Differential Calculus

Course at Paris VI University, 2010/2011 (v.3)

Pierre Schapira
http://www.math.jussieu.fr/˜schapira/lectnotes

Contents

Introduction

The aim of these Notes is to provide a short and self-contained presentation of the main concepts of differential calculus.

Our point of view is to work in the abstract setting of a real normed space, and when necessary to specialize to the case of a finite dimensional space endowed with a basis. The main notion is that of the differential of a function: it is the first linear approximation of the function, and it is important since, in practice, linear functions are essentially the only ones which can be computed. Hence, the game is to deduce some knowledge of the function from that of its differential. We illustrate this point all over the course and particularly in two situations:
(i) when searching extrema of a real function, by using the method of Lagrange multipliers,
(ii) when studying embedded submanifolds and calculating their tangent spaces.

These Notes use the basic notions of general topology and are the continuation of [4]. Among the vaste literature dealing with differential calculus, let us only quote the books [1] and [3].

For the French students who would learn Mathematical English, see [2].

Bibliography

[1] J-P. Marco (sous la direction de) *Mathématiques L3 Analyse.* Pearson Education (2009)

[2] J. Nekovar, *Introduction to Mathematical English.*
`http://people.math.jussieu.fr/ nekovar/co/en/`

[3] J. Saint Raymond, *Topologie, Calcul Différentiel et variable complexe.* Calvage et Mounet (2007)

[4] P. Schapira, *General Topology.*
`http://www.math.jussieu.fr/ schapira/lectnotes/`

Chapter 1

Differentiable maps

1.1 Differential

In all these Notes, unless otherwise specified, we will consider *real* normed vector spaces. Recall that if E and F are two such spaces, $L(E,F)$ denotes the linear space of continuous linear maps from E to F. The space $L(E,F)$ is endowed with the norm $||u|| = \sup_{||x||\leq 1} ||u(x)||$. Here, $u \in L(E,F)$, $x \in E$ and $||\cdot||$ denotes either the norm on E, or on F or on $L(E,F)$.

In this Chapter, we will often encounter the following situation: U is an open subset of E and $f\colon U \to F$ is a map. We write for short

(1.1) $$E \supset U \xrightarrow{f} F \text{ or also } f\colon E \supset U \to F.$$

Definition 1.1.1. One says that f is differentiable at $a \in U$ if there exists $u \in L(E,F)$ such that

(1.2) $$f(a+h) - f(a) - u(h) = ||h|| \cdot \varepsilon(a,h)$$

where $\varepsilon(a,h)$ goes to 0 when h goes to 0.

Lemma 1.1.2. *If u exists, then it is unique.*

Proof. Assume we have u_1 and u_2 which satisfy (1.2). Set $v = u_1 - u_2$. Then $v(h) = ||h||\varepsilon(a,h)$, where $\varepsilon(a,h)$ goes to 0 when h goes to 0. Let us fix h and choose a sequence $(\lambda_n)_n$ in $\mathbb{R}_{>0}$ which goes to 0 when n goes to ∞. We have

$$\begin{aligned} v(h) &= \frac{1}{\lambda_n} v(\lambda_n h) = \frac{1}{\lambda_n} ||\lambda_n h|| \varepsilon(a, \lambda_n h) \\ &= ||h|| \varepsilon(a, \lambda_n h) \end{aligned}$$

which goes to 0 when n goes to ∞. Since $v(h)$ does not depend on n, $v(h) = 0$.
q.e.d.

Definition 1.1.3. (i) If u exists, one denotes it by $f'(a)$, or else $df(a)$, and calls it the differential of f at a.

(ii) If $f'(a)$ exists for all $a \in U$ one says that f is differentiable on U. If moreover, the map $U \ni a \mapsto f'(a) \in L(E,F)$ is continuous, then one says that f is of class C^1 on U and one denotes by $C^1(U;F)$ the space of such functions.

- If f is differentiable at a then f is continuous at a.
- If f is constant, then f is differentiable at each $a \in E$ and $f'(a) = 0$.
- If $f \in L(E,F)$, then f is differentiable at each $a \in E$ and $f'(a) = f$.
- If f is differentiable at a and $\lambda \in \mathbb{R}$, then λf is differentiable at a and $(\lambda f)'(a) = \lambda f'(a)$.
- If f and g are differentiable at a, then $f + g$ is differentiable at a and $(f+g)'(a) = f'(a) + g'(a)$.
- If $F = F_1 \times F_2$ and $f = (f_1, f_2)$, then f is differentiable at a if and only if f_1 and f_2 are differentiable at a and in this case $f'(a) = (f_1'(a), f_2'(a))$. A similar remark holds when $F = F_1 \times \cdots \times F_m$.

Example 1.1.4. (i) If $E = F = \mathbb{R}$, then f is differentiable at $a \in \mathbb{R}$ if and only if it is derivable at a. For example, the map $x \mapsto |x|$ from $\mathbb{R}$ to $\mathbb{R}$ is not differentiable at 0.
(ii) Let K be a compact space and let $E = C^0(K;\mathbb{R})$ be the Banach space of real continuous functions on K. Recall that E is a Banach space for the sup-norm, $||f|| = \sup_{x\in K} |f(x)|$. Consider the map

$$f\colon E \to E, \quad f(\varphi) = \varphi^2,\ \varphi \in E.$$

Then f is differentiable at each $\varphi \in E$ and $f'(\varphi) \in L(E,E)$ is given by

$$f'(\varphi)(\psi) = 2\varphi \cdot \psi.$$

To check it, note that

$$(\varphi + \psi)^2 - \varphi^2 = 2\varphi \cdot \psi + \psi^2$$

and $\psi^2 = ||\psi|| \cdot \varepsilon((\varphi, \psi)$ with $\varepsilon(\varphi,\psi) = \frac{\psi}{||\psi||}\psi$ and $\varepsilon(\varphi,\psi)$ (which does not depend on φ is this case) goes to 0 when ψ goes to 0.
(iii) Let A be a Banach algebra and let $f\colon A \to A$ be the map $a \mapsto a^2$. Then $f'(a)(b) = ab + ba$. The proof goes as in (ii).

Proposition 1.1.5. *Let E, F, G be normed spaces, let U an open subset of E and V an open subset of F. Let $f\colon E \supset U \to V \subset F$ and $g\colon F \supset V \to G$ be two maps. Let $a \in U$ and set $b = f(a)$. Assume that $f'(a)$ exists and $g'(b)$ exists. Set $h := g \circ f$. Then $h'(a)$ exists and*

$$h'(a) = g'(b) \circ f'(a).$$

Proof. Set $y := f(a+x) - f(a)$. Then $f(a+x) = b + y$ and

$$\begin{aligned} y = f(a+x) - f(a) &= f'(a)x + ||x||\alpha(a,x), \\ g(b+y) - g(b) &= g'(b)y + ||y||\beta(b,y), \\ h(a+x) - h(a) &= g(f(a+x)) - g(f(a)) = g(b+y) - g(b) \\ &= g'(b)\big(f'(a)x + ||x||\alpha(a,x)\big) + ||\big(f'(a)x + ||x||\alpha(a,x)\big)|| \cdot \beta(b,y) \\ &= (g'(b) \circ f'(a))x + ||x||\gamma(a,x) \end{aligned}$$

with

$$\gamma(a,x) = g'(b)(\alpha(a,x)) + ||\big(f'(a)\frac{x}{||x||} + \alpha(a,x)\big)|| \cdot \beta(b,y).$$

Here, $\alpha(a,x)$ goes to 0 when x goes to 0 and $\beta(b,y)$ goes to 0 when y goes to 0. Since f is continuous at a, y goes to 0 when x goes to 0. Hence, $\beta(b,y)$ goes to 0 when x goes to 0 and thus $\gamma(a,x)$ goes to 0 when x goes to 0. q.e.d.

Example 1.1.6. Let $E = C^0(K;\mathbb{R})$ for a compact space K. Let $U = \{\varphi \in E; \varphi > 0\}$. The set U is open in E. Indeed, let $\varphi \in U$. Let ε denote the minimum of φ on K. Since K is compact, there exists $t \in K$ with $\varphi(t) = \varepsilon$. Thus $\varepsilon > 0$. It follows that the open ball $B(\varphi, \varepsilon)$ is contained in U. Now take $K = [0,1]$ and consider the map

$$f\colon U \to \mathbb{R}, \quad \varphi \mapsto \int_0^1 \ln(\varphi(t))dt.$$

Then $f = g \circ h$ with $h(\varphi) = \ln(\varphi)$ and $g(\psi) = \int_0^1 \psi(t)dt$. Let us calculate h'. Let $[a,b]$ be a compact interval contained in $\mathbb{R}_{>0}$. One has

$$\ln(x+y) = \ln(x) + \frac{y}{x} + |y|\varepsilon(x,y)$$

with $\varepsilon(x,y)$ goes to 0 when $x \in [a,b]$ and y goes to 0. Hence

$$\ln((\varphi+\psi)(t)) = \ln(\varphi(t)) + \frac{\psi(t)}{\varphi(t)} + |\psi(t)|\varepsilon(\varphi(t),\psi(t)). \tag{1.3}$$

However, this result is not precise enough since we need that $||\varepsilon(\varphi,\psi)||$ goes to 0 when $||\psi||$ goes to 0. Let us recall a well-known formula for real functions of one variable, a result which will be generalized in Corollary 1.3.6:

> Let I be an open interval of $\mathbb{R}$ and let $f\colon I\to\mathbb{R}$ be a function which admits a continuous derivative on I. Then $|f(x+y)-f(x)-f'(x)\cdot y|\leq|y|\sup_{0\leq|t|\leq 1}|f'(x+ty)-f'(x)|$.

Applying this result to the function $\ln(x)$, we find a constant $C\geq 0$ not depending on $t\in[0,1]$ such that $|\varepsilon(\varphi,\psi)(t)|\leq C\cdot|\psi(t)|$ where $\varepsilon(\varphi,\psi)$ is defined in (1.3). Hence

$$\ln(\varphi+\psi)=\ln(\varphi)+\frac{\psi}{\varphi}+||\psi||\alpha(\varphi,\psi)$$

where $\alpha(\varphi,\psi)\in E$ is the function $\alpha(\varphi,\psi)(t)=\dfrac{\psi(t)}{||\psi||}\varepsilon(\varphi(t),\psi(t))$. Then $\alpha(\varphi,\psi)$ goes to 0 when ψ goes to 0 and we get that h is differentiable at each $\varphi\in U$ and

$$h'(\varphi)(\psi)=\frac{\psi}{\varphi}.$$

Since g is linear, $g'(\varphi)=g$ and we get

$$f'(\varphi)(\psi)=\int_0^1\frac{\psi(t)}{\varphi(t)}dt.$$

Example 1.1.7. Let A be a Banach algebra. Denote by Ω the set of invertible elements. Then Ω is open in A. Indeed, if $x\in\Omega$, then $x+y=x(\mathbf{1}+x^{-1}y)$ and $\mathbf{1}+x^{-1}y$ is invertible as soon as $||y||\cdot||x^{-1}||<1$, hence as soon as $||y||<(||x^{-1}||)^{-1}$.

Now denote by $J\colon\Omega\to\Omega$ the map $x\mapsto x^{-1}$. Then J is differentiable at each $a\in\Omega$ and

$$J'(a)h=-a^{-1}ha^{-1}. \tag{1.4}$$

(i) Let us first prove that J is continuous, that is, $(a+h)^{-1}-a^{-1}$ goes to 0 when h goes to 0. When multiplying $(a+h)^{-1}-a^{-1}$ by a, and noticing that $a(a+h)^{-1}=(a^{-1})^{-1}(a+h)^{-1}=((a+h)a^{-1})^{-1}$, this is equivalent to saying that $(\mathbf{1}-u)^{-1}$ goes to $\mathbf{1}$ when u goes to 0. Since

$$||\,\mathbf{1}-(\mathbf{1}-u)^{-1}||=||\sum_{n\geq 1}u^n||\leq\sum_{n\geq 1}||u||^n,$$

this follows from the fact that $\sum_{n\geq 1}\varepsilon^n$ goes to 0 when ε goes to 0.
(ii) Let us prove that J is differentiable and calculate its differential. We have

$$\begin{aligned}(a+h)^{-1}-a^{-1} &= a^{-1}a(a+h)^{-1}-a^{-1}(a+h)(a+h)^{-1}\\ &= -a^{-1}h(a+h)^{-1}\\ &= -a^{-1}ha^{-1}-a^{-1}h\big((a+h)^{-1}-a^{-1}\big).\end{aligned}$$

Therefore, $(a+h)^{-1}-a^{-1}-(-a^{-1}ha^{-1}) = ||h||\varepsilon(a,h)$, with

$$\varepsilon(a,h) = -a^{-1}\frac{h}{||h||}\big((a+h)^{-1}-a^{-1}\big).$$

Hence, $\varepsilon(a,h)$ goes to 0 when h goes to 0, by (i).

As a particular case, take $A = C^0(K,\mathbb{R})$ as in Example 1.1.6. Then the differential of the map $\varphi \mapsto \varphi^{-1}$ is given at φ by the linear map $\psi \mapsto -\psi/\varphi^2$.

Partial differential

Let E, F, U, a be as above and let N be a vector subspace of E, endowed with the norm induced by that of E. The set $a+N \subset E$ is called an affine subspace of E. It is the translated by a of the vector space N. Consider the maps

$$\iota\colon N \hookrightarrow E, \quad x \mapsto x, \quad \iota_a = \iota + a, \quad x \mapsto x+a. \tag{1.5}$$

The map ι_a is called an affine map, it is the sum of the linear map ι and the constant map $x \mapsto a$.

Definition 1.1.8. Let $E \supset U \xrightarrow{f} F$ be a map and let $a \in U$. One says that f is differentiable along N at a if $f \circ \iota_a\colon \iota_a^{-1}(U) \to F$ is differentiable at $0 \in N$ and one denotes by $f'_N(a)$ this differential.

- If f is differentiable along N at a, then $f'_N(a) \in L(N,F)$.
- Applying Proposition 1.1.5, we find that if f is differentiable at a, then f is differentiable along N at a and $f'_N(a) = f'(a)\circ\iota$, where ι is the linear embedding $N \hookrightarrow E$. Hence, $f'_N(a)$ is the restriction of the linear map $f'(a)$ to the subspace N of E.
- As a particular case, assume that N has dimension 1. Let $v \in N$, $v \neq 0$. In this case, the differential of f along N at a is also called the differential along v at a and denoted $f'_v(a)$ or else $D_v f(a)$. Then

$$f'_v(a) = \frac{d}{dt}f(a+tv)(0).$$

In other words, the differential along v at a is the derivative at 0 of the function $\mathbb{R} \ni t \mapsto f(a+tv) \in F$.

Now assume that $E = E_1 \times E_2$. (The extension to the case where $E = E_1 \times \cdots \times E_n$ is left to the reader.)

A vector $h \in E$ will be written as $h = (h_1, h_2) \in E_1 \times E_2$ or else $h = h_1 + h_2$ with $h_i \in E_i$ $(i = 1, 2)$ when identifying E_i with a subspace of E. A linear map $u\colon E \to F$ may thus be written as

$$u((h_1, h_2)) = u_1(h_1) + u_2(h_2), \tag{1.6}$$

where u_i is the restriction of u to E_i $(i = 1, 2)$.

Let $a = (a_1, a_2)$ and assume that f is differentiable at a. Then

$$\begin{aligned} f(a_1 + h_1, a_2 + h_2) - f(a_1, a_2) - f'(a_1, a_2)(h_1, h_2) \\ = (||h_1|| + ||h_2||)\varepsilon(a_1, a_2, h_1, h_2) \end{aligned} \tag{1.7}$$

where $\varepsilon(a_1, a_2, h_1, h_2)$ goes to 0 when $||h_1|| + ||h_2||$ goes to 0. Using (1.6), we get that

$$f'(a)(h_1, h_2) = f'_{E_1}(a)h_1 + f'_{E_2}(a)h_2.$$

Sometimes one denotes by dx_i the projection $E_1 \times E_2 \to E_i$ $(i = 1, 2)$. This notation makes sense since dx_i is the differential of the linear map $x = (x_1, x_2) \mapsto x_i$ (see below). One also writes

$$\frac{\partial f}{\partial x_j} := f'_{E_j} \text{ or else } \partial_j f := f'_{E_j} \text{ or else } f'_{x_j} := f'_{E_j}. \tag{1.8}$$

Then we get

$$\begin{aligned} f'(a) &= \frac{\partial f}{\partial x_1}(a)dx_1 + \frac{\partial f}{\partial x_2}(a)dx_2 \\ &= f'_{x_1}(a)dx_1 + f'_{x_2}(a)dx_2. \end{aligned}$$

One shall be aware that the existence of partial derivatives does not imply the existence of the differential.

Example 1.1.9. Let $E = \mathbb{R} \times \mathbb{R}$ and consider the function $f\colon E \to \mathbb{R}$ given by

$$f(x, y) = \frac{xy}{x^2 + y^2} \text{ for } (x, y) \neq 0, \quad f(0, 0) = 0.$$

Then f admits partial derivatives along $\{0\} \times \mathbb{R}$ and along $\mathbb{R} \times \{0\}$ at each $(x, y) \in \mathbb{R}^2$ but f is not differentiable at $(0, 0)$ since it is not continuous at this point.

Jacobian matrix

Assume that $E = E_1 \times \cdots \times E_n$. Let $a \in U \subset E$, with U open, and let $f\colon U \to F$ be a map. Assume $f'(a)$ exists. This is a linear map from E to F. Recall the notation 1.8. Hence

$$\frac{\partial f}{\partial x_j}(a) \in L(E_j, F)$$

and for $h = (h_1, \ldots, h_n) \in E_1 \times \cdots \times E_n$, we have

$$f'(a) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(a) h_j.$$

When E is endowed with a basis $(e_1, \ldots, e_n)$, $\dfrac{\partial f}{\partial x_j}(a)$ is the derivative at $t = 0$ of the function $\mathbb{R} \ni t \mapsto f(a + t \cdot e_j)$. Moreover, denote by $(e_1^*, \ldots, e_n^*)$ the dual basis to $(e_1, \ldots, e_n)$. Then e_j^* is the linear map

$$x = \sum_{i=1}^{n} x_i e_i \mapsto x_j,$$

that is, the projection $x \mapsto x_j$. Hence, it is natural to set

$$dx_j = e_j^*.$$

Then the differential $df(a) = f'(a)$ may be written as

$$df(a) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(a)\, dx_j.$$

Assume furthermore that F is a product, $F = F_1 \times \cdots \times F_p$. In this case $f = (f_1, \ldots, f_p)$.

Definition 1.1.10. The Jacobian matrix of f at a, denoted $J_f(a)$, is the matrix of the linear map $f'(a)$.

Hence $J_f(a) = \{\dfrac{\partial f_i}{\partial x_j}(a)\}_{\{i=1,\ldots,p;j=1,\ldots,n\}}$ where $\dfrac{\partial f_i}{\partial x_j}(a)$ belongs to the i-th row and j-th column of the matrix:

$$f'(a) \quad = \quad \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \cdot \quad \cdot \quad \cdot & \frac{\partial f_1}{\partial x_n}(a) \\ \cdot & \cdot \quad \cdot \quad \cdot & \cdot \\ \cdot & \cdot \quad \cdot \quad \cdot & \cdot \\ \frac{\partial f_p}{\partial x_1}(a) & \cdot \quad \cdot \quad \cdot & \frac{\partial f_p}{\partial x_n}(a) \end{pmatrix} \tag{1.9}$$

Now consider the situation of Proposition 1.1.5 and assume that $E = E_1 \times \cdots \times E_n$, $F = F_1 \times \cdots \times F_p$ and $G = G_1 \times \cdots \times G_l$. Let U an open subset of E, V an open subset of F, let $f\colon E \supset U \to V \subset F$ and let $g\colon F \supset V \to G$. Let $a \in U$ and set $b = f(a)$. Assume that $f'(a)$ exists and $g'(b)$ exists. Set $h := g \circ f$. We know that $h'(a)$ exists and is equal to $g'(b) \circ f'(a)$. By translating this result in matrices notations, we get that the Jacobian matrix of h is the composition of the Jacobian matrix of g and that of f:

$$J_h(a) = J_g(b) \circ J_f(a).$$

$\mathbb{C}$-differentiability

Assume that E and F are both normed vector spaces over $\mathbb{C}$ (hence, they are in particular real vector spaces). Let us denote as usual by $L(E, F)$ the space of $\mathbb{R}$-linear maps from E to F and let us denote by $L_{\mathbb{C}}(E, F)$ the $\mathbb{C}$-vector subspace consisting of $\mathbb{C}$-linear maps.

Definition 1.1.11. Let $U \subset E$ be an open subset and let $a \in U$. Let $f\colon U \to F$ be a map. One says that f is $\mathbb{C}$-differentiable at $a \in U$ if $f'(a)$ exists and belongs to $L_{\mathbb{C}}(E, F)$.

Let us consider the particular case where $E = \mathbb{C}$ and $F = \mathbb{C}$. Let $U \subset \mathbb{C}$ be an open subset and let $f\colon U \to \mathbb{C}$. Assume that $f'(a)$ exists. Then f is $\mathbb{C}$-differentiable at $a \in U$ if $f'(a)$ is $\mathbb{C}$-linear. Writting $\mathbb{C} = \mathbb{R} \times \mathbb{R}$ with coordinates $z = (x, y)$, and denoting by f'_x (resp. f'_y) the partial derivative with respect to x (resp. y), we have

$$f'(a) = f'_x(a)dx + f'_y(a)dy$$

where dx is the projection $z \mapsto x$ and dy is the projection $z \mapsto y$.

Now recall that an $\mathbb{R}$-linear map from $\mathbb{C}$ to $\mathbb{C}$ may be written as $u = adx + bdy$ and u is $\mathbb{C}$-linear if and only if $a + ib = 0$. Therefore f is $\mathbb{C}$-differentiable at $a \in U$ if and only if its partial derivatives $f'_x(a)$ and $f'_y(a)$ satisfy $f'_x(a) + if'_y(a) = 0$. Writting $\partial_x f$ instead of f'_x and similarly with $\partial_y f$ we get the Cauchy-Riemann equation:

$$((\partial_x + i\partial_y)f)(a) = 0.$$

1.2 Necessary conditions for extremum

In this section, we consider real valued functions $f\colon E \supset U \to \mathbb{R}$. Let $a \in U$.

- One says that f has a maximum at a if $f(x) \leq f(a)$ for any $x \in U$.
- One says that the maximum is strict if moreover $f(x) < f(a)$ for any $x \in U$, $x \neq a$.
- One says that f has a local maximum at a if $f(x) \leq f(a)$ for any x in a neighborhood of a.
- One says that f has a strict local maximum at a if $f(x) < f(a)$ for any x in a neighborhood of a.
- Replacing $<$ with $>$, we get similar notions, replacing maximum with minimum.
- An extremum is either a maximum or a minimum, and similarly for a local/strict extremum.

Theorem 1.2.1. *Assume that $f'(a)$ exists and that f has a local extremum at a. Then $f'(a) = 0$.*

First proof. Let N be a line in E (that is, a vector subspace of dimension one) and set $N_a = a + N$. Then the function $f|_{N_a}$ has a local extremum at a. This function is a real function defined on an interval of the real line N_a. Hence, its derivative is 0. Therefore $f'(a)|_N = 0$ for all line $N \subset E$ which impies $f'(a) = 0$. q.e.d.

Second proof. Let us give a proof which does not reduce to the 1-dimensional case.

Assume for example that $f(a+x) - f(a) \geq 0$ in a neighborhood of $x = 0$. Using the definition of the differential, we get

$$f'(a) \cdot x + ||x|| \cdot \varepsilon(a, x) \geq 0$$

where $\varepsilon(a, x)$ goes to 0 when x goes to 0. Set $x = ty$ with $t \in \mathbb{R}, t > 0$. Then

$$f'(a) \cdot y + ||y||\varepsilon(a, ty) \geq 0$$

Making $t \to 0$, we get $f'(a) \cdot y \geq 0$. replacing y with $-y$, this implies $f'(a) = 0$. q.e.d.

Lagrange's multipliers

First, we need a lemma of linear algebra.

Lemma 1.2.2. *Let E be a vector space over $\mathbb{R}$, let $u_1, \ldots, u_d$ be d continuous linear maps from E to $\mathbb{R}$. Set*

$$N = \{x \in E; u_j(x) = 0 \text{ for all } j = 1, \ldots, d\}. \tag{1.10}$$

In other words, $N = \bigcap_{j=1}^{d} u_j^{-1}(0)$. Let $u\colon E \to \mathbb{R}$ be a continuous linear map and assume that $u|_N = 0$. Then there exist $\lambda_1, \ldots, \lambda_d$ in $\mathbb{R}$ such that $u = \sum_{j=1}^{d} \lambda_j u_j$.

Proof. We shall only give the proof when $\dim E < \infty$. In this case, we set $E^* = L(E, \mathbb{R})$. Let $L \subset E^*$ be the linear space generated by $u_1, \ldots, u_d$ and let $L^\perp$ denote its orthogonal:

$$L^\perp = \{x \in E; \langle u, x\rangle = 0 \text{ for all } u \in L\}.$$

Clearly, $x \in L^\perp$ if and only if $u_j(x) = 0$ for $j = 1, \ldots, d$. Hence, $N = L^\perp$. The hypothesis $u|_N = 0$ means that $u \in N^\perp$. Since $(L^\perp)^\perp = L$, we get that $u \in L$, that is, u belongs to the space generated by $(u_1, \ldots, u_d)$. q.e.d.

Corollary 1.2.3. *Let $f\colon E \supset U \to \mathbb{R}$ be a map and let $a \in U$. Let $N \subset E$ be the vector subspace given by the equations* (1.10) *and let $N_a = a + N$ be the affine space passing through a and parallel to N. Assume that $f'(a)$ exists and $f|_{N_a}$ has a local extremum at a on $N_a \cap U$. Then there exist $\lambda_1, \ldots, \lambda_d$ in $\mathbb{R}$* (*the Lagrange multipliers*) *such that $f'(a) = \sum_{j=1}^{d} \lambda_j u_j$.*

Proof. Let $\iota_a\colon N \to E$ be the map $x \mapsto a+x$. Then the map $f \circ \iota_a$ has a local extremum at 0 on N. Applying Theorem 1.2.1, we get that $(f \circ \iota_a)'(0) = 0$. Hence, $f'(a)|_N = 0$ and it remains to apply Lemma 1.2.2. q.e.d.

Theorem 1.2.4 below will be easily deduced from Corollary 1.2.3 in Chapter 3.

Theorem 1.2.4. (The Lagrange's multipliers Theorem.) *Let E be a finite dimensional vector space, U an open subset and let $a \in U$. Let $f\colon E \supset U \to \mathbb{R}$ be a map and assume that $f'(a)$ exists. Also assume that there exist d functions $\{\varphi_j\}_{j=1,\ldots,d}$ of class C^1 on U such that their differentials at a, $\{\varphi_j'(a)\}_{j=1,\ldots,d}$, are linearly independent. Set $S = \{x \in U; \varphi_j(x) = 0 \text{ for all } j = 1, \ldots, d\}$. Now assume that $f|_S$ has a local extremum at a. Then there exist $\lambda_1, \ldots, \lambda_d$ in $\mathbb{R}$* (*the Lagrange multipliers*) *such that $f'(a) = \sum_{j=1}^{d} \lambda_j \varphi_j'(a)$.*

Example 1.2.5. Let $E = C^0([0,1];\mathbb{R})$, $N = \{\varphi \in E; \int_0^1 \varphi(t)dt = 1\}$ and $U = E$. Consider the function

$$f\colon E \to \mathbb{R}, \quad \varphi \mapsto \int_0^1 \varphi^2(t)dt.$$

Let us find the extrema of f on N. Denote by $u\colon E \to \mathbb{R}$ the linear map $u\colon \varphi \mapsto \int_0^1 \varphi(t)dt$. Then $N = \{\varphi \in E; u(\varphi) = 1\}$. On the other hand, $f = h \circ g$ where $g(\varphi) = \varphi^2$ and $h(\psi) = \int_0^1 \psi(t)dt$. Since $g'(\varphi) \cdot \psi = 2\varphi\psi$ and h is linear, we get

$$f'(\varphi)\psi = 2\int_0^1 \varphi(t)\psi(t)dt.$$

If f has an extremum at φ on N, then there exists $\lambda \in \mathbb{R}$ such that $f'(\varphi) = \lambda u'(\varphi) = \lambda u$. Hence

$$f'(\varphi)\psi = \lambda \int_0^1 \psi(t)dt$$

for all $\psi \in E$. Choosing $\psi = 2\varphi - \lambda$, we get $\int_0^1 (2\varphi - \lambda)^2 dt = 0$, hence $\varphi = \frac{1}{2}\lambda$. Since $\varphi \in N$, this implies $\varphi \equiv 1$. On the other hand, one has the Cauchy-Schwarz's inequality

$$(\int_0^1 \varphi(t)\psi(t)dt)^2 \leq \int_0^1 \varphi^2(t)dt \int_0^1 \psi^2(t)dt.$$

Choosing $\varphi \equiv 1$, we find for $\psi \in N$:

$$1 \leq \int_0^1 \psi^2(t)dt,$$

that is, $1 \leq f(\psi)$. Therefore, $\varphi \equiv 1$ is a global minimum.

Example 1.2.6. Let $E = C^0([0,1];\mathbb{R})$, $N = \{\varphi \in E; \int_0^1 \varphi(t)dt = 1\}$ and $U = \{\varphi \in E; \varphi > 0$. Recall that U is open in E. Indeed, if $\varphi \in U$, there exists $c > 0$ such that $\varphi(t) \geq c$ for all $t \in [0,1]$ and $||\psi|| < c$ implies that $\varphi + \psi$ still belongs to U.

Consider the function

$$f\colon E \to \mathbb{R}, \quad \varphi \mapsto \int_0^1 \ln(\varphi(t))dt.$$

Then

$$f'(\varphi)\psi = \int_0^1 \frac{\psi(t)}{\varphi(t)} dt.$$

Therefore, if f has a local extremum at φ, we get

$$\int_0^1 \frac{\psi(t)}{\varphi(t)} dt = \lambda \int_0^1 \psi(t)dt$$

for all $\psi \in E$. Choose $\psi = \dfrac{1}{\varphi} - \lambda$. One gets

$$\int_0^1 (\frac{1}{\varphi} - \lambda)^2 dt = 0,$$

hence $\varphi = \lambda^{-1}$, and since $\varphi \in N$, $\varphi \equiv 1$.

Let us show that $\varphi \equiv 1$ is a strict local maximum. Set $\psi = 1 + \theta$. Then $\psi \in N$ if and only if $\int_0^1 \theta(t)dt = 0$. We have

$$\ln(1+\theta) = \theta - \frac{\theta^2}{2} + \theta^2 \varepsilon(\theta),$$

where $||\varepsilon(\theta)||$ goes to 0 when $||\theta||$ goes to 0. In particular, there exists $\eta > 0$ such that $||\theta|| \leq \eta$ implies $\varepsilon(\theta) \leq \dfrac{1}{4}$. Then

$$\begin{aligned} f(1+\theta) &= \int_0^1 \ln(1+\theta)dt \\ &= \int_0^1 \theta(t)dt - \frac{1}{2}\int_0^1 \theta(t)^2(1 - 2\varepsilon(\theta))dt. \end{aligned}$$

For $\theta \neq 0$, $||\theta|| \leq \eta$, the function $\theta(t)^2(1-2\varepsilon(\theta))$ is positive and not identically 0. Since $\int_0^1 \theta(t)dt = 0$, we get $f(1+\theta) < 0$.

Example 1.2.7. In this example, we shall prove the inequality

$$(1.11) \qquad (x_1 \cdots x_n)^{\frac{1}{n}} \leq \frac{(x_1 + \cdots + x_n)}{n} \text{ for all } (x_1, \dots, x_n) \in (\mathbb{R}_{>0})^n.$$

Note first that (1.11) is satisfied at $(x_1, \dots, x_n)$ if and only if it is satisfied at $(\lambda x_1, \dots, \lambda x_n)$ for some $\lambda \in \mathbb{R}_{>0}$. Hence, we are reduce to prove that

$$(1.12) (x_1 \cdots x_n)^{\frac{1}{n}} \leq \frac{1}{n} \text{ for all } (x_1, \dots, x_n) \in (\mathbb{R}_{>0})^n \text{ with } x_1 + \cdots + x_n = 1.$$

Denote by K the compact set

$$K = \{x \in \mathbb{R}^n; x_i \geq 0 \text{ for all } i, \sum_{i=1}^{n} x_i = 1\}.$$

Consider the function $f\colon K \to \mathbb{R}$, $f(x) = (x_1 \cdots x_n)^{\frac{1}{n}}$. This function being continuous, it takes its maximum at some point $a \in K$. Since $f \neq 0$ and $f \equiv 0$ on the boundary ∂K, we have $a \in \mathrm{Int(K)}$, that is, $a \in (\mathbb{R}_{>0})^n$. Set

$$U = (\mathbb{R}_{>0})^n \subset E = \mathbb{R}^n, N = \{x \in \mathbb{R}^n; \sum_{i=1}^{n} x_i = 1\}.$$

Applying Corollary 1.2.3, there exists $\lambda \in \mathbb{R}$ such that $f'(a) = \lambda(\sum_{i=1}^{n} dx_i)$. Hence

$$\sum_{i=1}^{n} (a_1 \cdots a_n)^{\frac{1}{n}} \frac{dx_i}{a_i} = \lambda(\sum_{i=1}^{n} dx_i).$$

This implies $a_1 = \cdots = a_n$ and since $\sum_{i=1}^{n} a_i = 1$, we get $a_i = \frac{1}{n}$ for all i. At this point a, we have $f(a) = ((\frac{1}{n})^n)^{\frac{1}{n}} = \frac{1}{n}$, which proves (1.12).

Example 1.2.8. Consider the plane curve Γ defined by the equation

$$\Gamma = \{(x, y) \in \mathbb{R}^2; \varphi(x, y) = 0\}, \text{ where } \varphi(x, y) = 3x^2 + 2xy + 2y^2 - 1.$$

First remark that since $|\varphi(x, y)| \geq x^2 + y^2$, the function $|\varphi(x, y)|$ goes to infinity when (x, y) goes to infinity. This implies that Γ is bounded, hence compact since φ is continuous.

We search the Euclidian distance of 0 to Γ. Hence, the problem is to find the minimum of the function $\sqrt{x^2 + y^2}$ on the compact set Γ.

Since $d\varphi \neq 0$ on Γ, we may apply Theorem 1.2.4 to the function $f(x, y) = x^2 + y^2$. If this function admits a minimum on Γ at (x, y), we will have

$$2(xdx + ydy) = \lambda((6x + 2y)dx + (4y + 2x)dy.$$

Hence

$$\begin{vmatrix} x & y \\ 6x + 2y & 4y + 2x \end{vmatrix} = 2x^2 + 4xy - 2y^2 - 6xy = 0.$$

Since $(x, y) \in \Gamma$, we get the system

$$3x^2 + 2xy + 2y^2 - 1 = 0, \quad x^2 - y^2 - xy = 0,$$

which gives $5x^2 = 1$, hence $x = \dfrac{\pm 1}{\sqrt{5}}$. Then one calculates y and among the four solutions for (x, y) one chooses that which minimizes $x^2 + y^2$.

Example 1.2.9. let us calculate the Euclidian distance in $\mathbb{R}^2$ between a line $\Delta = \{(x,y); \alpha x + \beta y + \gamma = 0\}$ and a parabola $\Gamma = \{(x,y); y - \delta x^2 = 0\}$. A first method would be to calculate $d((x,y),\Gamma)$ and then to mimimize this function on Δ. But a simpler method is the following. Denote by $a = (x,y)$ a point of Γ and $a' = (x',y')$ a point of Δ. The distance $d(a,a')$ is $\sqrt{(x-x')^2 + (y-y')^2}$. Hence, it is enough to search the minimum of the function $(x-x')^2 + (y-y')^2$ on the set $\Gamma \times \Delta \subset \mathbb{R}^4$. We set

$$\begin{aligned}
&f(x,y,x',y') = (x-x')^2 + (y-y')^2,\\
&\varphi_1(x,y,x',y') = -\delta x^2 + y,\\
&\varphi_2(x,y,x',y') = \alpha x' + \beta y' + \gamma,\\
&\Gamma \times \Delta := S = \{(x,y,x',y') \in \mathbb{R}^4; \varphi_1(x,y,x',y') = \varphi_2(x,y,x',y') = 0\}.
\end{aligned}$$

If $f|_S$ has an extremum at some point A, there exist λ_1 and λ_2 such that

$$f'(a) = \lambda_1 \varphi_1'(a) + \lambda_2 \varphi_2'(a).$$

In other words, the system of 3 vectors in $\mathbb{R}^4$, $f'(a), \varphi_1'(a), \varphi_2'(a)$ has rank 2, that is, the matrix

$$\begin{pmatrix} -2\delta x & 1 & 0 & 0 \\ 0 & 0 & \alpha & \beta \\ x - x' & y - y' & x' - x & y' - y \end{pmatrix} \text{ has rank 2.}$$

We get

$$\alpha(y-y') - \beta(x-x') = 0, \quad 2\delta x(y'-y) + (x'-x) = 0.$$

The first equation tells us that the line passing through (x,y) and (x',y') is orthogonal to Δ and the second equation tells us that this line is orthogonal to the tangent to Γ at (x,y). Assuming for simplicity that $\beta \neq 0$, we find $x = \frac{-\alpha}{2\delta\beta}$.

Let us be more explicit and assume $\Delta = \{y - x + 1 = 0\}$, $\Gamma = \{y - x^2 = 0\}$. Then $x = \frac{1}{2}$, $y = \frac{1}{4}$ and $x' = \frac{7}{8}$, $y' = \frac{-1}{8}$. Therefore, the value of the distance is $d = ((x'-x)^2 + (y'-y)^2)^{\frac{1}{2}} = \frac{3}{8}\sqrt{2}$.

1.3 The mean value theorem

In this section, we consider a normed space F, an interval $I = [a,b]$ in $\mathbb{R}$ with $-\infty < a < b < +\infty$, a continuous function $f\colon I \to F$ and a continuous function $g\colon I \to \mathbb{R}$.

[1] "Le théorème des accroissements finis" in French

Lemma 1.3.1. *Assume that the right derivatives $f'_d(x)$ and $g'_d(x)$ exist for all $x \in]a,b[$ and that $||f'_d(x)|| \leq g'_d(x)$ for all $x \in]a,b[$. Then $||f(b) - f(a)|| \leq g(b) - g(a)$.*

Proof. Let $\varepsilon > 0$ and consider the relation

$$(*)_\varepsilon \qquad ||f(x) - f(a)|| \leq g(x) - g(a) + \varepsilon(x - a) + \varepsilon.$$

Let Z be the set of $x \in [a,b]$ where $(*)_\varepsilon$ is satisfied at x. Since f and g are continuous, Z is closed. Let $U = [a,b] \setminus Z$, that is, U is the set of $x \in [a,b]$ where $(*)_\varepsilon$ is not satisfied at x. We shall show that $U = \emptyset$. Assume $U \neq \emptyset$ and denote by c its lower bound. If $c \in U$ then $c = a$ (since U is open). But $a \notin U$. Hence $c \notin U$. On the other hand, $c \neq b$, otherwise $U = \emptyset$. Hence, $c \in]a,b[$. Therefore, $||f'_d(c)|| \leq g'_d(c)$. Hence, there exists $\eta > 0$ such that for all $x \in [c, c+\eta]$:

$$\begin{aligned} ||\frac{(f(x) - f(c))}{(x - c)}|| - \frac{\varepsilon}{2} &\leq ||f'_d(c)||, \\ g'_d(c) &\leq \frac{(g(x) - g(c))}{(x - c)} + \frac{\varepsilon}{2}. \end{aligned}$$

Using the hypothesis that $||f'_d(c)|| \leq g'_d(c)$ we get

$$||f(x) - f(c)|| \leq g(x) - g(c) + \varepsilon(x - c). \tag{1.13}$$

Since $c \notin U$, we have

$$||f(c) - f(a)|| \leq g(c) - g(a) + \varepsilon(c - a) + \varepsilon. \tag{1.14}$$

By adding (1.13) and (1.14), we get

$$||f(x) - f(a)|| \leq g(x) - g(a) + \varepsilon(x - a) + \varepsilon.$$

Hence, $x \notin U$, which contredicts the fact that c is the lower bound of U. Therefore, U is empty. Since $(*)_\varepsilon$ holds for any $\varepsilon > 0$, it holds for $\varepsilon = 0$. q.e.d.

Remark 1.3.2. Contrarily to the case $F = \mathbb{R}$, if $\dim F > 1$ then, in general, there exists no $c \in [a,b]$ such that $f(b) - f(a) = (b-a)f'(c)$. For example, consider the map $f\colon \mathbb{R} \to \mathbb{R}^2$ given by $f(t) = (\cos(t), \sin(t))$.

Now let E be a real normed space and let $U \subset E$ be an open subset. Recall that a subset $A \subset U$ is convex if for any $x, y \in A$, the interval $[x,y] = \{tx + (1-t)y\}_{t\in[0,1]}$ is contained in A.

Theorem 1.3.3. *Let $f\colon U \to F$ be a map and assume that f is differentiable on U. Let $A \subset U$ be a convex subset and assume that $||f'|| \leq M$ on A for some $M \geq 0$. Then $||f(x) - f(y)|| \leq M||x - y||$ for all $x, y \in A$.*

Proof. Set $h(t) = f((1-t)x + ty)$. Then $h'(t) = f'((1-t)x + ty) \cdot (y - x)$. Hence, $||h'(t)|| \leq M \cdot ||x - y||$. Therefore $||h(1) - h(0)|| \leq M \cdot ||x - y||$ by Lemma 1.3.1. q.e.d.

Corollary 1.3.4. *Let $f\colon U \to F$ be a map. Assume that f is differentiable on U and U is connected. If $f' \equiv 0$ then f is constant. If $f' \equiv u$ with $u \in L(E, F)$ then there exists $b \in F$ such that $f(x) = u(x) + b$ for all $x \in U$ (f is affine).*

Proof. (i) Assume that $f' \equiv 0$. Since U is connected, it is enough to prove that f is locally constant. For each $a \in U$ there is a non empty open ball $B(a, \varepsilon) \subset U$ and such a ball is convex. Then f is constant on $B(a, \varepsilon)$ by Theorem 1.3.3.
(ii) If $f' = u$, set $g = f - u$. Then $g' = 0$ hence g is constant q.e.d.

Corollary 1.3.5. *Let $f\colon U \to F$ be a map and assume that f is differentiable on U. Let $a \in U$ and let $\varepsilon > 0$ such that the closed ball $\overline{B}(a, \varepsilon)$ is contained in U. Then for any $h \in E$ with $||h|| \leq \varepsilon$, one has*

$$||f(a+h) - f(a)|| \leq ||h|| \cdot \sup_{0 \leq t \leq 1} ||f'(a+th)||.$$

Proof. Set $M = \sup_{0 \leq t \leq 1} ||f'(a+th)||$ and apply Theorem 1.3.3 with $x = a$, $y = a + h$ and $A = [a, a+h]$. q.e.d.

Corollary 1.3.6. *Let $f\colon U \to F$ be a map and assume that f is differentiable on U. Let $a \in U$ and let $\varepsilon > 0$ such that the closed ball $\overline{B}(a, \varepsilon)$ is contained in U. Then for any $h \in E$ with $||h|| \leq \varepsilon$, one has*

$$||f(a+h) - f(a) - f'(a) \cdot h|| \leq ||h|| \cdot \sup_{0 \leq t \leq 1} ||f'(a+th) - f'(a)||.$$

Proof. Set $g(x) = f(x) - f'(a) \cdot x$. Then

$$\begin{aligned} &g(a+h) - g(a) = f(a+h) - f(a) - f'(a) \cdot h, \\ &g'(a+th) = f'(a+th) - f'(a). \end{aligned}$$

Applying Corollary 1.3.5, we get

$$||g(a+h) - g(a)|| \leq ||h|| \cdot \sup_{0 \leq t \leq 1} ||g'(a+th)||.$$

q.e.d.

Remark 1.3.7. It is possible to generalize Theorem 1.3.3 by weakening the hypothesis that A is convex with the hypothesis that A is connected. Assuming that $||f'|| \leq M$ on A for some $M \geq 0$, we get

$$||f(x) - f(y)|| \leq M d_A(x, y) \text{ for all } x, y \in A$$

where $d_A(x, y)$ is the infimum of all lengths of polygonial lines contained in A and going from x to y.

Application to partial derivatives

Assume that $E = E_1 \times E_2$ (the reader will extend the results to the case of a finite product). Let $U \subset E$ be an open subset as above. Recall Definition 1.1.3.

We know by Example 1.1.9 that there exist functions which admit partial derivatives at each $a \in U$ but which are not continuous. However, such a phenomena cannot happen if the partial derivatives are continuous functions of $a \in U$.

Theorem 1.3.8. *Let $f\colon U \to F$ be a map. Then f is in $C^1(U, F)$ if and only if the partial derivatives $f'_{x_i}(a)$ $(i = 1, 2)$ exist on U and are continuous, that is, the maps $U \to L(E_i, F)$, $a \mapsto f'_{x_i}(a)$ $(i = 1, 2)$ are continuous.*

Proof. (i) Denote by $u_i\colon E_i \to E$ the linear inclusion. Since $f'_{x_i}(a) = f'(a) \circ u_i$, the condition is clearly necessary.
(ii) Assume that the maps $a \mapsto f'_{x_i}(a)$ $(i = 1, 2)$ are continuous and let us prove that $f'(a)$ exists. For $a = (a_1, a_2) \in U$ we have

$$\begin{aligned}
&||f(a_1 + h_1, a_2 + h_2) - f(a_1, a_2 + h_2) - f'_{x_1}(a_1, a_2 + h_2) \cdot h_1|| \\
&\qquad \leq ||h_1|| \cdot \sup_{0 \leq t \leq 1} ||f'_{x_1}(a_1 + th_1, a_2 + h_2) - f'_{x_1}(a_1, a_2 + h_2)||, \\
&||f'_{x_1}(a_1, a_2 + h_2) \cdot h_1 - f'_{x_1}(a_1, a_2) \cdot h_1|| \\
&\qquad \leq ||h_1|| \cdot ||f'_{x_1}(a_1, a_2 + h_2) - f'_{x_1}(a_1, a_2)||, \\
&||f(a_1, a_2 + h_2) - f(a_1, a_2) - f'_{x_2}(a_1, a_2) \cdot h_2|| \\
&\qquad \leq ||h_2|| \cdot \sup_{0 \leq t \leq 1} ||f'_{x_2}(a_1, a_2 + th_2) - f'_{x_2}(a_1, a_2)||.
\end{aligned}$$

The first and third inequalities follow from Corollary 1.3.6 and the second one is obvious. By adding these inequalities we obtain

$$\begin{aligned}
&||f(a_1 + h_1, a_2 + h_2) - f(a_1, a_2) - \big(f'_{x_1}(a_1, a_2) \cdot h_1 + f'_{x_2}(a_1, a_2) \cdot h_2\big)|| \\
&\qquad \leq (||h_1|| + ||h_2||)(\varepsilon_1 + \varepsilon_2 + \varepsilon_3)
\end{aligned}$$

with

$$\begin{aligned}
\varepsilon_1 &= \sup_{0\le t\le 1} ||f'_{x_1}(a_1+th_1, a_2+h_2) - f'_{x_1}(a_1, a_2+h_2)|| \\
\varepsilon_2 &= ||f'_{x_1}(a_1, a_2+h_2) - f'_{x_1}(a_1, a_2)|| \\
\varepsilon_3 &= \sup_{0\le t\le 1} ||f'_{x_2}(a_1, a_2+th_2) - f'_{x_2}(a_1, a_2)||.
\end{aligned}$$

Therefore, $\varepsilon_1+\varepsilon_2+\varepsilon_3$ goes to 0 when $||h||$ goes to 0.
(iii) To conclude, let us prove that $a \mapsto f'(a)$ is continuous. By the hypothesis, the maps $a \mapsto f'_{x_i}(a)$ $(i = 1, 2)$ are continuous. Therefore the map $a \mapsto f'(a) = f'_{x_1}(a)\,dx_1 + f'_{x_2}(a)\,dx_2$ is continuous. (Recall that dx_i is the projection $E \to E_i$ $(i = 1, 2)$.) q.e.d.

Corollary 1.3.9. *Let U be an open subset of $E = E_1 \times E_2 \times \cdots \times E_n$ and let $F = F_1 \times F_2 \times \cdots \times F_p$. Then a map $f = (f_1, \ldots, f_p)\colon U \to F$ is C^1, that is, f is in $C^1(U, F)$ if and only if the maps $\frac{\partial f_i}{\partial x_j}$ are C^1.*

Sequences of differentiable functions

Let E and F be Banach spaces, U an open subset of E.

Theorem 1.3.10. *Let $(f_n)_n$ be a sequence of functions $f_n\colon U \to F$. Assume:*

(a) *U is connected and there exists $a \in U$ such that the sequence $(f_n(a))_n$ has a limit $b \in F$,*

(b) *the sequence $(f'_n)_n$ converges uniformly on U to a function $g\colon U \to L(E, F)$.*

Then

(i) *for any $x \in U$, the sequence $(f_n(x))_n$ has a limit $f(x)$,*

(ii) *the sequence $(f_n)_n$ converges to f uniformly on any ball of E contained in U,*

(iii) *the function $f\colon x \mapsto f(x)$ is differentiable and $f' = g$.*

Proof. (A) First, assume that U is an open ball of radius R and denote by B this ball. (We do not assume that B is centered at a.) Applying Theorem 1.3.3 to $f_n - f_p$ we find for $x, y \in B$:

$$||f_n(x) - f_p(x) - (f_n(y) - f_p(y))|| \le ||x-y|| \cdot \sup_{z\in B} ||f'_n(z) - f'_p(z)||. \tag{1.15}$$

Hence, choosing $y = a$

$$(1.16)\ ||f_n(x) - f_p(x)|| \leq ||f_n(a) - f_p(a)|| + ||x - a|| \cdot \sup_{z \in B} ||f'_n(z) - f'_p(z)||.$$

By the hypotheses, for any $\varepsilon > 0$ there exists N such that for $n, p \geq N$, we have

$$||f_n(a) - f_p(a)|| \leq \frac{\varepsilon}{2}, \quad \sup_{z \in B} ||f'_n(z) - f'_p(z)|| \leq \frac{\varepsilon}{4R}.$$

Since $||x - a|| \leq 2R$, we find $||f_n(x) - f_p(x) \leq \varepsilon$ and therefore the sequence $(f_n)_n$ is a Cauchy sequence for the norm of uniform convergence on B. Since F is complete, it follows that the sequence $(f_n)_n$ converges uniformly on B to a continuous function f.

Let us show that $f'(x_0)$ exists and $f'(x_0) = g(x_0)$ for any $x_0 \in B$. We have

$$\begin{aligned} &||f(x) - f(x) - g(x_0) \cdot (x - x_0)|| \leq \\ &\quad ||f(x) - f(x_0) - (f_n(x) - f_n(x_0))|| \\ &\quad + ||(f_n(x) - f_n(x_0)) - f'_n(x_0) \cdot (x - x_0)|| \\ &\quad + ||(f'_n(x_0) - g(x_0)) \cdot (x - x_0)||. \end{aligned} \tag{1.17}$$

Let $\varepsilon > 0$.
(i) By making p going to ∞ in (1.15) (with $y = x_0$), we see that there exists $N_1 \geq 0$ such that $n \geq N_1$ implies:

$$||f(x) - f(x_0) - (f_n(x) - f_n(x_0))|| \leq \varepsilon ||x - x_0||. \tag{1.18}$$

(ii) Since $f'_n(x_0) \xrightarrow{n} g(x_0)$, there exists $N_2 \geq 0$ such that $n \geq N_2$ implies:

$$||(f'_n(x_0) - g(x_0)) \cdot (x - x_0)|| \leq \varepsilon ||x - x_0||. \tag{1.19}$$

(iii) Choose $n \geq \sup(N_1, N_2)$. There exists $\eta > 0$ such that $||x - x_0|| \leq \eta$ implies:

$$||(f_n(x) - f_n(x_0)) - f'_n(x_0) \cdot (x - x_0)|| \leq \varepsilon ||x - x_0||. \tag{1.20}$$

By adding (1.18), (1.19) and (1.20), we obtain that for for each $\varepsilon > 0$, there exists $\eta > 0$ such that $||x - x_0|| \leq \eta$ implies that the left hand side of (1.17) is less or equal to $3\varepsilon ||x - x_0||$.

(B) We have proved the theorem when U is an open ball. Assume that $U = B_1 \cup B_2$ where B_1 and B_2 are open balls and $B_1 \cap B_2 \neq \emptyset$. Assume for example that the point a given in the statement belongs to B_1. Then the result holds on B_1. Let $b \in B_1 \cap B_2$. Since $f_n(b) \xrightarrow{n} f(b)$, we get the result on B_2.

(C) By induction, we have proved the theorem when U is a finite union of open balls

$$U = \bigcup_{j=1}^{N} B_j, \text{ with } B_i \cap B_{i+1} \neq \emptyset \text{ for } j < N. \tag{1.21}$$

(D) To conclude, remark that U being connected, for any pair $a, b \in U$ there exists a finite set of open balls as in (1.21) whose union contains both a and b. To check this point, notice that U being connected and locally arcwise connected, it is arcwise connected. Hence, there exists a continuous map $\gamma\colon [0,1] \to U$ such that $\gamma(0) = a$ and $\gamma(1) = b$. The set $K = \gamma([0,1]) \subset U$ being compact, there exists a finite covering of K by open balls conained in U. q.e.d.

Exercises to Chapter 1

Exercise 1.1. We follow the notations of Example 1.1.7: A is a Banach algebra, Ω denotes the open set of invertible elements and $J\colon \Omega \to \Omega$ is the map $x \mapsto x^{-1}$. Without using the fact that $J \circ J = \mathrm{id}$, check directly that $J'(J(a)) \circ J'(a) = \mathrm{id}_A$ for any $a \in \Omega$.

Exercise 1.2. Consider the map $f = (f_1, f_2)\colon \mathbb{R}^2 \to \mathbb{R}^2$ given by:

$$f_1(x,y) = x, \qquad f_2(x,y) = \begin{cases} 0 \text{ if } (x,y) = (0,0) \\ y\dfrac{y^2 - x^4}{y^2 + x^4} \text{ otherwise.} \end{cases}$$

(i) Show that $2x^2|y| \leq x^4 + y^2$ then that $|y - f_2(x,y)| \leq x^2$.
(ii) Deduce that f_2 and f are differentiable at each $(x,y) \in \mathbb{R}^2$ and calculate the differential of f at $(0,0)$.

Exercise 1.3. Let A, B, C be 3 points on the unit circle of $\mathbb{R}^2$. How to choose these points in order that the perimeter of the triangle (A, B, C) is maximum?

Exercise 1.4. Let $\{b_i\}_{i=1}^n$ be positive numbers with $b_i \neq b_j$ for $i \neq j$ and let $f(x) = \sum_{i=1}^n b_i x_i^2$. find the extrema of f on the unit sphere $\mathbb{S}^{n-1} = \{x \in \mathbb{R}^n; \sum_{i=1}^n x_i^2 = 1\}$.

Exercise 1.5. Find the local extrema of the following functions:
(i) $f(x, y) = x^2 - xy + y^2 - 2x + y$,
(ii) $f(x, y) = x^2y^3(6 - x - y)$,
(iii) $f(x, y) = x^3 + y^3 - 3xy$,
(iv) $f(x, y) = \exp(2x + 3y)(8x^2 - 6xy + 3y^2)$,
(v) $f(x, y) = \sin(x)\sin(y)\sin(x + y)$ $(0 \le x, y \le \pi)$

Exercise 1.6. Find the local extrema of $u(x, y, z) = \sin(x)\sin(y)\sin(z)$ when $x + y + z = \frac{\pi}{2}$ and $x, y, z > 0$.

Exercise 1.7. Find the values of a and b in order that the integral

$$\int_0^1 (\frac{1}{x^2 + 1} - ax - b)^2 dx$$

be as small as possible.

Chapter 2

Higher differentials

As in Chapter 1, E and F will denote real normed vector spaces.

2.1 The Schwarz lemma

Multilinear linear maps

Let $E_1, \cdots, E_n$ and F be real normed spaces. Recall that a map

$$u\colon E_1 \times \cdots \times E_n \to F, \quad x = (x_1, \ldots, x_n) \mapsto u(x)$$

is multilinear if u is separately linear, that is, $u(a_1, \ldots, a_{i-1}, x_i, a_{i+1}, \ldots, a_n)$ is linear with respect to $x_i \in E_i$ for any $(a_1, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n)$ in $E_1 \times \cdots \times E_{i-1} \times E_{i+1} \times \cdots \times E_n$.

If $n = 2$, a multilinear map is called a bilinear map.

If F is the field $\mathbb{R}$ of $\mathbb{C}$ (here, the field is $\mathbb{R}$), one says "a multilinear form" instead of "a multilinear map".

Similarly as for linear maps, one proves that

- the multilinear map u is continuous if and only if it is continuous at 0,
- this last condition is also equivalent to the existence of a constant $M \geq 0$ such that

$$||u(x_1, \ldots, x_n)|| \leq M \cdot ||x_1|| \cdots ||x_n||.$$

- If all E_i's are finite dimensional, any multilinear map is continuous.

One endows the vector space of multilinear linear maps from $E_1 \times \cdots \times E_n$ to F with the norm

$$||u|| = \sup_{||x_i|| \leq 1, 1 \leq i \leq n} ||u(x_1, \ldots, x_n)||.$$

Therefore

$$||u(x_1,\dots,x_n)|| \leq ||u|| \cdot ||x_1|| \cdots ||x_n||.$$

If $E_i = E$ for all i, a multilinear map is called an n-linear map. For $n \in \mathbb{N}$, we shall denote by $L^n(E,F)$ the space of continuous n-linear maps from E to F.

Note that for $n = 1$, $L^1(E,F) = L(E,F)$. For $n = 0$ one sets $L^0(E,F) = F$. Also note that

$$L^n(E,F) \simeq L(E, L^{n-1}(E,F)). \tag{2.1}$$

In other words, a continuous n-multilinear map from E to F is a continuous linear map from E to the space of continuous $(n-1)$-multilinear maps from E to F.

One says that u is symmetric if it satisfies:

$$u(x_1,\dots,x_n) = u(x_{\sigma(1)},\dots,x_{\sigma(n)}) \tag{2.2}$$

for any permutation σ of the set $\{1,\dots,n\}$. One denotes by $L^n_s(E,F)$ the subspace of $L^n(E,F)$ consisting of symmetric maps.

If $u \in L^n_s(E,F)$ is a symmetric n-multilinear map on E with values in F, one associates to it a polynomial $\widetilde{u}$ homogeneous of order n by setting for $x \in E$

$$\widetilde{u}(x) = u(x,\dots,x) = u \cdot x^n. \tag{2.3}$$

Lemma 2.1.1. *Let $u \in L^n_s(E,F)$. Then the map $\widetilde{u}$ is differentiable on E and*

$$\widetilde{u}'(x) \cdot h = n \cdot u \cdot x^{n-1} \cdot h. \tag{2.4}$$

Proof. The map $\widetilde{u}$ is the composition of the linear map

$$\delta \colon E \to E^n, \quad x \mapsto (x,\dots,x)$$

and the n-linear map u. It follows that

$$\widetilde{u}'(x) \cdot h = \sum_{j=1}^{n} u'_{x_j}(x) \cdot h = n \cdot u(x,\dots,x,h).$$

q.e.d.

Differential of order n

Let U be an open subset of E and let $f\colon E\supset U\to F$ be a map. One defines by induction the notion of differential of order m at $a\in U$.

Definition 2.1.2. (a) One says that f admits a differential of order m at $a\in U$, denoted $f^{(m)}(a)$ if there exists an open neighborhood V of a such that $f^{(m-1)}(x)$ exists for any $x\in V$ and the function $f^{(m-1)}$ is differentiable at a. One then defines $f^{(m)}(a)=(f^{(m-1)})'(a)$.

(b) If f admits a differential of order m at any $x\in U$ and the function $x\mapsto f^{(m)}(x)$ from U to $L^m(E,F)$ is continuous, one says that f if of class C^m on U. One denotes by $C^m(U,F)$ the $\mathbb{R}$-vector space of functions of class C^m on U.

(c) If f admits a differential of order m for all $m\in\mathbb{N}$ one says that f is of class C^∞ and one denotes by $C^\infty(U,F)$ the $\mathbb{R}$-vector space of functions of class C^∞ on U.

Recall that $f^{(1)}(a)=f'(a)\in L(E,F)$. Hence, $f^{(n-1)}$ is a map $V\to L^{n-1}(E,F)$ and $f^{(n)}(a)\in L^n(E,F)$ by (2.1). One sets for short

$$f''=f^{(2)}.$$

For $x,y\in E$ we write $f''(a)xy$ instead of $f''(a)(x,y)$.

Theorem 2.1.3. (The Schwarz Lemma.) If $f^{(n)}(a)$ exists, then the multilinear map $f^{(n)}(a)$ is symmetric, that is, $f^{(n)}(a)\in L^n_s(E,F)$.

Proof. (i) Let us reduce the problem to the case $n=2$. We argue by induction and assume the result is proved for $n-1$, with $n\geq 2$. Therefore

$$f^{(n)}(a)\in L(E,L^{n-1}_s(E,F)).$$

Hence, $f^{(n)}(a)(x_1,\dots,x_n)$ is symmetric with respect to $(x_1,\dots,x_{n-1})$. On the other hand, $f^{(n)}(a)=(f^{(n-2)})^{(2)}(a)$. Hence

$$f^{(n)}(a)(x_1,\dots,x_n)=(f^{(n-2)})^{(2)}(a)(x_1,\dots,x_{n-2},x_{n-1},x_n)$$

is symmetric with respect to (x_{n-1},x_n). To conclude, notice that is an n-multilinear map $u(x_1,\dots,x_n)$ is symmetric with respect to $(x_1,\dots,x_{n-1})$ and with respect to $x_{n-1},x_n)$, then it is symmetric. Indeed, any permutation σ of the set $\{1,\dots,n\}$ is a composition of permutations of the set $\{1,\dots,n-1\}$ and transposition of the set $\{n-1,n\}$.

(ii) Now we assume $n = 2$. Set

$$g(x, y) = f(a + x + y) - f(a + x) - f(a + y) + f(a).$$

Theorem 1.3.3 applied to $h(y) = g(x, y) - f''(a)xy$ gives (noticing that $h(0) = 0$):

$$||g(x, y) - f''(a)xy|| \leq ||y|| \cdot \sup_{0 \leq t \leq 1} ||f'(a + x + ty) - f'(a + ty) - f''(a) \cdot x||.$$

Since f' is differentiable at a, we have

$$\begin{aligned} f'(a + ty) - f'(a) &= f''(a) \cdot ty + ||ty||\varepsilon_1(ty), \\ f'(a + x + ty) - f'(a) &= f''(a) \cdot (x + ty) + ||x + ty||\varepsilon_2(x + ty), \end{aligned}$$

where $\varepsilon_1(ty)$ and $\varepsilon_2(x + ty)$ go to 0 when $||x|| + ||y||$ goes to 0. Therefore

$$\begin{aligned} f'(a + x + ty) - f'(a + ty) &= f''(a) \cdot (x + ty) - f''(a) \cdot ty + (||y|| + ||x||)\varepsilon_3(x, y), \\ &= f''(a) \cdot x + (||y|| + ||x||)\varepsilon_3(x, y), \end{aligned}$$

and we get

$$||g(x, y) - f''(a)xy|| \leq ||y|| \cdot (||x|| + ||y||)\varepsilon_3(x, y).$$

By inverting the roles of x and y, we obtain

$$||g(x, y) - f''(a)yx|| \leq ||x|| \cdot (||x|| + ||y||)\varepsilon_3(x, y).$$

Therefore,

$$||f''(a)yx - f''(a)xy|| \leq (||x|| + ||y||)^2\varepsilon_3(x, y)$$

where $\varepsilon_3(x, y)$ goes to 0 when $||x|| + ||y||$ goes to 0.
(iii) To conclude, note that if a bilinear map u satisfies

$$||u(x, y)|| \leq (||x|| + ||y||)^2\varepsilon_(x, y)$$

where $\varepsilon(x, y)$ goes to 0 when $||x|| + ||y||$ goes to 0, then $u = 0$. Indeed, consider x_0 and y_0 and let $t > 0$. Then

$$||u(tx_0, ty_0)|| = t^2||u(x_0, y_0)|| \leq t^2(||x_0|| + ||y_0||)^2\varepsilon_(tx_0, ty_0)$$

Dividing by t^2 both terms, we get that $||u(x_0, y_0)||$ goes to 0 when t goes to 0. Since $u(x_0, y_0)$ does not depend on t, it is 0. q.e.d.

Notation 2.1.4. Let $E = E_1 \times \cdots \times E_n$ and let $f\colon E \supset U \to F$. Recall that we have denoted by $\dfrac{\partial f}{\partial x_j}(a)$ the restriction of $f'(a)$ to E_j. Assume that $f''(a)$ exists. Then $\dfrac{\partial^2 f}{\partial x_i \partial x_j}(a)$ is the restriction of $(\dfrac{\partial f}{\partial x_j})'(a)$ to E_i. This is a bilinear map on $E_i \times E_j$.

Corollary 2.1.5. *Let $E = E_1 \times \cdots \times E_n$ and let $f\colon E \supset U \to F$. Assume that $f''(a)$ exists. Then $\dfrac{\partial^2 f}{\partial x_i \partial x_j}(a) = \dfrac{\partial^2 f}{\partial x_j \partial x_i}(a)$ for any $i, j \in \{1, \dots, n\}$.*

Proof. Let $h = (h_1, \dots, h_n) \in E_1 \times \cdots \times E_n$. One has

$$f'(a) \cdot h = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(a) \cdot h_j,$$

$$f''(a) \cdot h \cdot h = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a) \cdot h_j \cdot h_i.$$

The bilinear map $f''(a)$ is represented by the $(n \times n)$-matrix

$$f''(a) = \left(\ \frac{\partial^2 f}{\partial x_i \partial x_j}(a) \ \right)_{1 \le i,j \le n} \tag{2.5}$$

Since the bilinear map $f''(a)$ is symmetric, the matrix is symmetric. q.e.d.

Definition 2.1.6. The matrix in (2.5) is called the Hessian matrix of f at a.

One shall not confuse the Jacobian matrix (an $(p \times n)$-matrix) and the Hessian matrix (a symmetric $(n \times n)$-matrix).

Notation 2.1.7. Let $E = E_1 \times \cdots \times E_n$ and let $f\colon E \supset U \to F$. We set for short

$$\frac{\partial f}{\partial x_j} = f'_{x_j}, \quad \frac{\partial^2 f}{\partial x_i \partial x_j} = f''_{x_i,x_j}, \qquad \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k} = f^{(3)}_{x_i,x_j,x_k},$$

and so on.

Second differential of a composition

We shall consider the situation of Proposition 1.1.5 and calculate the second derivative. First we need a lemma.

Lemma 2.1.8. *Let E, F_1, F_2, F_3 be normed spaces, let U be an open subset of E and let $a \in U$. Let $u\colon U \to L(F_1, F_2)$ and $v\colon U \to L(F_2, F_3)$ be two maps. For $x \in U$, set $w(x) = v(x) \circ u(x) \in L(F_1, F_3)$. Assume that $u'(a) \in L(E, L(F_1, F_2))$ and $v'(a) \in L(E, L(F_2, F_3))$ exist. Then $w'(a) \in L(E, L(F_1, F_3))$ exists and for $h \in E$, one has:*

$$w'(a) \cdot h = (v'(a) \cdot h) \circ u(a) + v(a) \circ (u'(a) \cdot h).$$

Proof. The proof is left as an exercise. q.e.d.

Proposition 2.1.9. *Let E, F, G be normed spaces, let U an open subset of E and V an open subset of F. Let $f\colon E \supset U \to V \subset F$ and $g\colon F \supset V \to G$ be two maps. Let $a \in U$ and set $b = f(a)$. Assume that $f''(a)$ exists and $g''(b)$ exists. Set $h := g \circ f$. Then $h''(a)$ exists and*

$$h''(a) = g''(b) \circ f'(a) + g'(b) \circ f''(a).$$

Note that $f'(a) \in L(E, F)$, $g'(b) \in L(F, G)$, $f''(a) \in L_s^2(E, F)$, $g''(b) \in L_s^2(F, G)$, $h''(a) \in L_s^2(E, G)$ and recall that

$$\begin{aligned}(g''(b) \circ f'(a))(h_1, h_2) &= g''(b)(f'(a)h_1, f'(a)h_2),\\ (g'(b) \circ f''(a))(h_1, h_2) &= g'(b)(f''(a)(h_1, h_2)).\end{aligned}$$

Proof. By applying Proposition 1.1.5 we have $h'(x) = g'(f(x)) \circ f'(x)$, and the result follows from Lemma 2.1.8. q.e.d.

2.2 The Taylor formula

Let U be an open subset of E and let $f\colon E \supset U \to F$ be a map as above. Let $a \in U$. Assume that f admits differentials of order $n-1$ in a neighborhood of a and that $f^{(n)}(a)$ exists.

Set $h^n := (h, \ldots, h) \in E^n$ and $f^{(n)}(a) \cdot h^n = f^{(n)}(a)(h, \ldots, h) \in F$. Define $R_n(a, h) \in F$ by the formula:

$$f(a+h) = f(a) + \frac{1}{1!} f'(a) \cdot h + \cdots + \frac{1}{n!} f^{(n)}(a) \cdot h^n + R_n(a, h). \tag{2.6}$$

Theorem 2.2.1. (i) *One has $R_n(a, h) = ||h||^n \varepsilon(a, h)$ where $\varepsilon(a, h)$ goes to 0 when h goes to 0.*

(ii) *Assume moreover that $f^{(n+1)}(x)$ exists on the interval $[a, a+h] \subset U$ and $||f^{(n+1)}(x)|| \leq M$ for some constant M on this interval. Then*

$$||R_n(a, h)|| \leq M \cdot \frac{||h||^{n+1}}{(n+1)!}.$$

(iii) *Assume now that f is of class C^{n+1} on U and F is complete. Then if the interval $[a, a+h]$ is contained in U, one has*

$$R_n(a,h) = \int_0^1 \frac{(1-t)^n}{n!} f^{(n+1)}(a+th)h^{n+1}dt.$$

Proof. (i) We argue by induction on n. For $n=1$ this is nothing but the definition of the differential. Assume the theorem is proved for $n-1$. Set

$$g(x) = f(a+x) - \big(f(a) + \frac{1}{1!}f'(a)\cdot x + \cdots + \frac{1}{n!}f^{(n)}(a)\cdot x^n\big).$$

Then

$$g'(x) = f'(a+x) - \big(f'(a) + \frac{1}{1!}f''(a)\cdot x + \cdot x + \cdots + \frac{1}{(n-1)!}f^{(n)}(a)\cdot x^{n-1}\big).$$

Note that $g(x) \in F$ and $g'(x) \in L(E,F)$. Also note that $g(0) = 0$ and $g(h) = R_n(a,h)$.

The induction hypothesis, applied to the function $f'(a+x)$, gives

$$g'(x) = ||x||^{n-1}\alpha(x)$$

where $\alpha(x)$ goes to 0 when x goes to 0. In other words, for any $\varepsilon > 0$ there exists $\eta > 0$ such that $||x|| \leq \eta$ implies $\alpha(x) \leq \varepsilon$. Hence,

$$||g'(x)|| \leq \varepsilon ||x||^{n-1} \text{ for } ||x|| \leq \varepsilon.$$

Applying Theorem 1.3.3 (more precisely, applying Corollary 1.3.5), we get

$$||g(x)|| \leq \varepsilon ||x||^n \text{ for } ||x|| \leq \eta.$$

This shows that the function $\frac{g(h)}{||h||^n}$ goes to 0 when $|h|$ goes to 0 with $h \neq 0$.

(ii) First, consider a function $w\colon [0,1] \to F$, w of class C^n and such that $w^{(n+1)}(t)$ exists for $t \in [0,1]$ and satisfying:

$$\sup_{t\in[0,1]} ||w^{(n+1)}(t)|| \leq C, \quad w(0) = w'(0) = \cdots = w^{(n)}(0) = 0.$$

Let us prove that

$$||w(t)|| \leq \frac{t^{n+1}}{(n+1)!}C \tag{2.7}$$

We argue by induction on n. By assuming that the result is proved to the order $n-1$, we get

$$||w'(t)|| \leq \frac{t^n}{n!}C.$$

Then (2.7) follows by applying Lemma 1.3.1.

Now consider a function $v\colon [0,1] \to F$, v of class C^n and such that $v^{(n+1)}(t)$ exists for $t \in [0,1]$ and satisfying $\sup_{t\in[0,1]} ||v^{(n+1)}(t)|| \leq C$. By applying (2.7) to the function

$$w(t) = v(t) - \big(v(0) + \frac{1}{1!}v'(0) + \cdots + \frac{1}{n!}v^{(n)}(0)\big)$$

we get

$$||v(t) - \big(v(0) + \frac{t}{1!}v'(0) + \cdots + \frac{t^n}{n!}v^{(n)}(0)\big)|| \leq \frac{t^{n+1}}{(n+1)!}C. \tag{2.8}$$

To conclude, let us apply (2.8) to the function $v(t) = f(a+th)$. Since $v^{(p)}(t) = f^{(p)}(a+th)\cdot h^p$, we have

$$||v^{(n+1)}(t)|| \leq M \cdot ||h^{n+1}||$$

and we get the result.

(iii) Set $\varphi(t) = f(a+th)$. Then $\varphi^{(p)}(t) = f^{(p)}(a+th)\cdot h^p$ and

$$\frac{\partial}{\partial t}\big(\varphi(t) + \frac{(1-t)}{1!}\varphi'(t) + \cdots + \frac{(1-t)^n}{n!}\varphi^{(n)}(t)\big) = \frac{(1-t)^n}{n!}\varphi^{(n+1)}(t). \tag{2.9}$$

Let us integrate both terms of (2.9) from 0 to 1. Since

$$\int_0^1 \frac{\partial}{\partial t}\big(\frac{(1-t)^p}{p!}\varphi^{(p)}(t)\big)dt = \begin{cases} \varphi(1) - \varphi(0) & \text{if } p = 0,\\ -\dfrac{\varphi^{(p)}}{p!}(0) & \text{if } p > 0,\end{cases}$$

we obtain

$$\varphi(1) - \big(\varphi(0) + \frac{1}{1!}\varphi'(0) + \cdots + \frac{1}{n!}\varphi^{(n)}(0)\big) = \int_0^1 \frac{(1-t)^n}{n!}\varphi^{(n+1)}(t)dt.$$

Since $\varphi^{(p)}(t) = f^{(p)}(a+th)\cdot h^p$, we get the result. q.e.d.

Let us consider the particular case where $E = \mathbb{R}^n$. In this case $f'(a) \cdot h = \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(a)$. Denote by θ_h the differential operator

$$\theta_h = \sum_{i=1}^n h_i \frac{\partial}{\partial x_i}$$

Then $f'(a) \cdot h = \theta_h(f)(a)$, $f''(a) \cdot h^2 = \theta_h(\theta_h(f))(a)$ and more generally:

$$f^{(p)}(a) \cdot h^p = \theta_h^p(f)(a).$$

Notation 2.2.2. Let us recall some classical and useful notations. For $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$, one sets

$$|\alpha| = \alpha_1 + \cdots + \alpha_n, \quad \alpha! = \alpha_1! \dots \alpha_n! \tag{2.10}$$

and for a ring A and $y = (y_1, \dots, y_n) \in A^n$, one sets

$$y^\alpha = y_1^{\alpha_1} \cdots y_n^{\alpha_n} \tag{2.11}$$

One then has the generalized binomial formula:

$$(y_1 + \cdots + y_n)^p = \sum_{|\alpha|=p} \frac{p!}{\alpha!} y^\alpha. \tag{2.12}$$

Using these notations and formula 2.12, we get for $h = (h_1, \dots, h_n)$:

$$\frac{1}{p!} f^{(p)}(a) \cdot h^p = \sum_{|\alpha|=p} \frac{h^\alpha}{\alpha!} f_{x^\alpha}^{(p)}(a).$$

Example 2.2.3. (i) Assume $E = \mathbb{R}^2$, $a = 0$. The Taylor formula to the order 3 gives:

$$\begin{aligned} f(x,y) &= f(0,0) \\ &\quad + x f'_x(0,0) + y f'_y(0,0) \\ &\quad + \frac{x^2}{2!} f''_{xx}(0,0) + \frac{xy}{1!1!} f''_{xy}(0,0) + \frac{y^2}{2!} f''_{yy}(0,0) \\ &\quad + \frac{x^3}{3!} f^{(3)}_{xxx}(0,0) + \frac{x^2 y}{2!1!} f^{(3)}_{xxy}(0,0) + \frac{xy^2}{1!2!} f^{(3)}_{xyy}(0,0) + \frac{y^3}{3!} f^{(3)}_{yyy}(0,0) \\ &\quad + (|x| + |y|)^3 \varepsilon(x,y). \end{aligned}$$

(ii) Now assume $E = \mathbb{R}^3$, $a = 0$. Taylor's formular to the order 2 gives:

$$\begin{aligned} f(x,y,z) &= f(0,0,0) \\ &\quad + x f'_x(0,0) + y f'_y(0,0) + z f'_z(0,0) \\ &\quad + \frac{x^2}{2!} f''_{xx}(0,0,0) + \frac{y^2}{2!} f''_{yy}(0,0,0) + \frac{z^2}{2!} f''_{zz}(0,0,0) \\ &\qquad + \frac{xy}{1!1!} f''_{xy}(0,0,0) + \frac{xz}{1!1!} f''_{xz}(0,0,0) + \frac{yz}{1!1!} f''_{yz}(0,0,0) \\ &\quad + (|x| + |y| + |z|)^2 \varepsilon(x,y,z). \end{aligned}$$

2.3 Applications to extremum

In Section 1.2, we have given a necessary condition for a real valued function to have a local extremum. Here, we shall refined it and also give a sufficient condition.

Positive definite forms

Definition 2.3.1. Let $(E, ||\cdot||)$ be a real normed space and let $u\colon E \times E \to \mathbb{R}$ be a symmetric bilinear form.

(i) One says that u is positive semidefinite if $u(h,h) \geq 0$ for any $h \in E$.

(ii) One says that u is positive definite if there exists a constant $c > 0$ such that $u(h,h) \geq c \cdot ||h||^2$ for any $h \in E$.

(iii) One says that u is negative semidefinite (resp. negative definite) if $-u$ is positive semidefinite (resp. positive definite).

We shall concentrate our study on positive forms.

Lemma 2.3.2. *Assume that E is finite dimensional. Then u is positive definite if and only if $u(h,h) > 0$ for any $h \neq 0$.*

Proof. The unit sphere being compact, there exists $c > 0$ such that $u(h,h) > c$ for any h with $||h|| = 1$. For $h \neq 0$ one has

$$u(h,h) = ||h||^2 \cdot u(\frac{h}{||h||}, \frac{h}{||h||})$$

and the result follows. q.e.d.

Now consider a real vector space E endowed with a basis $(e_1, \dots, e_n)$ (*e.g.*, $E = \mathbb{R}^n$) and let u be a symmetric bilinear form on E. To u one associates the symmetric $(n \times n)$-matrix with entries $u(e_i, e_j)$. Conversely, if $A = \{a_{ij}\}_{i,j=1,\dots,n}$ is a symmetric $(n \times n)$-matrix, it defines a bilinear form on E by setting for $h = (h_1, \dots, h_n)$:

$$u(h,h) = (h_1, \dots, h_n) \begin{pmatrix} a_{11} & \cdot & \cdot & \cdot & a_{1n} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{n1} & \cdot & \cdot & \cdot & a_{nn} \end{pmatrix} \begin{pmatrix} h_1 \\ \cdot \\ \cdot \\ h_n \end{pmatrix} \tag{2.13}$$

For a square matrix $A = \{a_{ij}\}_{i,j=1,\dots,n}$ and for $1 \le p \le n$ let us denote by Δ_p the determinant of the matrix $\{a_{ij}\}_{1 \le i,j \le p}$.

The next classical result will be admitted without proof.

Lemma 2.3.3. *Let u be a symmetric bilinear form on the vector space E endowed with a basis $(e_1, \dots, e_n)$ and let A be the associated symmetric matrix. Then u is positive definite if and only if $\Delta_p > 0$ for all p with $1 \le p \le n$.*

One says that a symmetric square matrix is positive semidefinite (resp. positive definite) if the bilinear form it defines on $\mathbb{R}^n$ is positive semidefinite (resp. positive definite) and similarly with negative instead of positive.

One shall be aware that for a matrix, to be negative definite does not imply and is not implied by the condition that $\Delta_p < 0$ for all p with $1 \le p \le n$.

Necessary and sufficient conditions for extremum

Theorem 2.3.4. *Let $E \supset U \xrightarrow{f} \mathbb{R}$ be a map and let $a \in U$ with $f'(a) = 0$.*

(i) *Assume that $f''(a)$ is positive definite. Then f has a local strict minimum at a.*

(ii) *Assume that f has a local minimum at a. Then $f''(a)$ is positive semidefinite.*

Proof. (i) One has

$$f(a+x) - f(a) = \frac{1}{2} f''(a) \cdot x^2 + ||x||^2 \varepsilon(x) \tag{2.14}$$

where $\varepsilon(x)$ goes to 0 when x goes to 0. There exists $c > 0$ such that $f''(a) \cdot x^2 \ge c \cdot ||x||^2$ and there exists $\eta > 0$ such that $||\varepsilon(x)|| \le \frac{c}{4}$ for $||x|| \le \eta$. Therefore,

$$f(a+x) \ge f(a) + \frac{c}{4} ||x||^2 \text{ for } ||x|| \le \eta.$$

(ii) Let us argue by contradiction and assume there exists $h \in E$ with $f''(a) \cdot h^2 = c < 0$. Then $f''(a) \cdot (th)^2 = t^2 c$. Let us choose $\eta > 0$ such that $|\varepsilon(x)| \leq \frac{c}{4||h||^2}$ for $||x|| \leq \eta$, where $\varepsilon(x)$ is given in (2.14). For $||th|| \leq \eta$ we have:

$$\begin{aligned} f(a+th) - f(a) &= \frac{1}{2}t^2 f''(a) \cdot h^2 + ||th||^2 \varepsilon(th) \\ &= t^2 \frac{c}{2} + t^2 ||h||^2 \varepsilon(th). \end{aligned}$$

Since $|t^2||h||^2\varepsilon(th)| \leq t^2 \frac{c}{4}$, we get $f(a+th) - f(a) \leq t^2 \frac{c}{4} < 0$ and a is not a local minimum. q.e.d.

Corollary 2.3.5. *Let $\mathbb{R}^n \supset U \xrightarrow{f} \mathbb{R}$ be a map and let $a \in U$. Assume that $f'(a) = 0$ and the Hessian $H_f(a)$ is positive definite. Then f has a local strict minimum at a.*

The conditions in Theorem 2.3.4 (i) and (ii) are not necessary and sufficient conditions as shown by the examples below.

Example 2.3.6. (i) Consider the map $f\colon \mathbb{R}^2 \to \mathbb{R}$, $(x, y) \mapsto x^4 + y^4$. Then f has a strict minimum at $(0, 0)$. However, $f''(0, 0) = 0$ is not positive definite.
(ii) Consider the map $f\colon \mathbb{R}^2 \to \mathbb{R}$, $(x, y) \mapsto x^2 - y^4$. Then $f''(0, 0)$ is positive semidefinite, but $(0, 0)$ is not a local minimum of f.

Exercises to Chapter 2

Exercise 2.1. We follow the notations of Example 1.1.7: A is a Banach algebra, Ω denotes the open set of invertible elements and $J\colon \Omega \to \Omega$ is the map $x \mapsto x^{-1}$. Calculate $J''(a) \cdot h^2$ for $a \in \Omega$ and $h \in A$.

Exercise 2.2. Calculate the Taylor series at $(0, 0)$ of the function $f(x, y) = \frac{1}{1-x^2+y^2}$.

Exercise 2.3. Calculate the Taylor series to the order 3 at $(0, 0)$ of the functions $f(x, y) = xy \exp(\frac{x}{1-y})$ and $g(x, y) = (\dfrac{1 - x^2 - y^2}{1 - y})^{\frac{1}{2}}$.

Exercise 2.4. Search the local extrema of the functions from $\mathbb{R}^2$ to $\mathbb{R}$ given by $f(x, y) = x^4 + y^2$, $g(x, y) = xy$ and $h(x, y) = x^2 - y^2 - \frac{1}{4}y^4$ and discuss their nature (local/global, minimum/maximum).

Exercise 2.5. Consider the function $f\colon \mathbb{R}^2 \to \mathbb{R}$ given by $f(x, y) = x^4 + 2y^4 - 2x^2y^2 - x^2 - 2y^2$.
(i) Show that $(u - v)^2 + v^2 \geq \frac{1}{4}(u^2 + v^2)$ $(u, v \in \mathbb{R})$.
(ii) Deduce that
(a) $f(x, y)$ goes to $+\infty$ when $|x| + |y|$ goes to $+\infty$,
(b) the set $\{(x, y); f(x, y) \leq 0\}$ is compact,
(c) f reaches its minimum on $\mathbb{R}^2$.
(iii) (a) Find all critical points of f, that is, the points where the differential of f vanishes.
(b) For each of this critical point of f, write the Hessian matrix of f.
(c) At which points of $\mathbb{R}^2$ does f have a local minimum (respectively, a local maximum)? Find the minimum of f.

Exercise 2.6. Consider the function $f\colon \mathbb{R}^2 \to \mathbb{R}$ given by $f(x, y) = x^2 + y^2 - xy - 2x + y$.
(i) Show that
(a) $f(x, y)$ goes to $+\infty$ when $|x| + |y|$ goes to $+\infty$,
(b) the set $\{(x, y); f(x, y) \leq 0\}$ is compact,
(c) f reaches its minimum on $\mathbb{R}^2$.
(ii) Calculate this minimum.

Chapter 3

Submanifolds

3.1 The local inversion theorem

Let E and F be two Banach spaces. A continuous linear map $u \in L(E, F)$ is a continuous linear isomorphism if there exists $v \in L(F, E)$ such that $v \circ u = \mathrm{id}_E$ and $u \circ v = \mathrm{id}_F$. We shall admit here that if u is continuous and if there exists a linear map v such that $v \circ u = \mathrm{id}_E$ and $u \circ v = \mathrm{id}_F$, then v is continuous. (This result is no more true when removing the hypothesis that E and F are complete.)

Definition 3.1.1. Let E and F be two Banach spaces and let $U \subset E$ and $V \subset F$ be two open subsets. For $k \in \mathbb{N} \sqcup \{\infty\}$, a C^k-isomorphism $f \colon U \xrightarrow{\sim} V$ is a bijective map of class C^k such that the inverse map f^{-1} is also of class C^k.

- A C^0-isomorphism is nothing else but a topological isomorphism, also called a homeomorphism.
- Ley U and V be as above and let W be an open subset of a Banach space G. If $f \colon U \xrightarrow{\sim} V$ and $g \colon V \xrightarrow{\sim} W$ are C^k-isomorphisms, then $g \circ f \colon U \to W$ is a C^k-isomorphism.
- A linear C^0-isomorphism $u \in L(E, F)$ is a C^∞-isomorphism. Indeed, a continuous linear map is of class C^∞.
- $E \supset U \xrightarrow{f} F$ be a map of class C^1. Then the set of $a \in U$ such that $f'(a)$ is a linear C^0-isomorphism is open in U.

Example 3.1.2. The map $t \mapsto t^3$ from $\mathbb{R}$ to $\mathbb{R}$ is a C^0-isomorphism and is of class C^∞, but it is not a C^1-isomorphism.

Now assume that $f\colon U \xrightarrow{\sim} V$ is a C^k-isomorphism with $k \geq 1$ and let $a \in U$. Set $b = f(a)$ and $g = f^{-1}$. Then $g \circ f = \mathrm{id}_E$ implies $g'(b) \circ f'(a) = \mathrm{id}_E$, and similarly, $f'(a) \circ g'(b) = \mathrm{id}_F$. Therefore, $f'(a) \in L(E, F)$ is a continuous linear isomorphism. There is a converse statement.

Theorem 3.1.3. The local inversion theorem. *Let E and F be two Banach spaces and let $E \supset U \xrightarrow{f} F$ be a map of class C^k with $k \geq 1$, defined on an open set U of E. Let $a \in U$. Assume that $f'(a)$ is a linear C^0-isomorphism. Then there exists an open neighborhood V of $f(a)$ and an open neighborhood W of a with $W \subset U$ such that the restriction of f to W is a C^k-isomorphism $f|_W\colon W \xrightarrow{\sim} V$.*

Proof. (i) Since $f'(a)$ is an isomorphism, we may replace f with $f'(a)^{-1} \circ f$ and assume from the beginning that $F = E$ and $f'(a) = \mathrm{id}_E$. We set $b = f(a)$. Morover, since $f'(x)$ is invertible in a neighborhood of a and by replacing U with another neighborhood of a we may assume from the beginning that $f'(x)$ is invertible for $x \in U$.
(ii) Set $\Phi(x) = f(x) - x$. Then Φ is C^k and $\Phi'(a) = 0$. Hence, for any $C > 0$ there exists $r > 0$ such that $||x - a|| \leq r$ implies $||\Phi'(x)|| \leq C$. Applying Theorem 1.3.3, we get

$$x, x' \in B(a, r) \Rightarrow ||\Phi(x) - \Phi(x')|| \leq C \cdot ||x - x'||.$$

We shall repeat the argument of the fixed point theorem. Let us choose C with $0 < C < 1$. For $y \in B(b, (1 - C)r)$, we define a sequence $(x_n)_n$ as follows. We start with $x_0 = a$ and we set

$$x_{n+1} = y - \Phi(x_n). \tag{3.1}$$

Let us show that $x_n \in B(a, r)$ for all n. We have

$$||x_{n+1} - x_n|| \leq ||\Phi(x_n) - \Phi(x_{n-1})|| \leq C \cdot ||x_n - x_{n-1}||.$$

Therefore

$$||x_{n+1} - x_n|| \leq C^n \cdot ||x_1 - x_0||,$$

and

$$\begin{aligned} ||x_{n+1} - x_0|| &\leq ||x_{n+1} - x_n|| + ||x_n - x_{n-1}|| + \cdots + ||x_1 - x_0|| \\ &\leq (C^n + C^{n-1} + \cdots + 1)||x_1 - x_0|| \end{aligned}$$

Recall that $x_0 = a$, $y = \Phi(a) + x_1 = b - a + x_1$, so that $y - b = x_1 - x_0$. Therefore

$$||x_{n+1} - x_0|| \leq \frac{1 - C^n}{1 - C}||y - b|| < r.$$

This shows that $x_{n+1} \in B(a,r)$. Moreover, $(x_n)_n$ is a Cauchy sequence since

$$||x_{n+p} - x_n|| \leq (C^n + \cdots + C^{n+p})||x_1 - x_0||$$

and $C^n + \cdots + C^{n+p} \leq \frac{C^n}{1-C}$ goes to 0 when n goes to ∞. Since E is complete, the sequence $(x_n)_n$ has a limit x. Since

$$||x_n - x_0|| \leq \frac{1}{1-C}||y-b||,$$

we get $||x - x_0|| \leq \frac{1}{1-C}||y-b|| < r$ which shows that $x \in B(a,r)$. Making n goes to infinity in (3.1), we get

$$x = y - \Phi(x), \text{ or equivalently } y = f(x).$$

Moreover, x is unique in $B(a,r)$ since $x = y - \Phi(x)$ and $x' = y - \Phi(x')$ implies $x - x' = \Phi(x) - \Phi(x')$ hence $||x - x'|| \leq C \cdot ||x - x'||$ and $x = x'$. To summarize, we have proved that for any $y \in B(b,(1-C)r)$, there exists a unique $x \in B(a,r)$ such that $y = f(x)$.

(iii) Set $V := B(b,(1-C)r)$ and $W := f^{-1}(V) \cap B(a,r)$. Then W is open and $f\colon W \to V$ is continuous and bijective. Let us show that the inverse map f^{-1} is continuous. Let $y = f(x)$ and $y' = f(x')$. One has $y - y' = x - x' + \Phi(x) - \Phi(x')$ and $||\Phi(x) - \Phi(x')|| \leq C \cdot ||x - x'||$. Therefore

$$||y - y'|| \geq (1-C)||x - x'|| \tag{3.2}$$

or equivalently

$$||f^{-1}(y) - f^{-1}(y')|| \leq \frac{1}{1-C}||y - y'||.$$

We have thus proved that $f\colon W \to V$ is a C^0-isomorphism.

(iv) Let us show that f^{-1} is C^1. Set for short $g = f^{-1}$. For $x_0 \in W$, we have

$$f(x) - f(x_0) - f'(x_0) \cdot (x - x_0) = ||x - x_0|| \cdot \varepsilon_1(x_0, x) \tag{3.3}$$

where $\varepsilon_1(x_0,x)$ goes to 0 when x goes to x_0. Recall that $f'(x)$ is invertible for $x \in U$. Applying $f'(x_0)^{-1}$ to (3.3), we get

$$x - x_0 - f'(x_0)^{-1} \cdot (f(x) - f(x_0)) = ||x - x_0|| \cdot \varepsilon_2(x_0, x) \tag{3.4}$$

where $\varepsilon_2(x_0,x) = f'(x_0)^{-1}(\varepsilon_1(x_0,x))$ goes to 0 when x goes to x_0. Set $y = f(x)$, $y_0 = f(x_0)$, hence $x = g(y)$, $x_0 = g(y_0)$. Since f is a C^0-isomorphism, $\varepsilon_2(x_0,x) = \varepsilon_2(f^{-1}(y_0), f^{-1}(y)) = \varepsilon_3(y_0,y)$ goes to 0 when y goes to y_0. Using (3.2), we have

$$||x - x_0|| \leq \frac{1}{1-C} \cdot ||y - y_0||.$$

It follows that

$$||x - x_0|| \cdot \varepsilon_2(x_0, x) = ||y - y_0|| \cdot \varepsilon_4(y_0, y)$$

where $\varepsilon_4(y_0, y)$ goes to 0 when y goes to y_0. Then (3.4) reads as

$$g(y) - g(y_0) - f'(x_0)^{-1} \cdot (y - y_0) = ||y - y_0|| \cdot \varepsilon_4(y_0, y). \tag{3.5}$$

This shows that $g'(y_0)$ exists and

$$g'(y_0) = f'(x_0)^{-1}. \tag{3.6}$$

Recall Example 1.1.7 and denote by Ω the open subset of $L(E, E)$ consisting of invertible maps and by $J\colon \Omega \to \Omega$ the map $u \mapsto u^{-1}$. Then (3.6) reads as $g'(y_0) = f'^{-1}(f^{-1}(y_0))$, that is,

$$g' = J \circ f' \circ f^{-1}. \tag{3.7}$$

We have already seen in (1.4) that J is differentiable and that $J'(a)h = -a^{-1}ha^{-1}$. It follows easily that J is of class C^∞.

Now assume we have proved that when f is of class C^{k-1} then $g = f^{-1}$ is of class C^{k-1}. If f is of class C^k, it follows from (3.7) that g' is of class C^{k-1}. Hence, g is of class C^k. q.e.d.

Corollary 3.1.4. *Let E and F be Banach spaces and let $E \supset U \xrightarrow{f} F$ be a map of class C^1. Assume that $f'(x)$ is invertible for any $x \in U$. Then f is open, that is, for any V open in U, the set $f(V)$ is open in F.*

Proof. It is enough to prove that $f(U)$ is open. For any $a \in U$ there exists an open neighborhood W of a and an open neighborhood V of $f(a)$ such that $f(W) = V$. Therefore, $f(U)$ is a neighborhood of each of its point, hence, is open. q.e.d.

Corollary 3.1.5. *Let E and F be Banach spaces and let $E \supset U \xrightarrow{f} V \subset F$ with U and V open and f of class C^k $(k \geq 1)$. Assume that f is bijective and $f'(x)$ is invertible for any $x \in U$. Then f is a C^k-isomorphism.*

Example 3.1.6. Let $U = \mathbb{R}^2 \setminus \{0\}$ and let $f\colon U \to \mathbb{R}^2$ be the map

$$f\colon (x, y) \mapsto (x^2 - y^2, 2xy).$$

The Jacobian matrix of f is invertible at each $(x, y) \in U$. Therefore, $f(U)$ is open. Note that, setting $z = x + iy$, f is the map $z \mapsto z^2$.

Example 3.1.7. Let $E = C^0([0,1],\mathbb{R})$. Denote by U the subset of E consisting of functions φ such that $\varphi(t) > 0$ for all $t \in [0,1]$. We have already noticed that U is open in E. Let us show that for $||\psi|| \ll 1$, the equation

$$\psi = \ln\varphi \int_0^1 \varphi(t)dt$$

admits a solution $\varphi \in U$.

Consider the map

$$f\colon U \to E, \quad \varphi \mapsto \ln\varphi \int_0^1 \varphi(t)dt.$$

Then

$$f'(\varphi)\cdot\psi = \frac{\psi}{\varphi}\int_0^1 \varphi(t)dt + \ln\varphi \int_0^1 \psi(t)dt.$$

Hence, $f'(\mathbf{1}) = \mathrm{id}_E$. (We denote by $\mathbf{1}$ the function $t \mapsto 1$.) Applying Theorem 3.1.3, we deduce that there exists a neighborhood W of $\mathbf{1}$ and a neighborhood V of 0 such that $f|_W$ is an isomorphism $W \xrightarrow{\sim} V$.

The implicit functions theorem

Before to state the theorem, let us take an example. Consider the circle $\{(x,y) \in \mathbb{R}^2; x^2+y^2 = 1\}$, that is, the set $\{(x,y) \in \mathbb{R}^2; f(x,y) = 0\}$ where $f(x,y) = x^2+y^2-1$. For $y > 0$ it is the graph of the function $y = \sqrt{1-x^2}$, for $y < 0$ it is the graph of the function $y = -\sqrt{1-x^2}$, for $x > 0$, it is the graph of the function $x = \sqrt{1-y^2}$ and for $x < 0$, it is the graph of the function $x = -\sqrt{1-y^2}$. One sees that, *locally on* $\mathbb{R}^2$, the equation $f(x,y) = 0$ is equivalent to either an equation $y = g(x)$ or an equation $x = h(y)$.

In the sequel, for a function f defined on an open subset of $E \times F$, we write for short f'_x and f'_y instead of f'_E and f'_F.

Theorem 3.1.8. The implicit functions theorem. *Let E, F, G be three Banach spaces, U an open subset of $E \times F$, $f\colon U \to G$ a map of class C^k $(k \geq 1)$ and let $(a,b) \in U$ with $f(a,b) = 0$. Assume that $f'_y(a,b) \in L(F,G)$ is invertible. Then there exists an open neighborhood V of (a,b) in U, an open neighborhood W of a in E and a map $g\colon W \to F$ of class C^k such that*

$$(x,y) \in V, f(x,y) = 0 \Leftrightarrow x \in W, y = g(x).$$

Moreover $g'(x) = -{f'_y}^{-1}(x,g(x)) \circ f'_x(x,g(x))$ *for* $(x,y) \in V$.

Proof. Consider the map

$$\widetilde{f}\colon U \to E \times G, \quad (x,y) \mapsto (x, f(x,y)).$$

The differential is given by the matrix $\widetilde{f}'(x,y) = \begin{pmatrix} \mathrm{id}_E & 0 \\ f'_x & f'_y \end{pmatrix}$. This matrix is invertible at (a,b). Indeed, recall that for an invertible linear map u, the inverse of the matrix $\begin{pmatrix} \mathrm{id}_E & 0 \\ v & u \end{pmatrix}$ is the matrix $\begin{pmatrix} \mathrm{id}_E & 0 \\ -u^{-1}\circ v & u^{-1} \end{pmatrix}$.

Therefore, we may apply Theorem 3.1.3 to $\widetilde{f}$. There exists an open neighborhood V of (a,b) in $E \times F$ and an open subset $\widetilde{W} \subset E \times G$ such that $\widetilde{f}$ induces a C^k-isomorphism $\widetilde{f}\colon V \xrightarrow{\sim} \widetilde{W}$. Set $\Phi = \widetilde{f}^{-1}$. Since $\widetilde{f}(x,y) = (x, f(x,y))$, we have

$$\Phi(u,v) = (u, \varphi(u,v)) \text{ for } (u,v) \in \widetilde{W} \subset E \times G.$$

Since $(\widetilde{f} \circ \Phi)(x,v) = (x, f(x, \varphi(x,v))) = (x,v)$, we get $f(x,\varphi(x,v)) = v$. Then we set

$$W = \widetilde{W} \cap E \times \{0\}, \quad g(x) = \varphi(x,0).$$

Since $\varphi(u,v)$ is C^k, g is C^k. Finally, let us calculate the differential of $0 \equiv f(x,g(x))$. We have

$$f'(x,y) = f'_x dx + f'_y dy = f'_x dx + f'_y \circ g'(x) dx,$$

hence

$$0 = f'_x(x,g(x)) + f'_y(x,g(x)) \circ g'(x)$$

and the result for $g'(x)$ follows. q.e.d.

3.2 The constant rank theorem

In this section, the vector spaces E, F, etc. are finite dimensional.

Definition 3.2.1. Let U be an open subset of E, let $a \in U$ and let $E \supset U \xrightarrow{f} F$ be a map. Assume that $f'(a)$ exists. The rank of f at a, denoted $\mathrm{rk(f)(a)}$ is the rank of the linear map $f'(a)$.

- Recall that the rank of a linear map $u\colon E \to F$ is

$$\mathrm{rk(u) := \dim u(E) = \dim E - \dim(Ker\, u)}.$$

- If E and F are endowed with basis, then the Jacobian matrix $J_f(a)$ is well-defined, and the rank of f at a is the rank of this matrix.

- Assume that f is C^1 on U. If $\mathrm{rk(f)(a) = r}$, then $\mathrm{rk(f)(x) \geq r}$ for x in a neighborhood of a. Indeed, we may endow E and F with basis. Then J_f has rank $\geq r$ at a if there exists a square $(r \times r)$-matrix extracted from $J_f(x)$ with non zero-determinant at $x = a$. This determinant will remain different from 0 for x in a neighborhood of a.

Consider another map $E \supset \widetilde{U} \xrightarrow{\widetilde{f}} F$ with $\widetilde{U}$ open in E and let $\widetilde{a} \in \widetilde{U}$.

Definition 3.2.2. Let us say that the germ of f at a and the germ of $\widetilde{f}$ at $\widetilde{a}$ are C^k-equivalent if there exist:

(i) open neighborhoods W of a and $\widetilde{W}$ of $\widetilde{a}$ in E,

(ii) open subsets V and $\widetilde{V}$ in F,

(iii) C^k-isomorphisms $\varphi\colon W \xrightarrow{\sim} \widetilde{W}$ and $\psi\colon V \xrightarrow{\sim} \widetilde{V}$

such that $\varphi(a) = \widetilde{a}$ and $\widetilde{f} = \psi \circ f \circ \varphi^{-1}$.

This is visualized by the commutative diagram

$$
\begin{array}{ccc}
E \supset W & \xrightarrow{\quad f \quad} & V \subset F \\
{\scriptstyle \varphi}\big\downarrow{\scriptstyle \sim} & & {\scriptstyle \sim}\big\downarrow{\scriptstyle \psi} \\
E \supset \widetilde{W} & \xrightarrow{\quad \widetilde{f} \quad} & \widetilde{V} \subset F.
\end{array}
\tag{3.8}
$$

Clearly, if the germ of f at a and the germ of $\widetilde{f}$ at $\widetilde{a}$ are C^k-equivalent, then $\mathrm{rk(f)(a)} = \mathrm{rk}(\widetilde{\mathrm{f}})(\widetilde{\mathrm{a}})$.

Definition 3.2.3. Let $E \subset U \xrightarrow{f} F$ be a map of class C^k $(k \geq 1)$ and let $a \in U$. One says that

(i) f is a submersion at a if $\mathrm{rk(f)(a) = \dim F}$, that is, if $f'(a)$ is surjective,

(ii) f is an immersion at a if $\mathrm{rk(f)(a) = \dim E}$, that is, if $f'(a)$ is injective.

Note that is f is either a submersion or an immersion at a, then f has constant rank in a neighborhood of a. Also note that f is both an immersion and a submersion if and only if f is a local C^k-isomorphism.

Theorem 3.2.4. *Let $E \supset U \xrightarrow{f} F$ be a map of class C^k $(k \geq 1)$ and assume that f is a submersion at a. Then there exist a vector space G of dimension $\dim E - \dim F$, an open neighborhood W of a and a C^k-isomorphism $\varphi\colon E \supset W \to \widetilde{W} \subset F \times G$ such that $f|_W = \pi \circ \varphi$ where $\pi\colon F \times G \to F$ is the projection.*

This is visualized by the commutative diagram

$$\begin{array}{ccc} E \supset W & & \\ \varphi \big\downarrow \wr & \searrow^{f} & \\ F \times G \supset \widetilde{W} & \xrightarrow[\pi]{} & \widetilde{V} \subset F. \end{array} \tag{3.9}$$

Proof. Let us choose a basis on E and a basis on F. For $x \in E$ we write $x = (x_1, \ldots, x_n)$ and we write $f = (f_1, \ldots, f_p)$. After eventually re-ordering the basis on E we may assume that the matrix

$$\frac{D(f_1, \ldots, f_p)}{D(x_1, \ldots, x_p)} := \left(\frac{\partial f_i}{\partial x_j}\right)_{\{1 \leq i,j \leq p\}} \tag{3.10}$$

is invertible. Set

$$\varphi(x_1, \ldots, x_n) = (f_1, \ldots, f_p, x_{p+1}, \ldots, x_n).$$

Then, denoting by I_{n-p} the unit $((n-p) \times (n-p))$-matrix and setting $\frac{D(f_1,\ldots,f_p)}{D(x_{p+1},\ldots,x_n)} := \left(\frac{\partial f_i}{\partial x_j}\right)_{\{1 \leq i \leq p, p+1 \leq j \leq n\}}$ we have

$$J_\varphi = \begin{pmatrix} \frac{D(f_1,\ldots,f_p)}{D(x_1,\ldots,x_p)} & \frac{D(f_1,\ldots,f_p)}{D(x_{p+1},\ldots,x_n)} \\ 0 & \mathrm{I}_{n-p} \end{pmatrix}.$$

Therefore, φ is a C^k-isomorphism of $\mathbb{R}^n$ in a neighborhood W of a by Theorem 3.1.3. Set $G = \mathbb{R}^{n-p}$ and denote by $\pi\colon \mathbb{R}^p \times \mathbb{R}^{n-p} \to \mathbb{R}^p$ the projection, we have $f = \pi \circ \varphi$. q.e.d.

Theorem 3.2.5. *Let $E \supset U \xrightarrow{f} F$ be a map of class C^k $(k \geq 1)$ and assume that f is an immersion at a. Then there exist a vector space G of dimension $\dim F - \dim E$, an open neighborhood V of $a \times \{0\}$ in $E \times G$ and a C^k-isomorphism $\psi\colon E \times G \supset V \to \widetilde{V} \subset F$ such that setting $W = (E \times \{0\}) \cap V$, $f|_W = \psi \circ \iota$ where ι is the linear embedding $E \to E \times G$, $x \mapsto (x, 0)$.*

This is visualized by the commutative diagram

$$\begin{array}{ccc} E \supset W & \xrightarrow{\iota} & V \subset E \times G \\ & \searrow_{f} & \wr \big\downarrow \psi \\ & & \widetilde{V} \subset F \end{array} \tag{3.11}$$

Proof. Let us choose a basis on E and a basis on F. For $x \in E$ we write $x = (x_1, \dots, x_n)$ and we write $f = (f_1, \dots, f_p)$. After eventually re-ordering the basis on F we may assume that the matrix

$$\frac{D(f_1, \dots, f_n)}{D(x_1, \dots, x_n)} := \left(\frac{\partial f_i}{\partial x_j}\right)_{\{1 \le i,j \le n\}} \tag{3.12}$$

is invertible. Set $G = \mathbb{R}^{p-n}$ and let $y = (y_{n+1}, \dots, y_p) \in G$. Consider the map $\psi\colon \mathbb{R}^n \times \mathbb{R}^{p-n} \to \mathbb{R}^p$, $\psi(x, y) = (f_1(x), \dots, f_n(x), y_{n+1} + f_{n+1}(x), \dots, y_p + f_p(x))$. Using notations similar as in the proof of Theorem 3.2.4, we have

$$J_\psi = \begin{pmatrix} \frac{D(f_1,\dots,f_n)}{D(x_1,\dots,x_n)} & 0 \\ \frac{D(f_{n+1},\dots,f_p)}{D(x_1,\dots,x_n)} & \mathrm{I}_{p-n} \end{pmatrix}.$$

Therefore, ψ is invertible in a neighborhood of $(a, 0)$. Moreover, setting $\iota(x) = (x, 0) \in E \times G$, we have $f = \psi \circ \iota$. q.e.d.

One can formulate both Theorems 3.2.4 and 3.2.5 in a single statement.

Theorem 3.2.6. (The constant rank Theorem.) *Let $E \supset U \xrightarrow{f} F$ be a map of class C^k $(k \ge 1)$ and assume that* rk(f)(x) $\equiv$ r *for any x in a neighborhood of a. Then there exist $\varphi\colon E \to E$ defined in a neighborhood of a and $\psi\colon F \to F$ defined in a neighborhood of $f(a)$ such that φ and ψ are C^k-isomorphisms and the map $\psi \circ f \circ \varphi^{-1}$ is linear of rank r.*

This theorem asserts that if a map f of class C^k $(k \ge 1)$ has constant rank in a neighborhood of a, then the germ of f at a is equivalent to a linear map.

Proof. We may write $F = F_1 \times F_2$ and $f = (f_1, f_2)$ such that $\dim F_1 = r$ and f_1 is a submersion at a. By Theorem 3.2.4, there exists a C^k-isomorphism $\varphi\colon E \to E$ defined in a neighborhood of a such that $f_1 \circ \varphi^{-1}$ is a linear projection. Replacing f with $f \circ \varphi^{-1}$, we may assume from the beginning that $E = \mathbb{R}^r \times \mathbb{R}^{n-r}$, $F = \mathbb{R}^r \times \mathbb{R}^{p-r}$, $f = (f_1, f_2)$, f_1 is the map $x = (x_1, \dots, x_n) \mapsto (x_1, \dots, x_r)$. Let us write for short $x = (x', x'') \in E$ with $x' = (x_1, \dots, x_r)$, $x'' = (x_{r+1}, \dots, x_n)$ and $(x', y'') \in F$ with $y'' = (y_{r+1}, \dots, y_p)$. Since f has rank r, we have $\frac{\partial f_2(x)}{\partial x_j} = 0$ for $j > r$ in a neighborhood of a. Therefore, $f_2(x) = f_2(x')$ and $f(x', x'') = (x', f_2(x'))$. Define the map $\psi\colon F \to F$, $\psi(x', y'') = (x', y'' - f_2(x'))$. Then ψ is a local isomorphism in a neighborhood of $f(a)$ and $\psi \circ f$ is the map $E \ni (x', x'') \mapsto (x', 0) \in F$. q.e.d.

3.3 Submanifolds

Again in this section, the vector spaces E, F, etc. are finite dimensional.

Recall that a subset S of a topological space X is said locally closed if it is the intersection of an open subset U and a closed subset Z of X. Of course, U and Z are not unique. This is equivalent to saying that for each $x \in S$, there exists an open neighborhood U_x of x in X such that $S \cap U_x$ is closed in U_x. For example, an open interval of $\mathbb{R}$ is locally closed in $\mathbb{C}$.

In the sequel, k is an integer ≥ 1 or $k = +\infty$.

Definition 3.3.1. Let S be a subset of a real finite vector space E. One says that S is a submanifold of class C^k of E if for each $a \in S$ there exists an open neighborhood U of a, an open subset V of E, a C^k-isomorphism $\varphi\colon U \xrightarrow{\sim} V$ such that $\varphi(S \cap U) = N \cap V$ and N is a linear subspace.

Consider two C^k-isomorphisms $\varphi_j\colon U \xrightarrow{\sim} V_j$ $(j = 1, 2)$ such that $\varphi_j(S \cap U) = N_j \cap V_j$ and N_j is a linear subspace. Then $\theta\colon \varphi_2 \circ \varphi_1^{-1}$ is a C^k-isomorphism $V_1 \xrightarrow{\sim} V_2$ and $\theta(V_1 \cap N_1) = V_2 \cap N_2$. It follows that $\dim N_1 = \dim N_2$. This allows us to set

Definition 3.3.2. Let S be a submanifold and let $a \in S$. The dimension of S at a is the dimension of the linear space N given in Definition 3.3.1. The codimension of S at a is the dimension of E minus the dimension of S at a.

The dimension of S at x is a locally constant function. One denotes it by $\dim S$. One denotes by $\mathrm{codim} S$ the codimension.

Remark 3.3.3. (a) A submanifold S of E is locally closed in E.

(b) If S is an affine space, that is, $S = a + N$ where N is a linear subspace, then S is a submanifold.

(c) An open subset U of E is a submanifold of codimension 0. A discrete subset of E is a submanifold of dimension 0.

(d) Let $\theta\colon E \supset U \xrightarrow{\sim} V \subset F$ be a C^k-isomorphism from an open subset U of E to an open subset V of F. If S is a submanifold of E, then $\theta(S)$ is a submanifold of F of the same dimension.

(e) If S is a sumanifold of E and Z is a submanifold of F, then $S \times Z$ is a submanifold of $E \times F$.

We shall give many examples of submanifolds below.

Construction by submersion

Theorem 3.3.4. *Let $E \supset U \xrightarrow{f} F$ be a map of class C^k with $k \geq 1$. Let $b \in f(U)$ and set $S = f^{-1}(b)$. Assume that f is a submersion at each $x \in S$. Then S is a submanifold and* $\operatorname{codim} S = \dim F$.

Proof. The problem is local on S, that is, we have to prove that each $x \in S$ admits an open neighborhood for which the conclusion of the statement holds. Hence, applying Theorem 3.2.4, we may write $f = \pi \circ \varphi$ where φ is a C^k-isomorphism and π is linear and surjective. By Remark 3.3.3 (iv), we may assume from the beginning that f is linear and surjective. Then the result is clear. q.e.d.

Example 3.3.5. (i) The n-sphere $\mathbb{S}^n := \{x = (x_0, \dots, x_n) \in \mathbb{R}^{n+1}; \sum_{j=0}^{n} x_j^2 = 1\}$ is a submanifold of dimension n and of class C^∞.
(ii) Let $E \supset U \xrightarrow{f} F$ be a map of class C^k and denote by Γ_f the graph of f:

$$\Gamma_f = \{(x, y) \in U \times F; y = f(x)\}.$$

Then $\Gamma_f = \{(x, y) \in U \times F; g(x, y) = 0\}$ where $g\colon U \times F \to F$ is the map $g(x, y) = y - f(x)$. Then rk(g)(x, y) = dim F at each (x, y) which shows that g is a submersion. Hence Γ_f is a submanifold and its dimension is that of E.
(iii) The surface $\{(x, y, z) \in \mathbb{R}^3; z = xy\}$ is a submanifold since it is a graph. The change of variable $(x, y) \mapsto (x + y, x - y)$ interchanges it with the submanifold $\{(x, y, z) \in \mathbb{R}^3; z = x^2 - y^2\}$ which is easier to draw.

Construction by immersion

Theorem 3.3.6. *Let $E \supset U \xrightarrow{f} F$ be a map of class C^k with $k \geq 1$. Let $a \in U$ and assume that f is an immersion at a. Then there exists an open neighborhood V of a in U such that $f(V)$ is a submanifold of F. Moreover,* $\dim f(V) = \dim E$.

Proof. Applying Theorem 3.2.5 and Remark 3.3.3 (vi), we may assume that $f = \psi \circ \iota$ where ι is linear and injective. By Remark 3.3.3 (iv), we may assume from the beginning that f is linear and injective. Then the result is clear. q.e.d.

Remark 3.3.7. The map f may be an immersion at each $x \in U$ without being injective. In such a case, $f(U)$ will not be a submanifold. Moreover, even if f is an immersion and is injective, $f(U)$ may not be a submanifold. However, if $f\colon U \to V \subset F$ is an immersion, is injective and is proper (the inverse image by f of any compact set of V is compact in U), then one can show that $f(U)$ is a submanifold in this case.

Example 3.3.8. (i) One can also define the circle $\mathbb{S}^1$ as the image of $\mathbb{R}$ in $\mathbb{R}^2$ by the map $t \mapsto (\cos t, \sin t)$. For each $(x, y) \in \mathbb{S}^1 \subset \mathbb{R}^2$, there exists an open neighborhood U of (x, y) and an open interval $I \subset \mathbb{R}$ such that $U \cap \mathbb{S}^1 = f(I)$. Therefore, $\mathbb{S}^1$ is a submanifold of dimension 1.
(ii) The torus $\mathbb{T}^2 = \mathbb{S}^1 \times \mathbb{S}^1 \subset \mathbb{R}^3$ is the image of $\mathbb{R}^2$ by the map

$$(\varphi, \theta) \mapsto \quad \begin{cases} (R + r\cos\theta)\cos\varphi, \\ (R + r\cos\theta)\sin\varphi, \\ r\sin\theta. \end{cases}$$

Here $0 < r < R$. One proves as for the circle that $\mathbb{T}^2$ is a submanifold of dimension 2.
(iii) A graph $\Gamma_f \subset U \times F$, as in Example 3.3.5 (ii), may also be obtained by the immersion $E \supset U \hookrightarrow U \times F$, $x \mapsto (x, f(x))$.

3.4 Tangent space

Definition 3.4.1. Let E be a finite dimensional real vector space and let S be a submanifold. Let $a \in S$. A vector $v \in E$ is tangent to S at a if there exists a C^1-map $\gamma\colon I \to E$, where I is an open interval of $\mathbb{R}$ and $t_0 \in I$ such that $\gamma(t) \in S$ for all $t \in I$, $\gamma(t_0) = a$ and $\gamma'(t_0) = v$.

One denotes by T_aS the subset of E consisting in vectors tangent to S at a and calls it the tangent vector space to S at a.

Intuitively, one can think to γ as an object moving along S. Then a tangent vector v to S at a is the speed at time $t = t_0$ of this object, assuming it is at a at time t_0.

Lemma 3.4.2. *Let $E \supset U \xrightarrow{f} V \subset F$ be a map of class C^1. Let $S \subset U$ and $Z \subset V$ be submanifolds and assume that $f(S) \subset Z$. Then $f'(a) \cdot T_aS \subset T_{f(a)}Z$.*

Proof. Let $v \in T_aS$. There exists an open interval I, $t_0 \in I$ and a C^1-map $\gamma\colon I \to E$ such that $\gamma(I) \subset S$, $\gamma(t_0) = a$ and $\gamma'(t_0) = v$. Then $(f \circ \gamma)(I) \subset Z$, $(f \circ \gamma)(t_0) = f(a)$ and $f'(a) \cdot v = (f \circ \gamma)'(t_0)$. q.e.d.

Proposition 3.4.3. *Let S be a submanifold of E and let $a \in S$. Then T_aS is a linear subspace of E of dimension* $\dim S$.

Proof. There exists an open neighborhood W of a in U, an open subset $V \subset E$, a linear subspace N of E and a C^1-isomorphism $f\colon W \xrightarrow{\sim} V$ such that $f\colon S \cap W \xrightarrow{\sim} V \cap N$. Applying Lemma 3.4.2, we get that $f'(a)$ induces a bijection $T_aS \xrightarrow{\sim} T_{f(a)}N$. Hence, it is enough to check the result when $S = N$. In this case, this is clear since any vector $v \in N$ is the derivative at $t = 0$ of a C^1-map $\gamma\colon \mathbb{R} \to N$, namely the map $\gamma(t) = t \cdot v$. q.e.d.

Definition 3.4.4. The affine tangent space to S at a is the affine space $a + T_aS$.

Tangent space for submersions

Proposition 3.4.5. *Let S be a submanifold of E and let $a \in S$. Assume that there exists a submersion $E \supset U \xrightarrow{f} F$ of class C^1 and $b \in f(U)$ such that $S = f^{-1}(b)$. Then $T_aS = \operatorname{Ker} f'(a)$.*

Proof. Since $f(S) \subset \{b\}$, $f'(a)T_aS \subset \{0\}$ by Lemma 3.4.2. Hence, $T_aS \subset \operatorname{Ker} f'(a)$. Since both spaces have the same dimension, the result follows. q.e.d.

If $F = \mathbb{R}^p$ and $f = (f_1, \dots, f_p)$, then the affine tangent space $a+T_aS$ is given by the equations

$$a + T_aS = \{x \in E; \langle x - a, f_j'(a)\rangle = 0,\ 1 \le j \le p\}.$$

Example 3.4.6. Let

$$S = \{(x, y, z) \in \mathbb{R}^3; z^2 - x^2 - y^2 = 0, 2x + y + 1 = 0\}.$$

Consider the map $f\colon \mathbb{R}^3 \to \mathbb{R}^2$, $(x, y, z) \mapsto (z^2 - x^2 - y^2, 2x + y + 1)$, so that $S = f^{-1}(0)$. The Jacobian matrix of f is given by

$$J_f(x, y, z) \;=\; \begin{pmatrix} -2x & -2y & 2z \\ 2 & 1 & 0 \end{pmatrix}$$

Hence f is a submersion at each $(x, y, z) \in S$ and S is a submanifold of dimension 1.

Let $a = (-1, 2, \sqrt{5}) \in S$ and let us calculate the affine tangent space to S at a.

The tangent vector space to S at a is the kernel of $J_f(a)$, that is, the kernel of $\begin{pmatrix} 2 & -4 & 2\sqrt{5} \\ 2 & 1 & 0 \end{pmatrix}$. The tangent vector space is the line

$$\Delta = \{(x, y, z) \in \mathbb{R}^3; x - 2y + \sqrt{5}z = 0, 2x + y = 0\}.$$

The affine tangent space is $a + \Delta$, that is the set

$$\begin{aligned} a + \Delta = \{(x, y, z) \in \mathbb{R}^3;\ & (x + 1) - 2(y - 2) + \sqrt{5}(z - \sqrt{5}) = 0, \\ & 2(x + 1) + (y - 2) = 0\}. \end{aligned}$$

Tangent space for immersions

Proposition 3.4.7. *Let $E \supset U \xrightarrow{f} F$ be a map of class C^1, $a \in U$ and assume that f is an immersion at a. Let V be an open neighborhood of a in U such that $f(V)$ is a submanifold of F. Set $Z = f(V)$ and $b = f(a)$. Then $T_bZ = \operatorname{Im} f'(a)$.*

Proof. The tangent space to V is E. Since $f(V) \subset Z$, $f'(a)(E) \subset T_bZ$ by Lemma 3.4.2. Since both spaces have the same dimension, the result follows. q.e.d.

Applications to surfaces

Let $f : U \to F$ be a real function defined on an open set U of $\mathbb{R}^2$ with values in a normed space F (*e.g.*, $F = \mathbb{R}^3$). Let us write

$$f(x, y) = f(x_0, y_0) + x f'_x(x_0, y_0) + y f'_y(x_0, y_0) + \cdots$$

Set $\vec{\alpha} = f'_x(x_0, y_0)$ and $\vec{\beta} = f'_y(x_0, y_0)$.

(a) First assume that $\vec{\alpha}$ and $\vec{\beta}$ are linearly independant. Consider the affine plane

$$P = \{f(x_0, y_0) + \lambda\vec{\alpha} + \mu\vec{\beta}\}.$$

Then, for U a sufficiently small neighborhood of (x_0, y_0), P will be tangent to the surface $f(U)$.

(b) Now assume $\vec{\alpha} \neq 0$ but $\vec{\beta}$ is proportional to $\vec{\alpha}$. Then $\vec{\alpha}$ generates a line to which the surface $f(U)$ will be "tangent". For example, consider the surface in $\mathbb{R}^3$

$$(x, y) \mapsto (x^2, x^3, y + x^2) \text{ at } (0, 0).$$

Denoting by (u, v, w) the coordinates on $\mathbb{R}^3$, this surface is also defined by the equation $v = \pm|u|^{\frac{3}{2}}$. The w-axis is "tangent" to the surface. Note that f is no more an immersion, and the surface is not a submanifold. (It has "singularities".)

(c) Assume that $\vec{\alpha} = \vec{\beta} = 0$ but the three vectors $\vec{e} = f''_{xx}$, $\vec{f} = f''_{yy}$ and $\vec{g} = f''_{xy}$ are linearly independent. Then the surface $f(U)$ will be "tangent" to the cone

$$\gamma = f(x_0, y_0) + x^2\vec{e} + y^2\vec{f} + 2xy\vec{g}.$$

To check that γ is a cone, denote by (u, v, w) the coordinates on $\mathbb{R}^3$ in the basis $\vec{e}, \vec{f}, \vec{g}$ and assume for short that $f(x_0, y_0) = 0$. Then $\gamma = \{(u, v, w); w^2 =$

$4uv\}$. Setting $u = X + Y$, $v = X - Y$, $2Z = w$, we get $\gamma = \{(X, Y, Z); Z^2 = X^2 - Y^2\}$. Recall that the cone γ is not a C^0-submanifold in a neighborhood of $0 \in \mathbb{R}^3$ since $\gamma \setminus \{0\}$ has two connected components. A fortiori, γ is not a C^1-submanifold.

Applications to problems of extremum

We can now give another formulation of the the Lagrange's multipliers Theorem 1.2.4 and a complete proof.

Theorem 3.4.8. (The Lagrange's multipliers Theorem.) *Let E be a finite dimensional vector space, U an open subset and let $a \in U$. Let $f\colon E \supset U \to \mathbb{R}$ be a map and assume that $f'(a)$ exists. Let $S \subset U$ be a closed submanifold with $a \in S$. Assume that $f|_S$ has a local extremum at a. Then the restriction of $f'(a)$ to the tangent space T_aS vanishes, that is, $f'(a)|_{T_aS} = 0$.*

Proof. We may assume that there exist an open neighborhood V of a and a submersion $g\colon V \to \mathbb{R}^d$ such that $S \cap V = g^{-1}(0)$. By Theorem 3.2.4 (with other notations), there exists a local isomorphism $\psi\colon E \to E$ defined in a neighborhood of $b := \psi^{-1}(a)$ such that $g \circ \psi$ is linear. Set $N = \{x \in E; (g \circ \psi)(x) = 0\}$. Then $f \circ \psi$ has a local extremum on N at b. Let us write $g = (\varphi_1, \dots, \varphi_d)$. Applying Corollary 1.2.3, we find scalars $(\lambda_1, \dots, \lambda_p)$ such that

$$(f \circ \psi)'(b) = \sum_{j=1}^{d} \lambda_j \cdot (\varphi_j \circ \psi)'(b),$$

that is

$$f'(a) \circ \psi'(b) = \sum_{j=1}^{d} \lambda_j \cdot \varphi_j'(a) \circ \psi'(b).$$

Since the linear map $\psi'(b)$ is invertible, we get

$$f'(a) = \sum_{j=1}^{d} \lambda_j \cdot \varphi_j'(a),$$

which is equivalent to saying that $f'(a)|_{T_aS} = 0$. q.e.d.

One can also extend Theorem 2.3.4 to submanifolds.

Theorem 3.4.9. *Let E be a finite dimensional vector space, U an open subset and let $a \in U$. Let $f\colon E \supset U \to \mathbb{R}$ be a map. Assume that $f'(a)$ and $f''(a)$ exist and $f'(a) = 0$. Let $S \subset U$ be a closed submanifold with $a \in S$.*

(i) *Assume that $f''(a)|_{T_aS}$ is positive definite. Then $f|_S$ has a local strict minimum at a.*

(ii) *Assume that $f|_S$ has a local minimum at a. Then $f''(a)|_{T_aS}$ is positive semidefinite.*

Proof. The proof goes as for Theorem 3.4.8. We may assume that there exist an open neighborhood V of a and a submersion $g\colon V \to \mathbb{R}^d$ such that $S \cap V = g^{-1}(0)$. By Theorem 3.2.4 (with other notations), there exists a local isomorphism $\psi\colon E \to E$ defined in a neighborhood of $b := \psi^{-1}(a)$ such that $g \circ \psi$ is linear. Set $N = \{x \in E; (g \circ \psi)(x) = 0\}$. Then f has a local (resp. strict local) minimum at a if and only if $f \circ \psi$ has a local (resp. strict local) minimum on N at b.

On the other hand, $f''(a)|_{T_aS}$ is positive definite (resp. semidefinite) if and only if $(f \circ \psi)''(b)|_N$ is positive definite (resp. semidefinite). Indeed, Proposition 2.1.9 gives

$$(f \circ \psi)''(b) = f''(a) \circ \psi'(b) + f'(a) \circ \psi''(b)$$

and the hypothesis that $f'(a) = 0$ gives

$$(f \circ \psi)''(b) = f''(a) \circ \psi'(b).$$

To conclude, we apply Theorem 2.3.4 on N. q.e.d.

One shall be aware that the condition that $f'(a) = 0$ cannot be weaken by only assuming that $f'(a)|_{T_aS} = 0$. For example, consider $E = \mathbb{R}^2$, $a = (0,0)$, $S = \{(x,y) \in \mathbb{R}^2; y - x^2 = 0\}$ and $f(x,y) = x^2 - y$. Then $T_aS = \{(x,y); y = 0\}$, $f'(a)|_{T_aS} = 0$, $f''(a)|_{T_aS}(x,x) = x^2$ is positive definite, but f has not a strict minimum at a on S.

Exercises to Chapter 3

Exercise 3.1. (This exercise is a continuation of Exercise 1.2.) Consider the map $f = (f_1, f_2)\colon \mathbb{R}^2 \to \mathbb{R}^2$ given by:

$$f_1(x,y) = x, \qquad f_2(x,y) = \begin{cases} 0 \text{ si } (x,y) = (0,0) \\ y\dfrac{y^2 - x^4}{y^2 + x^4} \text{ sinon.} \end{cases}$$

(i) Show that $2x^2|y| \le x^4 + y^2$ then that $|y - f_2(x,y)| \le x^2$.
(ii) Deduce that f_2 and f are differentiable at each $(x,y) \in \mathbb{R}^2$ and calculate the differential of f at $(0,0)$.

(iii) Show that for any neighborhood U of $(0,0)$ in $\mathbb{R}^2$, there exist $a, b \in U$ with $a \neq b$ and $f(a) = f(b)$.
(iv) Is the function f of class C^1 ?

Exercise 3.2. (i) Show that the equation

$$u^3 + 2u + \exp(u - x - y^2) = \cos(x - y + u)$$

admits one and only one solution u for any given $(x, y) \in \mathbb{R}^2$.
(ii) Show that the function $u(x, y)$ defined by (i) is of class C^∞.
(iii) Calculate the Taylor series of $u(x, y)$ up to order 2 in a neighborhood of $(x, y) = (0, 0)$.

Exercise 3.3. Let $J = [-1, +1]$ and denote as usual by $\mathcal{C}^0(J, \mathbb{R})$ the Banach space of real valued continuous functions for the norm of uniform convergence $||f||_0 = \sup_{t \in J} |f(t)|$. Let us say that $f \in C^0(J)$ is $\mathcal{C}^1$ on J if f is C^1 on $]-1, +1[$ and f' admits a limit at -1 and at 1. Still denote by f' the function so defined on J and denote by $\mathcal{C}^1(J, \mathbb{R})$ the vector subspace of $\mathcal{C}^0(J, \mathbb{R})$ consisting of functions of class C^1 on J. One endows $\mathcal{C}^1(J)$ with the norm

$$\|f\|_1 = \|f\|_0 + \|f'\|_0.$$

Denote by $\mathcal{C}_0^1(J)$ the vector subspace of $\mathcal{C}^1(J)$ consisting of functions f satisfying$f(0) = 0$ and let us endow $\mathcal{C}_0^1(J)$ with the norm induced by that $\mathcal{C}^1(J)$.
(i) Show that $\mathcal{C}^1(J)$ and $\mathcal{C}_0^1(J)$ are Banach spaces (use Theorem 1.3.10).
(ii) Denote by $D\colon \mathcal{C}_0^1(J) \to \mathcal{C}^0(J)$ the linear map $f \mapsto f'$. Show that:
(a) $\|D(f)\|_0 \leq \|f\|_1$ for all $f \in \mathcal{C}_0^1(J)$,
(b) D is bijective,
(c) $|f(t)| \leq |t| \|D(f)\|_0$ for all $f \in \mathcal{C}_0^1(J)$ and all $t \in J$,
(d) D is an isomorphism of Banach spaces from $\mathcal{C}_0^1(J)$ onto $\mathcal{C}^0(J)$.
(iii) Define $\Phi\colon \mathcal{C}_0^1(J) \to \mathcal{C}^0(J)$ by setting for $f \in \mathcal{C}_0^1(J)$, $\Phi(f) = Df - f^2$. Show that Φ is C^1, calculate its differential and show that $\Phi'(0)$ is an isomorphism.
(iv) For $r > 0$, denote as usual by $B(0; r)$ the open ball of center 0 and radius r in $\mathcal{C}^0(J)$. Show that there exist $r > 0$ and a function $\Psi\colon B(0; r) \to \mathcal{C}_0^1(J)$ of class $\mathcal{C}^1$ such that for all $g \in B(0; r)$, $\Psi(g)$ is a solution of the differential equation $y' = y^2 + g$.

Exercise 3.4. Let $S = \{(x, y, z) \in \mathbb{R}^3; x^2 + y^2 - z^2 - 2x = 0\}$.
(i) Prove that S is a submanifold of class C^∞ of $\mathbb{R}^3$.
(ii) Calculate the equations of the affine tangent planes to S which are parallel to the 3 planes of equations $x = 0$, $y = 0$ and $z = 0$ respectively.

Exercise 3.5. Let a, b, c, R be strictly positive real numbers. Consider the two surfaces

$$S = \{(x,y,z) \in \mathbb{R}^3; x^2 + y^2 + (z - \frac{b^2+c^2}{c})^2 = R^2\},$$
$$\Gamma = \{(x,y,z) \in \mathbb{R}^3; \frac{x^2}{a^2} + \frac{y^2}{b^2} = \frac{z^2}{c^2}\}.$$

Calculate R in order that these two surfaces are tangent.

Exercise 3.6. Find the equation of an affine tangent plane to the ellipsoid

$$S = \{(x,y,z) \in \mathbb{R}^3; \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1\}$$

which intersects the coordinates axis along intervals of the same length.

Exercise 3.7. Consider the map $f\colon \mathbb{R}^4 \to \mathbb{R}^2$

$$(x,y,z,t) \mapsto (u = x^2 + 2y^2 + 3z^2 + 4t^2 - 4, v = x + y - z - t)$$

and let $S = \{(x,y,z,t) \in \mathbb{R}^4; f(x,y,z,t) = 0\}$.
(i) Prove that S is a compact submanifold of dimension 2.
(ii) Calculate the affine tangent plane $a + T_aS$ at $a = (1,0,1,0)$.
(iii) Prove that the equation $f(x,y,z,t) = 0$ in a neighborhood of a is equivalent to $(z,t) = g(x,y)$ for a C^∞-function g with values in $\mathbb{R}^2$ and calculate $\frac{\partial z}{\partial x}$, $\frac{\partial z}{\partial y}$, $\frac{\partial t}{\partial x}$ and $\frac{\partial t}{\partial y}$.