DERIVATIVE AND OPTIMIZATION

For PDEs applications, differential calculus is useful for

- elementary differential manipulations in PDEs;
- basis of optimization and constrained optimization possibly in infinite dimension (1st course, SM). This topic will be developed further in the fundamental course *Introduction to non-linear PDEs*;
- elementary differential geometry in the treatment of the boundary terms in PDEs defined in a domain of $\mathbb{R}^d$;
- applications of the Stokes/Gauss-Green formula and the mean value theorem to PDEs.

1. Derivative

For two Banach spaces X, Y, we denote by $\mathcal{L}(X,Y)$ the space of linear and **continuous** mapping between X and Y. We recall that $\mathcal{L}(X,Y)$ is a Banach space when endowed with the usual operator norm. We write $\mathcal{L}(X) := \mathcal{L}(X,X)$.

Definition 1.1 (Fréchet)**.** *Consider two Banach spaces X and Y, an open set $\Omega \subset X$ and a mapping $f : \Omega \to Y$.*
(i) We say that f is continuous in $u \in \Omega$ if

$$\|f(v) - f(u)\| = o(1), \text{ when } v \to u.$$

We say that f is continuous on Ω if it is continuous in every point of Ω.
(ii) We say that f is Fréchet-differentiable in $u \in \Omega$ if there exists $L \in \mathcal{L}(X,Y)$ s.t.

$$\|f(v) - f(u) - L(v-u)\| = o(\|v-u\|), \text{ when } v \to u,$$

we note $L = Df(u) = df_u$. We say that f is Fréchet-differentiable on Ω if it is Fréchet-differentiable in every point of Ω.
(iii) We say that f is C^1 if f is Fréchet-differentiable on Ω and $Df : \Omega \to \mathcal{L}(X,Y)$, $u \mapsto Df(u)$, is continuous on Ω.

Definition 1.2 (Gâteaux)**.** *Consider two Banach space X and Y, a subset $A \subset X$ (not necessarily open) and a mapping $f : A \to Y$.*
(iv) We say that f is Gâteaux-differentiable in $u \in A$ if there exists $L \in \mathcal{L}(X,Y)$ and for any direction $w \in X$ such that $u + tw \in A$ for any $t > 0$ small enough there holds

$$\|f(u+tw) - f(u) - tLw\| = o(t), \text{ when } t \to 0^+,$$

we note $L = f'(u)$. We say that f is Gâteaux-differentiable on Ω if it is Gâteaux-differentiable in every point of Ω.

Lemma 1.3. *Consider two Banach spaces X and Y, an open set $\Omega \subset X$ and a mapping $F : \Omega \to Y$. There holds:*
(1) F is Fréchet-differentiable at some point implies that F is continuous and is Gâteaux-differentiable at the same point with $DF(u) = F'(u)$.
(2) F is Gâteaux-differentiable on Ω, F and DF are continuous on Ω if, and only if, F is C^1.

Proof of Lemma 1.3. The point (1) being clear, we only establish the point (2), and more precisely that F is Gâteaux-differentiable on Ω, F and DF are continuous on Ω imply that F is Fréchet-differentiable on Ω, since then it is C^1. We take $u, v \in \Omega$ such that $[u, v] := \{u + t(v - u);\ t \in [0, 1]\} \in \Omega$, and we write

$$F(v) - F(u) = \int_0^1 \frac{d}{dt}[F(u + t(v - u))]\, dt = \int_0^1 DF(u + t(v - u))(v - u)\, dt,$$

which classically makes sense when $Y = \mathbb{R}$ (and we accept here that the integral makes sense also for Banach space valued functions). We deduce

$$\begin{aligned} &F(v) - F(u) - DF(u)(v - u) = \\ &= \int_0^1 [DF(u + t(v - u)) - DF(u)](v - u)dt \\ &= \mathcal{O}\Big(\|v - u\| \sup_{v' \in [u,v]} \|DF(v') - DF(u)\|\Big) = o(\|v - u\|), \end{aligned}$$

so that F is Fréchet-differentiable in u. $\square$

We wish to emphasize on the Taylor-Laplace expansion (of order one) for a C^1 function F:

$$F(v) = F(u) + \int_0^1 DF(u + t(v - u))(v - u)\, dt,$$

for any $[u, v] \subset \Omega$. Since the integrand is continuous, the integral here may be understood as a Riemann integral, i.e. defined as the limit of a Riemann sum.

We give now some examples.

• *Exemple 1.1.* When $X \simeq \mathbb{R}^m$, $Y \simeq \mathbb{R}^n$, we may introduce two basis $(e_1, \dots, e_m)$ of X and $(\varepsilon_1, \dots, \varepsilon_n)$ of Y respectively and for a differentiable function $f : \Omega \subset X \to Y$, we may abuse notations by writing

$$f(x) = \sum_{i=1}^n f_i(x)\varepsilon_i = \sum_{i=1}^n f_i(x_1, \dots, x_m)\varepsilon_i, \quad \forall\, x := \sum_{j=1}^m x_j e_j.$$

We then introduce the partial derivatives notation

$$\partial_j f_i(x) = \frac{\partial f_i}{\partial x_j}(x) := Df_i(x)e_j := \lim_{t \to 0^+} \frac{f_i(x + te_j) - f_i(x)}{t},$$

and we can make the identification

$$Df(x) \simeq \big(\partial_j f_i(x)\big)_{ij}.$$

It is worth mentioning that point (2) in Lemma 1.3 is nothing but the classical characterization of C^1 functions thanks to the continuity property of its first order partial derivatives.

• *Exemple 1.2.* Consider $X = H$ a Hilbert space with scalar product $(\cdot,\cdot)$, $A \in \mathcal{L}(H)$, and let us define $f : H \to \mathbb{R}$, $f(u) := (Au, u)$. Defining the adjoint A^* of A through the relation $(A^*u, v) = (u, Av)$ for any $u, v \in H$, we easily compute $Df(u) = A + A^*$, or in other words

$$Df(u) \in H', \quad Df(u)w = (Au, w) + (u, Aw).$$

To see this, we write

$$f(u + w) = f(u) + (Au, w) + (u, Aw) + (Aw, w),$$

we observe that $|(Aw, w)| = \mathcal{O}(\|w\|^2)$ and we come back to the very definition of $Df(u)$ in Definition 1.1-(ii).
In particular, $Df(u) = 2A$ when A is self-adjoint. When $H = L^2(\Omega)$, $g(u) := \|u\|^2_{L^2}$ and when $H = H^1_0 := \{u \in L^2;\, \nabla u \in (L^2)^d,\, u = 0 \text{ a.e. on } \partial\Omega\}$, $h(u) := \|\nabla u\|^2_{L^2}$, we aslo find

$$Dg(u)w = \int uw, \quad Dh(u)w = \int \nabla u \cdot \nabla w,$$

by performing the same kind of expansion. Alternatively, $H^1_0(\Omega)$ may be defined as the Hilbert space obtained by completion of the space $C^1_c(\Omega)$ for the norm $u \mapsto \|u\|_{H^1}$, with $\|u\|^2_{H^1} := \|u\|^2_{L^2} + \|\nabla u\|^2_{L^2}$. Anyway, in these notes, we manipulate a $H^1_0(\Omega)$ function u as if it was a $C^1(\bar\Omega)$ function such that $u = 0$ on the boundary $\partial\Omega$. We refer to the companion course *A review of functional analysis tools for PDEs* for details.

• *Exemple 1.3.* Consider $1 \le p < \infty$, $f : \Omega \times \mathbb{R} \to \mathbb{R}$ a continuous fonction such that $|f(x, s)| \le a(x) + b|s|^{p/q}$, with $1 \le q < \infty$, $a \in L^q(\Omega)$, $b \ge 0$, and define

$$A : L^p(\Omega) \to L^q(\Omega), \quad A(u)(x) := f(x, u(x)).$$

Then A is well-defined and continuous. Indeed, if $u_n \to u$ in L^p, we have $u_{n'} \to u$ a.e. and $|u_{n'}| \le u^* \in L^p$ for some subsequence $(u_{n'})$ (from the partial reverse of the Lebesgue convergence theorem), so that $A(u_{n'}) \to Au$ a.e. and $|A(u_{n'})| \le a + b|u^*|^{p/q}$ and thus $A(u_{n'}) \to Au$ in L^q (from the Lebesgue convergence theorem). Since this true for any subsequence $(u_{n'})$, we have $A(u_n) \to Au$ in L^q.

• *Exemple 1.4.* We consider $f : \Omega \times \mathbb{R} \to \mathbb{R}$ a continuous fonction and we set

$$F(x, s) := \int_0^s f(x, \sigma) d\sigma.$$

For $p \in (1, \infty)$ and $p' := p/(p-1)$, we assume that

$$\begin{aligned} |f(x, s)| &\le a_0(x) + b_0|s|^{p-1}, \quad a_0 \in L^{p'}(\Omega),\, b_0 \ge 0, \\ |F(x, s)| &\le a_1(x) + b_1|s|^p, \quad a_1 \in L^1(\Omega),\, b_1 \ge 0. \end{aligned}$$

We consider in $X = L^p(\Omega)$, the mapping

$$\mathcal{F} : X \to \mathbb{R}, \quad \mathcal{F}(u) := \int_\Omega F(x, u(x))\, dx.$$

Then F is C^1 and $D\mathcal{F}(u) = Au$, where A is the operator defined in Example 1.3.
- On the one hand, F is continuous (from X into $\mathbb{R}$) because the mapping $B : L^p \to L^1$, $u \mapsto (Bu)(x) := F(x, u(x))$, is continuous as proved in Example 1.3. Similarly,

$A : L^p \to L^{p'} = (L^p)'$ is continuous. On the other hand, for $u, w \in L^p$, we compute

$$\frac{\mathcal{F}(u+tw) - \mathcal{F}(u)}{t} = \int_\Omega \frac{F(x, u(x) + tw(w)) - F(x, u(x))}{t}\, dx = \int_\Omega f(x, v_t(x)) w(w)\, dx,$$

for some $v_t(x) \in [u(x), u(x)+tw(x)]$ from the mean value theorem. We thus observe that

$$f(x, v_t(x)) \to f(x, u(x)) \text{ when } t \to 0,$$
$$|f(t, v_t(x))| \le a_0(x) + b_0|v_t(x)|^{p-1} \le a_0(x) + b_0(|u(x)| + |w(x))|^{p-1} \in L^{p'}, \quad \forall\, t \in [0,1].$$

We deduce

$$\frac{\mathcal{F}(u+tw) - \mathcal{F}(u)}{t} \to \int_\Omega f(x, u)) w(x)\, dx, \quad \text{as } t \to 0,$$

from the Lebesgue convergence theorem and the Holder inequality.

For later reference, we observe that if $F : X \to Y$ and $G : Y \to Z$ are C^1 functions (for instance), then

$$D_u G \circ F = D_{F(u)} G \circ D_u F,$$

with $D_u F \in \mathcal{L}(X, Y)$, $D_{F(u)} G \in \mathcal{L}(Y, Z)$ and $D_u G \circ F \in \mathcal{L}(X, Z)$. In particular, when $X = \mathbb{R}$, $u : \mathbb{R} \to X$, $F : X \to Y$, we have

$$\frac{d}{dt}\, F \circ u(t) = DF(u(t)) u'(t), \tag{1.1}$$

with $(F \circ u)'(t) \in Y = \mathcal{L}(\mathbb{R}, Y)$, $DF(u(t)) \in \mathcal{L}(X, Y)$ and $u'(t) \in X = \mathcal{L}(\mathbb{R}, X)$. We also observe that when $F : X_1 \times X_2 \to Y$, we may write

$$DF(u)(w_1, w_2) = D_1 F(u) w_1 + D_2 F(u) w_2, \tag{1.2}$$

with $D_i F(u) \in \mathcal{L}(X_i, Y)$ and $w_i \in X_i$.

For two Banach spaces X, Y, we denote by $\mathcal{B}(X, Y)$ the space of **bilinear** and **continuous** mapping from X^2 into Y. We recall that $\mathcal{B}(X, Y) \simeq \mathcal{L}(X, \mathcal{L}(X, Y))$.

Definition 1.4 (class C^2)**.** *Consider two Banach spaces X and Y, an open set $\Omega \subset X$ and a mapping $f : \Omega \to Y$. We say that f is C^2 if both f and Df are C^1. We note $D^2 f := D(Df) \in \mathcal{B}(X, Y)$.*

Lemma 1.5 (Schwarz theorem and Taylor-Laplace expansion)**.** *Consider a Banach space X, an open set $\Omega \subset X$ and $f : \Omega \to \mathbb{R}$ of class C^2.*
(1) - Then $D^2 f$ is symmetric, or in other words

$$D^2 f(u)(v, w) = D^2 f(u)(v, w), \quad \forall\, u \in \Omega,\ \forall\, v, w \in X.$$

In particular, in finite dimension $X \simeq \mathbb{R}^m$, the Hessian $D^2 f(u) := (\partial^2_{ij} f(u))_{ij}$ is a symmetric matrix.
The Taylor-Laplace expansion (of order two) for a C^2 function f writes:

$$f(u+w) = f(u) + \langle Df(u), w\rangle + \int_0^1 (1-t) D^2 f(u_t)(w, w) dt, \tag{1.3}$$

for any $u \in \Omega$, $w \in X$ such that $[u, u+w] \subset \Omega$ and we denote $u_t := u + tw$.

2. Optimization and convex functions

Lemma 2.1. *Consider X a Banach space, $K \subset X$ a compact set and $F : K \to \mathbb{R}$ a lsc function. Then, the associated minimizer problem has at least one solution, namely:*

$$\exists\, u^* \in K, \quad F(u^*) = \min_{u \in K} F(u) \in \mathbb{R}. \tag{2.1}$$

If F is Gâteaux-differentiable and $u^* \in \mathring{K} \neq \emptyset$ satisfies (2.1)*, then the following Euler equation holds*

$$F'(u^*) = 0. \tag{2.2}$$

If additionally F is C^2, there holds

$$D^2 F(u^*) \ge 0. \tag{2.3}$$

Proof of Lemma 2.1. Take $u_n \in K$ such that $F(u_n) \to I := \inf F \in \mathbb{R} \cup \{-\infty\}$. By compactness, there exists a subsequence $(u_{n'})$ such that $u_{n'} \to u^* \in K$. Because F is lsc, we find

$$-\infty < F(u^*) \le \liminf F(u_{n'}) = I.$$

Because $u^* \in K$, we find $F(u^*) \ge I$, and thus (2.1) holds. For proving (2.2), we fixe arbitrarily $w \in X$, and we compute

$$\langle F'(u^*), w \rangle = \lim_{t \to 0^+} \frac{1}{t}(F(u^* + tw) - F(u^*)) \ge 0.$$

We conclude to (2.2) by writing the same inequality for $-w$ instead of w. When furthermore F is C^2, the Lagrange expansion (1.3) implies

$$0 \le \frac{1}{t^2}(f(u^* + tw) - f(u^*)) = \int_0^1 (1-s) D^2 f(u^*_{st})(w,w)\, ds$$

for any $w \in X$ and $t > 0$ small enough, where $u^*_{st} := u^* + stw$. By passing to the limit $t \to 0$, we get

$$0 \le \frac{1}{2} D^2 f(u^*)(w,w), \quad \forall\, w \in X,$$

which precisely means that $D^2 f(u^*)$ is positive. □

(Strong) compact set in infinity dimensional space are scarce. That explains the importance of convex optimization in which framework we may overcome this difficulty.

Definition 2.2 (Convexity)**.** *Consider a Banach space X.*
(i) We say that $K \subset X$ is convex if

$$\forall\, u, v \in K, \quad [u,v] \subset K.$$

(ii) For a convex set $K \subset X$, we say that $f : K \to \mathbb{R}$ is convex if

$$\forall\, u, v \in K,\ \forall\, t \in [0,1], \quad f(tu + (1-t)v) \le t f(u) + (1-t) f(v).$$

Similarly, we say that $f : K \to \mathbb{R}$ is strictly convex if

$$\forall\, u, v \in K,\ u \neq v,\ \forall\, t \in]0,1[, \quad f(tu + (1-t)v) < t f(u) + (1-t) f(v).$$

We have the following characterization of convex functions.

Lemma 2.3 (Convex functions). *Consider a Banach space X, a convex set $K \subset X$ and a Gâteaux-differentiable function $F : K \to \mathbb{R}$. The following properties are equivalent:*

(i) F is convex;

(ii) For any $u, v \in K$, $F(v) \geq F(u) + \langle F'(u), v - u\rangle$;

(iii) F' is monotonous, which means

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

When furthermore K is an open set and F is C^2, these properties are equivalent to (iv) $D^2F(u) \geq 0$, $\forall\, u \in K$.

Proof of Lemma 2.3. (i) $\Rightarrow$ (ii). We write the convexity inequality as

$$F(u + t(v - u)) - F(u) \leq t(F(v) - F(u)),$$

we divide by t and we pass to the limit $t \to 0$.

(ii) $\Rightarrow$ (iii). Exchanging the role of u and v in (ii), we have

$$\begin{aligned} F(v) &\geq F(u) + \langle F'(u), v - u\rangle \\ F(u) &\geq F(v) + \langle F'(v), u - v\rangle, \end{aligned}$$

for any $u, v \in K$. We conclude by summing up these inequalities.

(iii) $\Rightarrow$ (i). We define $\varphi(t) := F(w_t)$, $w_t := u + t(v - u))$. Because

$$\varphi'(t) = \langle F'(w_t), v - u\rangle,$$

the condition (iii) implies

$$\varphi'(t) - \varphi'(s) = \frac{1}{t - s}\langle F'(w_t) - F'(w_s), w_t - w_s\rangle \geq 0,$$

for $t > s$. The mapping $t \mapsto \varphi'(t)$ is thus increasing, and in particular $\varphi'(ts) \leq \varphi'(s)$, for any $t, s \in [0, 1]$. As a consequence, we have

$$\varphi(t) - \varphi(0) = t \int_0^1 \varphi'(ts)\, ds \leq t \int_0^1 \varphi'(s)\, ds = t(\varphi(1) - \varphi(0)),$$

which means that F is convex.

(iv) $\Rightarrow$ (ii). From the Taylor-Laplace expansion (1.3), we have

$$\begin{aligned} F(u + w) &= F(u) + \langle DF(u), w\rangle + \int_0^1 (1 - t) D^2F(u_t)(w, w) dt \\ &\geq F(u) + \langle DF(u), w\rangle, \end{aligned}$$

when (iv) holds.

(iii) $\Rightarrow$ (iv). From the very definition

$$D^2F(u)(w, w) = \lim_{t \to 0^+} \frac{1}{t^2}\langle DF(u + tw) - DF(u), tw\rangle \geq 0,$$

when (iii) holds. $\square$

Similarly, we have the following characterization of strictly convex functions.

Lemma 2.4 (Strictly convex functions)**.** *Consider a Banach space X, a convex set $K \subset X$ and a Gâteaux-differentiable function $F : K \to \mathbb{R}$. The following properties are equivalent:*

(i) F is strictly convex;

(ii) For any $u, v \in K$, $F(v) > F(u) + \langle F'(u), v - u\rangle$.

(iii) F' is strictly monotonous, which means

$$\langle F'(v) - F'(u), v - u\rangle > 0, \quad \forall\, u, v \in K, u \neq v.$$

When we assume furthermore that K is an open set and F is C^2, these properties are a consequence of

(iv) $D^2F(u) > 0$, $\forall\, u \in K$.

We accept the following result (for which we refer to a functional analysis course).

Theorem 2.5 (Banach-Alaoglu)**.** *Consider a bounded sequence (u_n) in a Hilbert space H. Then, there exists a subsequence $(u_{n'})$ and $u^* \in H$ such that*

$$\forall\, v \in H, \quad (u_{n'}, v) \to (u^*, v) \quad \textit{as} \quad n' \to \infty.$$

We say that $(u_{n'})$ converges weakly to u^*, we note $u_{n'} \rightharpoonup u^*$.

We formulate a simple form of optimization result in infinite dimension.

Theorem 2.6. *Consider a Hilbert space H and a convex and Gâteaux-differentiable function $F : H \to \mathbb{R}$ such that $F(u) \to +\infty$ when $|u| \to \infty$. Then, there exists at least one $u^* \in H$ such that*

$$F(u^*) = \min_{u \in H} F(u) \in \mathbb{R}, \quad F'(u^*) = 0. \tag{2.4}$$

When furthermore F is strictly convex, then u^* is unique.

Proof of Theorem 2.6. By assumption, $\{u \in H;\, F(u) \leq F(0)\} \subset B(0, R)$ for some $R > 0$. Consider next a sequence (u_n) such that

$$\lim F(u_n) = \inf_{u \in H} F(u) = \inf_{u \in B_R} F(u).$$

We may thus assume (u_n) bounded and, from Theorem 2.5, there exists $u^* \in H$ and a subsequence $(u_{n'})$ such that $u_{n'} \rightharpoonup u^*$. From Lemma 2.3-(ii), we have

$$F(u^*) \leq F(u_{n'}) - \langle F'(u^*), u_{n'} - u^*\rangle,$$

and thus

$$F(u^*) \leq \liminf F(u_{n'}) = I.$$

The equation $F'(u^*) = 0$ comes from (2.2). When furthermore F is strictly convex, Lemma 2.4-(iii) implies that $F'(v) \neq 0$ if $v \neq u^*$, and thus $F(v) \neq I$ if $v \neq u^*$. □

For $f \in L^2(\Omega)$, $\Omega \subset \mathbb{R}^d$ an open and bounded set, we consider the mapping $\mathcal{E} : H_0^1(\Omega) \to \mathbb{R}$ defined by

$$\mathcal{E}(u) := \frac{1}{2} \int_\Omega |\nabla u|^2 - \int_\Omega u f.$$

The existence of a unique solution to the minimization problem

$$u \in H_0^1, \quad \mathcal{E}(u) = \min_{v \in H_0^1} \mathcal{E}(v)$$

can be established thanks to Theorem 2.6 and a functional inequalities (called the Poincaré inequality). We rather refer to the companion course *"A review of functional analysis tools for PDEs"* for details. We just emphasize that the associated Euler equation $\mathcal{E}'(u) = 0$ is nothing but

$$\mathcal{E}'(u)w = \int_\Omega \nabla u \cdot \nabla w - \int_\Omega wf = 0, \quad \forall\, w \in H^1_0(\Omega).$$

When u is smooth enough and taking advantage of the vanishing condition $w = 0$ on $\partial\Omega$, the Stokes formula (which will be the subject of the next course) implies

$$\begin{aligned} \int_\Omega \nabla u \cdot \nabla w &= \int_\Omega \operatorname{div}(w\nabla u) - \int_\Omega \Delta u\, w \\ &= \int_{\partial\Omega} w\nabla u \cdot n\, d\sigma - \int_\Omega \Delta u\, w = -\int_\Omega \Delta u\, w. \end{aligned}$$

As a consequence, we equivalently have

$$\mathcal{E}'(u)w = \int_\Omega (-\Delta u - f)w = 0, \quad \forall\, w \in H^1_0(\Omega),$$

so that the Euler equation is a weak formulation for the Laplace equation

$$-\Delta u = f \text{ in } \Omega, \quad u = 0 \text{ on } \partial\Omega.$$

3. Constrained optimization

In many cases, we wish to minimize a functional on a part of a vectorial space which writes as

$$K := \{v \in X;\ G(v) = 0\}, \tag{3.1}$$

for a given function $G : X \to \mathbb{R}^m$. A typical example is given by $G(v) = \|v\|^2 - 1$ when $m = 1$, so that K is nothing but the circle

$$K = \{v \in X;\ \|v\| = 1\}.$$

We start by recalling the implicit function theorem. First consider the simple situation when $f : \mathbb{R} \times X \to Z$ and we want to find solutions to the equation

$$f(t, u) = 0.$$

We assume that $f(t_0, u_0) = 0$ for some $(t_0, u_0) \in \mathbb{R} \times X$ and we look for a solution (t, u) for any t in a neighborhood of t_0. We may reformulate the problem as finding a smooth mapping $u : I \to X$, $I \subset \mathbb{R}$, such that

$$f(t, u(t)) = 0, \quad \forall\, t \in I.$$

Differentiating and using (1.1) and (1.2), we find

$$D_1F(t, u(t)) + D_2F(t, u(t))u'(t) = 0, \quad \forall\, t \in I,$$

where $D_1F(t, u(t)) \in Z = \mathcal{L}(\mathbb{R}, Z)$ and $D_2F(t, u(t)) \in \mathcal{L}(X, Z)$. In other word, the problem reduces to the ODE

$$u'(t) = -[D_2F(t, u(t))]^{-1} D_1F(t, u(t)), \quad u(t_0) = u_0,$$

provided that $D_2F(t, u(t))$ is invertible.

Theorem 3.1 (implicit function)**.** *Consider three Banach spaces X, Y and Z, $\Omega \subset X \times Y$ an open set and $f \in C^1(\Omega, Z)$. Assume that $(u_0, v_0) \in \Omega$ satisfies $f(u_0, v_0) = 0$ and $D_2 f(u_0, v_0)$ is invertible. Then, there exists an open set $U \subset X$ such that $u_0 \in U$ and a unique $\varphi : U \to Y$ of class C^1 such that $\varphi(u_0) = v_0$ and*

$$f(u, \varphi(u)) = 0, \quad \forall\, u \in U.$$

Moreover, there exists an open set $V \subset Y$ such that $v_0 \in V$ and if $(u, v) \in U \times V$ satisfies $f(u, v) = 0$ then $v = \varphi(u)$. Finally, we have

$$D\varphi(u) = -[D_2 f(u, \varphi(u))]^{-1} \circ D_1 f(u, \varphi(u)).$$

Theorem 3.2 (Lagrange multiplier)**.** *Consider a Banach space X and two smooth mapping $F : X \to \mathbb{R}$, $G : X \to \mathbb{R}^m$. If u_0 is a solution to the constrained minimization problem*

$$u_0 \in K, \quad F(u_0) = \min_{v \in K} F(u),$$

where K is defined by (3.1) and $G_1'(u_0), \dots, G_m'(u_0)$ are linearly independant, then there exist $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ such that

$$F'(u_0) = \sum_{i=1}^{m} \lambda_i G_i'(u_0). \tag{3.2}$$

In particular, when $m = 1$, that means $G'(u_0) \neq 0$ and there exists $\lambda \in \mathbb{R}$ such that

$$F'(u_0) = \lambda G'(u_0). \tag{3.3}$$

Proof of Theorem 3.2. We only consider the case $m = 1$. The condition $G'(u_0) \neq 0$, with $G'(u_0) \in \mathcal{L}(X, \mathbb{R}) \simeq X'$, implies that there exists $a \in X$ such that $\langle G'(u_0), a\rangle = 1$, and thus $X = X_0 \otimes \mathbb{R}a$, with $X_0 := \ker G'(u_0)$. Consider the mapping

$$\Phi : X_0 \times \mathbb{R} \to \mathbb{R}, \ \Phi(t, v) := G(u_0 + v + ta).$$

We see that $\Phi(0, 0) = G(u_0) = 0$ and

$$\partial_t \Phi(0, 0) = \langle G'(u_0), a\rangle = 1, \quad \partial_v \Phi(0, 0) = G'(u_0)_{|X_0} = 0.$$

The implicit function Theorem 3.1 implies that there exists a neighborhood $\omega \subset X_0$ such that $0 \in \omega$ and a C^1 function $\psi : \omega \to \mathbb{R}$ such that $\Phi(v, \psi(v)) = 0$ for any $v \in \omega$. By construction $\psi(0) = 0$, $\psi'(0) = 0$. We thus can find a neighborhood $\Omega \subset X$ of u_0 and a neighborhood $I \subset \mathbb{R}$ of 0 such that

$$\begin{aligned} u \in \Omega \text{ and } \ G(u) = 0 \quad &\Leftrightarrow \quad u = u_0 + v + ta, \ v \in \omega, \ t \in I, \ \Phi(t, v) = 0 \\ &\Leftrightarrow \quad u = u_0 + v + \psi(v)a, \ v \in \omega. \end{aligned}$$

Finally define $J : \omega \subset X_0 \to \mathbb{R}$ by $J(v) := F(u_0 + v + a\psi(v))$ so that its minimum on ω is reached in $v = 0$. We deduce that $J'(0) = 0$, or in other words

$$\langle J'(0), w\rangle = 0, \quad \forall\, w \in X_0.$$

From the definition of J and ψ, we have $J'(0) = F'(u_0)(I - a\psi'(0)) = F'(u_0)$, so that

$$\langle F'(u_0), w\rangle = 0, \quad \forall\, w \in X_0.$$

Since $F'(u_0) \in \mathcal{L}(X, \mathbb{R}) = X'$ is a linear form on X with $\ker F'(u_0) \supset \ker G'(u_0)$, there exists $\lambda \in \mathbb{R}$ such that (3.3) holds. $\square$

The proof of Theorem 3.2 in the case $m > 1$ is similar to the proof and it uses the classical algebraic lemma.

Lemma 3.3. *Consider a vector space X and $f_0, \dots, f_m \in X^*$ such that*

$$\bigcap_{1 \le i \le m} ker(f_i) \subset ker(f_0).$$

Then there exist $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ such that

$$f_0 = \sum_{i=1}^{m} \lambda_i f_i.$$

• We consider the problem

$$\min_{|u|=1} (Au, u)$$

for a symmetric matrix $A \in M_n(\mathbb{R})$. We may apply Theorem 3.2 with $G = \mathbb{R}^n \to \mathbb{R}$, $G(u) = |u|^2 - 1$, and $F : \mathbb{R}^n \to \mathbb{R}$, $F(u) = (Au, u)$, and we get $G'(u) = 2u$, $F'(u) = 2Au$. Thus any solution u_1 to the above constrained problem satisfies

$$Au_1 = \lambda u_1, \quad |u_1| = 1,$$

so that it is a solution to the first eigenvalue problem.
One can either proceed directly. By compacteness, there exists $u_1 \in S$ such that

$$\lambda_1 := (Au_1, u_1) \le (Au, u), \quad \forall\, u \in S.$$

For $w \in \mathbb{R}^n$ and $t > 0$ small enough, we have $u := (u_1 + tw)/|u_1 + tw| \in S$, and thus

$$\lambda_1 |u_1 + tw|^2 \le (A(u_1 + tw), u_1 + tw), \quad \forall\, t.$$

Expanding, we get

$$\lambda_1 (1 + 2t u_1 \cdot w + t^2 |w|^2) \le (Au_1, u_1) + 2t (Au_1, w) + t^2 (Aw, w), \quad \forall\, t,$$

and thus

$$\lambda_1 (u_1 \cdot w + \frac{t}{2} |w|^2) \le (Au_1, w) + \frac{t}{2} (Aw, w), \quad \forall\, t,$$

Passing to the limit $t \to 0$, we get

$$\lambda_1 (u_1, w) \le (Au_1, w), \quad \forall\, w \in \mathbb{R}^n,$$

and finally

$$Au_1 = \lambda_1 u_1.$$

• The same can be done for the minimization problem

$$\min_{u \in H_0^1,\, \|u\|_{L^2} = 1} \int_\Omega |\nabla u|^2.$$

Proceeding in the same way, we may define $F(u) := \| \nabla u\|_{L^2}^2$, $G(u) := \| u\|_{L^2}^2 - 1$, and the Lagrange multiplier Theorem 3.2 implies that any solution u_1 satisfies

$$\lambda_1 \int_\Omega u_1 w = \int \nabla u \cdot \nabla w, \quad \forall\, w \in H_0^1(\Omega).$$

That is a weak formulation of the first eigenvalue problem

$$u_1 \in H_0^1(\Omega), \quad -\Delta u_1 = \lambda_1 u_1.$$

• We now consider the minimization problem

$$\min_{f \in K} \mathcal{H}(f),$$

where $\mathcal{H}$ is the Boltzmann entropy

$$\mathcal{H}(f) := \int_{\mathbb{R}^d} f \log f \, dv$$

and K is the set of moments constraints

$$K := \{ \ f \geq 0, \ \mathcal{M}(f) := \int_{\mathbb{R}^d} f(1, v, |v|^2) dv = (1, 0, d) \}.$$

Without the first positivity constraint, the Lagrange multiplier Theorem 3.2 implies that any solution f_0 must satiisfy

$$\mathcal{H}'(f_0) = \Lambda \cdot \mathcal{M}'(f_0), \quad \Lambda \in \mathbb{R}^{d+2}.$$

In other words, we have

$$\int_{\mathbb{R}^d} [(1 + \log f_0) - \Lambda \cdot (1, v, |v|^2)] g \, dv = 0$$

for any g. That implies

$$1 + \log f_0 = \Lambda \cdot (1, v, |v|^2)$$

and thus

$$f_0 = e^{\lambda_0 - 1 + \lambda \cdot v - \kappa |v|^2}$$

with $\lambda_0, \kappa \in \mathbb{R}$, $\lambda \in \mathbb{R}^d$. Since $f_0 \in K$, we conclude that

$$f_0 = (2\pi)^{-d/2} e^{-|v|^2/2}.$$

4. References

- Chapter 1 of the book "Introduction à la théorie des points critiques: et applications aux problèmes elliptiques" (SMAI, 1993) by Otared Kavian

- Chapter 2 & Appendix C of the book "PDEs" (AMS, 1997) by Lawrence C. Evans