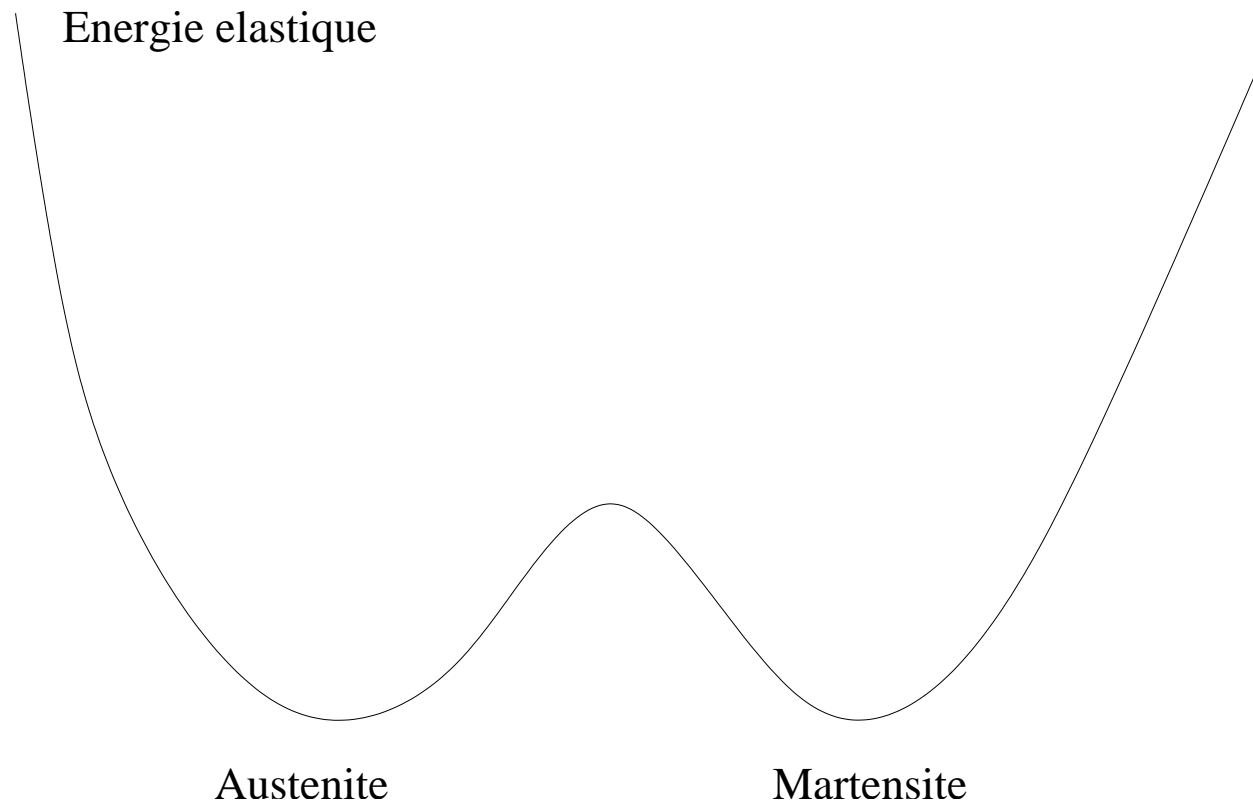


Cu-Al-Ni alloy (courtesy of YONG S. CHU)





J. Ball and R. James proposed the following mechanism: to sustain the applied forces, the alloy has a tendency to coexist under different phases, suitably aligned, which minimize the energy \Rightarrow **minimizing sequence !**



Convex analysis

To obtain the existence of minimizers we add a convexity assumption.

Definition. A set $K \subset V$ is said to be **convex** if, for any $x, y \in K$ and for any $\theta \in [0, 1]$, $(\theta x + (1 - \theta)y)$ belongs to K .

Definition. A function J , defined from a non-empty convex set $K \in V$ into \mathbb{R} is **convex** on K if

$$J(\theta u + (1 - \theta)v) \leq \theta J(u) + (1 - \theta)J(v) \quad \forall u, v \in K, \forall \theta \in [0, 1].$$

Furthermore, J is **strictly convex** if the inequality is strict whenever $u \neq v$ and $\theta \in]0, 1[$.

Existence result

Theorem. Let K be a non-empty closed convex set in a reflexive Banach space V , and J a **convex** continuous function on K , which is “infinite at infinity” in K , i.e.,

$$\forall (u^n)_{n \geq 0} \text{ sequence in } K, \quad \lim_{n \rightarrow +\infty} \|u^n\| = +\infty \implies \lim_{n \rightarrow +\infty} J(u^n) = +\infty .$$

Then, there exists a minimizer of J in K .

Remarks:

1. V reflexive Banach space $\Leftrightarrow (V')' = V$ (V' is the dual of V)
2. The theorem is still true if V is just the dual of a separable Banach space.
3. In practice, **this assumption is satisfied for all the functional spaces which we shall use**: for example, $L^p(\Omega)$ with $1 < p \leq +\infty$.

Uniqueness

Proposition. If J is **strictly convex**, then there exists at most one minimizer of J .

Proposition. If J is convex on the convex set K , then any **local minimizer** of J on K is a **global minimizer**.

Remark. For convex functions there is no difference between local and global minimizers.

Remark. Convexity is not the only tool to prove existence of minimizers. Another method is, for example, **compactness**.

Differentiability

Definition. Let V be a Banach space. A function J , defined from a neighborhood of $u \in V$ into \mathbb{R} , is said to be **differentiable in the sense of Fréchet** at u if there exists a continuous linear form on V , $L \in V'$, such that

$$J(u + w) = J(u) + L(w) + o(w) \quad , \quad \text{with} \quad \lim_{w \rightarrow 0} \frac{|o(w)|}{\|w\|} = 0 .$$

We call L the differential (or derivative, or gradient) of J at u and we denote it by $L = J'(u)$, or $L(w) = \langle J'(u), w \rangle_{V',V}$.

- ☞ If V is a Hilbert space, its dual V' can be identified with V itself thanks to the **Riesz representation theorem**. Thus, there exists a unique $p \in V$ such that $\langle p, w \rangle = L(w)$. We also write $p = J'(u)$.
- ☞ We use this identification $V = V'$ if $V = \mathbb{R}^n$ or $V = L^2(\Omega)$.
- ☞ In practice, it is often easier to compute the **directional derivative** $j'(0) = \langle J'(u), w \rangle_{V',V}$ with $j(t) = J(u + tw)$.

A basic example to remember

Consider the variational formulation

$$\text{find } u \in V \text{ such that } a(u, w) = L(w) \quad \forall w \in V$$

where a is a **symmetric** coercive continuous bilinear form and L is a continuous linear form.

Define the **energy**

$$J(v) = \frac{1}{2}a(v, v) - L(v)$$

Lemma. u is the unique minimizer of J

$$J(u) = \min_{v \in V} J(v)$$

Proof. We check that the optimality condition $J'(u) = 0$ is equivalent to the variational formulation.

Computing the directional derivative is simpler than computing $J'(v)$!

We define $j(t) = J(u + tw)$

$$j(t) = \frac{t^2}{2}a(w, w) + t\left(a(u, w) - L(w)\right) + J(u)$$

and we differentiate $t \rightarrow j(t)$ (a polynomial of degree 2 !)

$$j'(t) = ta(w, w) + \left(a(u, w) - L(w)\right).$$

By definition, $j'(0) = \langle J'(u), w \rangle_{V', V}$, thus

$$\langle J'(u), w \rangle_{V', V} = a(u, w) - L(w).$$

It is not obvious to deduce a formula for $J'(u)$...

but it is enough, most of the time, to know $\langle J'(u), w \rangle$.

Examples: (we use the "usual" scalar product in L^2)

$$1. J(v) = \int_{\Omega} \left(\frac{1}{2}v^2 - fv \right) dx \text{ with } v \in L^2(\Omega)$$

$$\langle J'(u), w \rangle = \int_{\Omega} (uw - fw) dx.$$

Thus

$$J'(u) = u - f \in L^2(\Omega) \text{ (identified with its dual)}$$

$$2. J(v) = \int_{\Omega} \left(\frac{1}{2}|\nabla v|^2 - fv \right) dx \text{ with } v \in H_0^1(\Omega)$$

$$\langle J'(u), w \rangle = \int_{\Omega} (\nabla u \cdot \nabla w - fw) dx.$$

Therefore, after integrating by parts,

$$J'(u) = -\Delta u - f \in H^{-1}(\Omega) = (H_0^1(\Omega))' \text{ (not identified with its dual)}$$

Remark (delicate). If instead of the "usual" scalar product in L^2 we rather use the H^1 scalar product, then we identify $J'(u)$ with a **different** function.

$$J(v) = \int_{\Omega} \left(\frac{1}{2} |\nabla v|^2 - fv \right) dx$$

From the directional derivative

$$\langle J'(u), w \rangle = \int_{\Omega} (\nabla u \cdot \nabla w - fw) dx,$$

using the H^1 scalar product $\langle \phi, w \rangle = \int_{\Omega} (\nabla \phi \cdot \nabla w + \phi w) dx$, we deduce

$$-\Delta J'(u) + J'(u) = -\Delta u - f, \quad J'(u) \in H_0^1(\Omega).$$

Here we identify $H_0^1(\Omega)$ with its dual.

Optimality conditions

Theorem (Euler inequality). Let $u \in K$ with K convex. We assume that J is differentiable at u . If u is a local minimizer of J in K , then

$$\langle J'(u), v - u \rangle \geq 0 \quad \forall v \in K .$$

If $u \in K$ satisfies this inequality and if J is convex, then u is a global minimizer of J in K .

Remarks.

- ➡ If u belongs to the interior of K , we deduce the **Euler equation** $J'(u) = 0$.
- ➡ The Euler inequality is usually just a necessary condition. It becomes **necessary and sufficient** for convex functions.

Minimization with equality constraints

$$\inf_{v \in V, F(v)=0} J(v)$$

with $F(v) = (F_1(v), \dots, F_M(v))$ differentiable from V into \mathbb{R}^M .

Definition. We call **Lagrangian** of this problem the function

$$\mathcal{L}(v, \mu) = J(v) + \sum_{i=1}^M \mu_i F_i(v) = J(v) + \mu \cdot F(v) \quad \forall (v, \mu) \in V \times \mathbb{R}^M.$$

The new variable $\mu \in \mathbb{R}^M$ is called **Lagrange multiplier** for the constraint $F(v) = 0$.

Lemma. The constrained minimization problem is equivalent to

$$\inf_{v \in V, F(v)=0} J(v) = \inf_{v \in V} \sup_{\mu \in \mathbb{R}^M} \mathcal{L}(v, \mu).$$

Stationarity of the Lagrangian

Theorem. Assume that J and F are continuously differentiable in a neighborhood of $u \in V$ such that $F(u) = 0$. If u is a local minimizer and if the vectors $(F'_i(u))_{1 \leq i \leq M}$ are **linearly independent**, then there exist Lagrange multipliers $\lambda_1, \dots, \lambda_M \in \mathbb{R}$ such that

$$\frac{\partial \mathcal{L}}{\partial v}(u, \lambda) = J'(u) + \lambda \cdot F'(u) = 0 \quad \text{and} \quad \frac{\partial \mathcal{L}}{\partial \mu}(u, \lambda) = F(u) = 0 .$$

Minimization with inequality constraints

$$\inf_{v \in V, F(v) \leq 0} J(v)$$

where $F(v) \leq 0$ means that $F_i(v) \leq 0$ for $1 \leq i \leq M$, with F_1, \dots, F_M differentiable from V into \mathbb{R} .

Definition. Let u be such that $F(u) \leq 0$. The set

$$I(u) = \{i \in \{1, \dots, M\}, F_i(u) = 0\}$$

is called the set of **active** constraints at u . The inequality constraints are said to be **qualified** at $u \in K$ if the vectors $(F'_i(u))_{i \in I(u)}$ are linearly independent.

Definition. We call **Lagrangian** of the previous problem the function

$$\mathcal{L}(v, \mu) = J(v) + \sum_{i=1}^M \mu_i F_i(v) = J(v) + \mu \cdot F(v) \quad \forall (v, \mu) \in V \times (\mathbb{R}^+)^M.$$

The new **non-negative** variable $\mu \in (\mathbb{R}^+)^M$ is called **Lagrange multiplier** for the constraint $F(v) \leq 0$.

Lemma. The constrained minimization problem is equivalent to

$$\inf_{v \in V, F(v) \leq 0} J(v) = \inf_{v \in V} \sup_{\mu \in (\mathbb{R}^+)^M} \mathcal{L}(v, \mu).$$

Stationarity of the Lagrangian

Theorem. We assume that the constraints are **qualified** at u satisfying $F(u) \leq 0$. If u is a local minimizer, there exist **Lagrange multipliers** $\lambda_1, \dots, \lambda_M \geq 0$ such that

$$J'(u) + \sum_{i=1}^M \lambda_i F'_i(u) = 0, \quad \lambda_i \geq 0, \quad \lambda_i = 0 \text{ if } F_i(u) < 0 \quad \forall i \in \{1, \dots, M\}.$$

This condition is indeed the stationarity of the Lagrangian since

$$\frac{\partial \mathcal{L}}{\partial v}(u, \lambda) = J'(u) + \lambda \cdot F'(u) = 0,$$

and the condition $\lambda \geq 0, F(u) \leq 0, \lambda \cdot F(u) = 0$ is equivalent to the Euler inequality for the **maximization** with respect to μ in the closed convex set $(\mathbb{R}^+)^M$

$$\frac{\partial \mathcal{L}}{\partial \mu}(u, \lambda) \cdot (\mu - \lambda) = F(u) \cdot (\mu - \lambda) \leq 0 \quad \forall \mu \in (\mathbb{R}^+)^M.$$

Interpreting the Lagrange multipliers

Define the Lagrangian for the minimization of $J(v)$ under the constraint $F(v) = c$

$$\mathcal{L}(v, \mu, c) = J(v) + \mu \cdot (F(v) - c)$$

We study the sensitivity of the minimal value with respect to variations of c .

Let $u(c)$ and $\lambda(c)$ be the minimizer and the optimal Lagrange multiplier. We assume that they are differentiable with respect to c . Then

$$\nabla_c \left(J(u(c)) \right) = -\lambda(c).$$

λ gives the derivative of the minimal value with respect to c without any further calculation ! Indeed

$$\nabla_c \left(J(u(c)) \right) = \nabla_c \left(\mathcal{L}(u(c), \lambda(c), c) \right) = \frac{\partial \mathcal{L}}{\partial c} (u(c), \lambda(c), c) = -\lambda(c)$$

because

$$\frac{\partial \mathcal{L}}{\partial v} (u(c), \lambda(c), c) = 0 \quad , \quad \frac{\partial \mathcal{L}}{\partial \mu} (u(c), \lambda(c), c) = 0 .$$

Duality and saddle point

Definition. Let $\mathcal{L}(v, q)$ be a Lagrangian. We call $(u, p) \in U \times P$ a **saddle point** (or mountain pass, or min-max) of \mathcal{L} in $U \times P$ if

$$\forall q \in P \quad \mathcal{L}(u, q) \leq \mathcal{L}(u, p) \leq \mathcal{L}(v, p) \quad \forall v \in U .$$

For $v \in U$ and $q \in P$, define $\mathcal{J}(v) = \sup_{q \in P} \mathcal{L}(v, q)$ and $\mathcal{G}(q) = \inf_{v \in U} \mathcal{L}(v, q)$. We call **primal problem**

$$\inf_{v \in U} \mathcal{J}(v) ,$$

and **dual problem**

$$\sup_{q \in P} \mathcal{G}(q) .$$

Example. $U = V$, $P = \mathbb{R}^M$ or \mathbb{R}_+^M , and $\mathcal{L}(v, q) = J(v) + q \cdot F(v)$. In this case $\mathcal{J}(v) = J(v)$ if $F(v) = 0$ and $\mathcal{J}(v) = +\infty$ otherwise, while there is no constraints for the dual problem (except the simple one, $q \in P$).

Lemma (weak duality). It always holds true that

$$\inf_{v \in U} \mathcal{J}(v) \geq \sup_{q \in P} \mathcal{G}(q).$$

Proof: $\inf \sup \mathcal{L} \geq \sup \inf \mathcal{L}$.

Theorem (strong duality). The couple (u, p) is a saddle point of \mathcal{L} in $U \times P$ if and only if

$$\mathcal{J}(u) = \min_{v \in U} \mathcal{J}(v) = \max_{q \in P} \mathcal{G}(q) = \mathcal{G}(p) .$$

Remark. The dual problem is often simpler than the primal one because it has no constraints. After solving the dual, the primal solution is obtained through an unconstrained minimization.

Application: dual or complementary energy

Very important for the sequel !

Let $f \in L^2(\Omega)$. We already know that solving

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

is equivalent to minimizing the (primal) energy

$$\min_{v \in H_0^1(\Omega)} \left\{ J(v) = \frac{1}{2} \int_{\Omega} |\nabla v|^2 dx - \int_{\Omega} f v dx \right\}$$

We introduce a dual or complementary energy

$$\max_{\substack{\tau \in L^2(\Omega)^N \\ -\operatorname{div} \tau = f \text{ in } \Omega}} \left\{ G(\tau) = -\frac{1}{2} \int_{\Omega} |\tau|^2 dx \right\}.$$

J is convex and G is concave.

Proposition. Let $u \in H_0^1(\Omega)$ be the unique solution of the p.d.e. Defining $\sigma = \nabla u$ we have

$$J(u) = \min_{v \in H_0^1(\Omega)} J(v) = \max_{\substack{\tau \in L^2(\Omega)^N \\ -\operatorname{div}\tau = f \text{ in } \Omega}} G(\tau) = G(\sigma),$$

and σ is the unique maximizer of G .

Proof. We define a Lagrangian in $H_0^1(\Omega) \times L^2(\Omega)^N$

$$\mathcal{L}(v, \tau) = -\frac{1}{2} \int_{\Omega} |\tau|^2 dx - \int_{\Omega} (f + \operatorname{div}\tau)v dx.$$

By integrating by parts

$$\mathcal{L}(v, \tau) = -\frac{1}{2} \int_{\Omega} |\tau|^2 dx - \int_{\Omega} f v dx + \int_{\Omega} \tau \cdot \nabla v dx.$$

v is the Lagrange multiplier for the constraint $-\operatorname{div}\tau = f$.

We check that the dual of the dual is the primal !

$$\max_{\tau} \mathcal{L}(v, \tau) = J(v).$$

End of the proof

By definition, if τ satisfies the constraint $-\operatorname{div}\tau = f$, we have

$$G(\tau) = \mathcal{L}(v, \tau) \quad \forall v$$

On the other hand,

$$\mathcal{L}(v, \tau) \leq \max_{\tau} \mathcal{L}(v, \tau) = J(v).$$

Besides, integrating by parts yields $\int_{\Omega} |\nabla u|^2 dx = \int_{\Omega} f u dx$, thus

$$J(u) = \frac{1}{2} \int_{\Omega} |\nabla u|^2 dx - \int_{\Omega} f u dx = -\frac{1}{2} \int_{\Omega} |\nabla u|^2 dx = G(\nabla u).$$

In other words, for any τ satisfying $-\operatorname{div}\tau = f$,

$$G(\tau) = \mathcal{L}(u, \tau) \leq J(u) = G(\sigma)$$

which means that $\sigma = \nabla u$ is a maximizer of G among all τ 's such that $-\operatorname{div}\tau = f$.

Numerical algorithms for minimization problems

A simplified classification:

☞ Stochastic algorithms: **global minimization**. Examples: Monte-Carlo, simulated annealing, genetic. See the last chapter and the last course.

Inconvenient: high CPU cost.

☞ Deterministic algorithms: **local minimization**. Examples: gradient methods, Newton.

Inconvenient: they require the gradient of the objective function.

Gradient descent with an optimal step

The goal is to solve

$$\inf_{v \in V} J(v) .$$

Initialization: choose $u^0 \in V$. **Iterations:** for $n \geq 0$

$$u^{n+1} = u^n - \mu^n J'(u^n) ,$$

where $\mu^n \in \mathbb{R}$ is chosen at each iteration such that

$$J(u^{n+1}) = \inf_{\mu \in \mathbb{R}^+} J(u^n - \mu J'(u^n)) .$$

Main idea: if $u^{n+1} = u^n - \mu w^n$ with a small $\mu > 0$, then

$$J(u^{n+1}) = J(u^n) - \mu \langle J'(u^n), w^n \rangle + \mathcal{O}(\mu^2),$$

thus, to decrease J , the best "first order" choice is w^n proportional to $J'(u^n)$.

Convergence

Theorem Assume that J is differentiable, strongly convex with $\alpha > 0$,

$$\langle J'(u) - J'(v), u - v \rangle \geq \alpha \|u - v\|^2 \quad \forall u, v \in V,$$

and J' is Lipschitzian on any bounded set of V , i.e.,

$$\forall M > 0, \quad \exists C_M > 0, \quad \|v\| + \|w\| \leq M \Rightarrow \|J'(v) - J'(w)\| \leq C_M \|v - w\|.$$

Then the gradient algorithm with an optimal step **converges**: for any u^0 , the sequence (u^n) converges to the unique minimizer u .

Remark. If J is not strongly convex:

- ☞ the algorithm may not converge because it oscillates between several minimizers,
- ☞ the algorithm may converge to a local minimizer,
- ☞ the minimizer obtained by the algorithm may vary with the initialization.

Gradient descent with a fixed step

The goal is to solve

$$\inf_{v \in V} J(v) .$$

Initialization: choose $u^0 \in V$. **Iterations:** for $n \geq 0$

$$u^{n+1} = u^n - \mu J'(u^n) ,$$

Theorem. Assume that J is differentiable, strongly convex, and J' is Lipschitzian on any bounded set of V . Then, if $\mu > 0$ is small enough, the gradient algorithm with fixed step converges: for any u^0 , the sequence (u^n) converges to the unique minimizer u .

Remark. An intermediate variant is: increase the step, $\mu_{n+1} = 1.1 \times \mu_n$, if J decreases, and reduce the step, $\mu_{n+1} = 0.5 \times \mu_n$, if J increases.

Projected gradient

Let K be a non-empty closed convex subset of V . The goal is to solve

$$\inf_{v \in K} J(v) .$$

Initialization: choose $u^0 \in K$. **Iterations:** for $n \geq 0$

$$u^{n+1} = P_K(u^n - \mu J'(u^n)) ,$$

where P_K is the projection on K .

Theorem. Assume that J is differentiable, strongly convex, and J' is Lipschitzian on any bounded set of V . Then, if $\mu > 0$ is small enough, the projected gradient algorithm with fixed step converges.

Remark. Another possibility is to **penalize** the constraints, i.e., for small $\epsilon > 0$ we replace

$$\inf_{v \in V, F(v) \leq 0} J(v) \quad \text{by} \quad \inf_{v \in V} \left(J(v) + \frac{1}{\epsilon} \sum_{i=1}^M [\max(F_i(v), 0)]^2 \right) .$$

Examples of projection operators P_K

☞ If $V = \mathbb{R}^M$ and $K = \prod_{i=1}^M [a_i, b_i]$, then for $x = (x_1, x_2, \dots, x_M) \in \mathbb{R}^M$

$$P_K(x) = y \quad \text{with} \quad y_i = \min(\max(a_i, x_i), b_i) \quad \text{pour} \quad 1 \leq i \leq M.$$

☞ If $V = \mathbb{R}^M$ and $K = \{x \in \mathbb{R}^M \mid \sum_{i=1}^M x_i = c_0\}$, then

$$P_K(x) = y \quad \text{with} \quad y_i = x_i - \lambda \quad \text{and} \quad \lambda = \frac{1}{M} \left(-c_0 + \sum_{i=1}^M x_i \right).$$

☞ Same if $V = L^2(\Omega)$ and $K = \{\phi \in V \mid a(x) \leq \phi(x) \leq b(x)\}$ or
 $K = \{\phi \in V \mid \int_{\Omega} \phi dx = c_0\}$.

For more general closed convex sets K , P_K can be very hard to determine. In such cases one rather uses the [Uzawa algorithm](#) which looks for a saddle point of the Lagrangian.

Newton algorithm (of order 2)

Main idea: if $V = \mathbb{R}^N$ and if $J'' \geq 0$

$$J(w) \approx J(v) + J'(v) \cdot (w - v) + \frac{1}{2} J''(v)(w - v) \cdot (w - v),$$

the minimizer of which is $w = v - (J''(v))^{-1} J'(v)$.

Algorithm: $u^{n+1} = u^n - (J''(u^n))^{-1} J'(u^n)$.

☞ It converges faster if u^0 is close form the minimizer u

$$\|u^{n+1} - u\| \leq C \|u^n - u\|^2 .$$

☞ It requires solving a linear system with the matrix $J''(u^n)$.

☞ It can be generalized in a quasi-Newton method (without computing J'') or to the constrained case.