

# Calcul différentiel

(licence de mathématiques)

Sylvie Benzoni

30 avril 2008



# Chapitre I

## Différentielle d'une fonction

### 1 Introduction

On considère deux  $\mathbb{R}$ -espaces vectoriels normés  $E$  et  $F$ , que l'on suppose **complets** : on dit aussi que ce sont des espaces de **Banach**. On notera en général  $\|\cdot\|_E$  et  $\|\cdot\|_F$  leurs normes respectives (et simplement  $\|\cdot\|$  lorsque  $E = F$  et qu'il n'y a pas d'ambiguïté possible.)

Bien souvent, ces espaces seront en fait de dimension finie. Ce seront même simplement des espaces  $\mathbb{R}^n$ , munis d'une quelconque norme  $\|\cdot\|_k$  pour  $k \in [1, \dots, \infty]$ , définie par

$$\|x\|_k = \left( \sum_{j=1}^n |x_j|^k \right)^{1/k} \quad \text{si } k < \infty \quad \text{et} \quad \|x\|_\infty = \sup_{1 \leq j \leq n} |x_j|,$$

où  $x_1, \dots, x_n$  désignent les composantes de  $x \in \mathbb{R}^n$ . Ils pourront éventuellement être des versions déguisées de  $\mathbb{R}^n$ , comme l'espace des matrices  $\mathcal{M}_{p,q}(\mathbb{R})$  à  $p$  lignes et  $q$  colonnes, que l'on peut munir de diverses normes. Par exemple, on peut prendre la norme subordonnée à la norme  $\|\cdot\|_k$  dans  $\mathbb{R}^q$  et  $\mathbb{R}^p$ , définie par

$$\|M\|_k = \sup_{x \in \mathbb{R}^q, x \neq 0} \frac{\|Mx\|_k}{\|x\|_k}.$$

**Exercice.** Exprimer cette norme à l'aide des coefficients de  $M$  lorsque  $p = q$  et  $k = 1, 2, \infty$ . On rappelle que de toutes façons, toutes les normes sont équivalentes en dimension finie.

On notera  $\mathcal{L}(E; F)$  l'espace des applications **linéaires continus** de  $E$  dans  $F$ , muni de la norme

$$\|\ell\| = \sup_{x \in E \setminus \{0\}} \frac{\|\ell(x)\|_F}{\|x\|_E}.$$

C'est un espace de Banach (**exercice**).

On suppose que  $U$  est un ouvert (non vide !) de  $E$ , et on considère une fonction

$$f : U \rightarrow F.$$

**Définition I.1** On dit que la fonction  $f$  est **différentiable** en un point  $x \in U$  si elle est **continue** au point  $x$  et s'il existe  $\ell \in \mathcal{L}(E; F)$  tel que

$$(1.1) \quad \lim_{h \xrightarrow{\neq} 0_E} \frac{\|f(x+h) - f(x) - \ell(h)\|_F}{\|h\|_E} = 0.$$

Cette définition est (volontairement) redondante : en fait, l'existence d'une application linéaire *continue*  $\ell$  telle qu'on ait (1.1) impose à  $f$  d'être continue en  $x$ . Car, en notant

$$\varepsilon(h) = \|f(x+h) - f(x) - \ell(h)\|_F / \|h\|_E \text{ si } h \neq 0_E \text{ et } \varepsilon(0) = 0,$$

on a par l'inégalité triangulaire

$$\|f(x+h) - f(x)\|_F \leq \varepsilon(h) \|h\|_E + \|\ell(h)\|_F,$$

où le membre de droite tend vers 0 lorsque  $h$  tend vers  $0_E$  grâce à (1.1) et à la continuité de  $\ell$  en  $0_E$ . Inversement, si on suppose  $f$  continue en  $x$  et s'il existe  $\ell$  linéaire telle qu'on ait (1.1), alors par l'inégalité triangulaire

$$\|\ell(h)\|_F \leq \varepsilon(h) \|h\|_E + \|f(x+h) - f(x)\|$$

tend vers 0 lorsque  $h$  tend vers  $0_E$ . (Et on rappelle que la continuité d'une application **linéaire** équivaut à sa continuité en 0.)

**Notation.** Bien sûr, l'application  $\ell$  dans la définition I.1 dépend du point  $x$ . On la notera désormais  $\ell = df_x$ , de sorte que (1.1) se réécrit

$$\lim_{h \xrightarrow{\neq} 0_E} \frac{\|f(x+h) - f(x) - df_x(h)\|_F}{\|h\|_E} = 0.$$

L'application  $df_x$  est appelée **différentielle** de  $f$  au point  $x$ .

**Définition I.2** On dit que la fonction  $f$  est **différentiable** sur  $U$  si elle est différentiable en tout point  $x \in U$ . Dans ce cas, on appelle **différentielle** de  $f$  la fonction

$$\begin{aligned} df : U &\rightarrow \mathcal{L}(E; F) . \\ x &\mapsto df_x \end{aligned}$$

Si de plus  $df$  est continue, on dit que  $f$  est **continûment différentiable**, ou de façon équivalente que  $f$  est de **classe**  $\mathcal{C}^1$ .

## 2 Premiers exemples

- Toute application **constante** est continûment différentiable, de **différentielle nulle**.
- Toute application **linéaire continue**  $\ell$  est continûment différentiable, et sa **différentielle** est **constante**, égale à  $\ell$  en tout point :

$$d\ell_x(h) = \ell(h) \text{ quels que soient } x \text{ et } h \in E .$$

• Plus généralement, toute application **multi-linéaire continue** est continûment différentiable. Ce dernier point demande quelques explications. Si  $E_1, \dots, E_n$  sont des espaces de Banach, le produit cartésien  $E = E_1 \times \dots \times E_n$ , muni de

$$\|(x_1, \dots, x_n)\|_E = \|x_1\|_{E_1} + \dots + \|x_n\|_{E_n},$$

est aussi un espace de Banach. On dit qu'une application  $\phi : E \rightarrow F$  est  **$n$ -linéaire** (et lorsqu'on ne veut pas préciser  $n$  on dit **multi-linéaire**) si pour tout  $j \in \{1, \dots, n\}$  et pour tout

$(x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n) \in E_1 \times \dots \times E_{j-1} \times E_{j+1} \times \dots \times E_n$  (avec les conventions naturelles lorsque  $j = 1$  ou  $j = n$ ), l'application partielle

$$x \in E_j \mapsto \phi(x_1, \dots, x_{j-1}, x, x_{j+1}, \dots, x_n)$$

est **linéaire**. Lorsque  $\phi$  est de plus continue, toutes les applications partielles sont continues, et  $\phi$  est continûment différentiable, de différentielle définie par :

$$(2.2) \quad d\phi_{(x_1, \dots, x_n)}(h_1, \dots, h_n) = \sum_{j=1}^n \phi(x_1, \dots, x_{j-1}, h_j, x_{j+1}, \dots, x_n).$$

(Dans cette écriture abusive, le premier terme de la somme est évidemment  $\phi(h_1, x_2, \dots, x_n)$  et le dernier  $\phi(x_1, \dots, x_{n-1}, h_n)$ .) La démonstration utilise le résultat suivant.

**Lemme I.1** *Si  $\phi$  est une application  $n$ -linéaire continue, il existe  $C \in \mathbb{R}^{+*}$  tel que*

$$(2.3) \quad \|\phi(h_1, \dots, h_n)\|_F \leq C \prod_{j=1}^n \|h_j\|_{E_j}$$

quels que soient les vecteurs  $h_j \in E_j$ .

**Dém.** D'après la continuité de  $\phi$  en  $0_E = (0_{E_1}, \dots, 0_{E_n})$ , il existe  $\eta > 0$  tel que si  $\max \|k_j\|_{E_j} \leq \eta$ ,

$$\|\phi(k_1, \dots, k_n)\|_F \leq 1.$$

Soient  $h_1 \in E_1, \dots, h_n \in E_n$ . Si l'un de ces vecteurs est nul,  $\phi(h_1, \dots, h_n) = 0_F$ . Sinon, d'après la multilinéarité de  $\phi$  on a

$$\phi(h_1, \dots, h_n) = \frac{1}{\eta^n} \left( \prod_{j=1}^n \|h_j\|_{E_j} \right) \phi\left(\frac{\eta}{\|h_1\|} h_1, \dots, \frac{\eta}{\|h_n\|} h_n\right).$$

On en déduit l'inégalité (2.3) avec  $C = 1/\eta^n$  (inégalité trivialement satisfaite lorsque l'un des vecteurs  $h_j$  est nul).  $\square$

**Remarque.** Réciproquement, une application  $n$ -linéaire  $\phi$  satisfaisant (2.3) est continue sur  $E = E_1 \times \dots \times E_n$ . De plus, l'ensemble des applications  $n$ -linéaires continues, noté  $\mathcal{L}(E_1, \dots, E_n; F)$  (attention à ne pas le confondre avec l'ensemble  $\mathcal{L}(E_1 \times \dots \times E_n; F)$  des applications linéaires sur  $E$ ), muni de

$$\|\phi\|_{\mathcal{L}(E_1, \dots, E_n; F)} = \max\{\|\phi(h_1, \dots, h_n)\|_F; \|h_j\|_{E_j} \leq 1\}$$

est un espace de Banach.

**Démonstration de (2.2).** (Le cas  $n = 1$  est celui des applications linéaires continues !) Pour comprendre ce qui se passe avec  $n \geq 2$ , on peut commencer par traiter le cas  $n = 2$ . Pour une application  $\phi$  bilinéaire :

$$\phi(x_1 + h_1, x_2 + h_2) - \phi(x_1, x_2) - \phi(h_1, x_2) - \phi(x_1, h_2) = \phi(h_1, h_2).$$

Or d'après le lemme I.1, il existe  $C > 0$  tel que

$$\|\phi(h_1, h_2)\| \leq C \|h_1\|_{E_1} \|h_2\|_{E_2} \leq \frac{C}{2} \|(h_1, h_2)\|_E^2.$$

Donc on a bien

$$\lim_{(h_1, h_2) \xrightarrow{E} (0_{E_1}, 0_{E_2})} \frac{\|\phi(x_1 + h_1, x_2 + h_2) - \phi(x_1, x_2) - \phi(h_1, x_2) - \phi(x_1, h_2)\|_F}{\|(h_1, h_2)\|_E} = 0.$$

De plus, la différentielle

$$\begin{aligned} d\phi : E_1 \times E_2 &\rightarrow \mathcal{L}(E_1 \times E_2; F) \\ (x_1, x_2) &\mapsto d\phi_{(x_1, x_2)} : (h_1, h_2) \mapsto \phi(h_1, x_2) + \phi(x_1, h_2) \end{aligned}$$

est bien continue (ceci est laissé en exercice).

Le cas général se traite par récurrence sur  $n$ . Admettons le résultat pour les applications  $n$ -linéaires, et considérons une application  $(n + 1)$ -linéaire  $\phi$ . Alors

$$\begin{aligned} &\phi(x_1 + h_1, \dots, x_{n+1} + h_{n+1}) - \phi(x_1, \dots, x_{n+1}) - \sum_{j=1}^{n+1} \phi(x_1, \dots, x_{j-1}, h_j, x_{j+1}, \dots, x_{n+1}) = \\ &\phi(x_1 + h_1, \dots, x_n + h_n, x_{n+1} + h_{n+1}) - \phi(x_1 + h_1, \dots, x_n + h_n, x_{n+1}) - \phi(x_1, \dots, x_n, h_{n+1}) \\ &+ \phi(x_1 + h_1, \dots, x_n + h_n, x_{n+1}) - \phi(x_1, \dots, x_{n+1}) - \sum_{j=1}^n \phi(x_1, \dots, x_{j-1}, h_j, x_{j+1}, \dots, x_{n+1}). \end{aligned}$$

Lorsqu'on divise cette égalité par  $\|(h_1, \dots, h_n, h_{n+1})\|_{E_1 \times \dots \times E_{n+1}}$ , qui par définition est supérieur à  $\|(h_1, \dots, h_n)\|_{E_1 \times \dots \times E_n}$ , le deuxième morceau de droite tend clairement vers 0, grâce à l'hypothèse de récurrence appliquée à l'application  $n$ -linéaire

$$(x_1, \dots, x_n) \mapsto \phi(x_1, \dots, x_n, x_{n+1}).$$

Le premier morceau se simplifie en :

$$\phi(x_1 + h_1, \dots, x_n + h_n, h_{n+1}) - \phi(x_1, \dots, x_n, h_{n+1}).$$

Or on peut montrer (ceci est laissé en exercice) que l'application

$$\begin{aligned} E_1 \times \dots \times E_n &\rightarrow \mathcal{L}(E_{n+1}; F) \\ (y_1, \dots, y_n) &\mapsto \phi(y_1, \dots, y_n, \cdot) \end{aligned}$$

(où  $\phi(y_1, \dots, y_n, \cdot)$  désigne l'application linéaire  $h_{n+1} \mapsto \phi(y_1, \dots, y_n, h_{n+1})$ ) est continue. Donc pour tout  $\varepsilon > 0$ , il existe  $\eta > 0$  tel que  $\max_{j \in \{1, \dots, n\}} \|h_j\|_{E_j} \leq \eta$  entraîne

$$\|\phi(x_1 + h_1, \dots, x_n + h_n, h_{n+1}) - \phi(x_1, \dots, x_n, h_{n+1})\|_F \leq \varepsilon \|h_{n+1}\|_{E_{n+1}}.$$

Finalement, on conclut par l'inégalité triangulaire que le rapport

$$\frac{\left\| \phi(x_1 + h_1, \dots, x_{n+1} + h_{n+1}) - \phi(x_1, \dots, x_{n+1}) - \sum_{j=1}^{n+1} \phi(x_1, \dots, x_{j-1}, h_j, x_{j+1}, \dots, x_{n+1}) \right\|_F}{\|(h_1, \dots, h_{n+1})\|_E}$$

tend bien vers 0 lorsque  $(h_1, \dots, h_{n+1})$  tend vers  $(0_{E_1}, \dots, 0_{E_n})$ . □

- Une fonction de variable réelle est **différentiable** si et seulement si elle est **dérivable** : la différentielle d'une fonction dérivable  $g : U \subset \mathbb{R} \rightarrow F$  est donnée par

$$dg_t(k) = k g'(t) \quad \text{quels que soient } t \in U \text{ et } k \in \mathbb{R}.$$

(Noter que  $k \in \mathbb{R}$  est un scalaire et  $g'(t) \in F$  est un vecteur en général.)

- « Inversement », quel que soit l'espace de Banach  $E$ , si  $f : U \subset E \rightarrow F$  est différentiable, alors quels que soient  $x \in U$  et  $h \in E$ , la fonction

$$g : t \mapsto g(t) := f(x + th)$$

est dérivable en  $t = 0$ , et

$$g'(0) = df_x(h).$$

On dit que c'est la **dérivée de  $f$  dans la direction  $h$**  (si  $h$  est non nul).

- **Fonctions à valeurs dans un espace produit** : une fonction

$$\begin{aligned} f : U &\rightarrow F = F_1 \times \cdots \times F_n \\ x &\mapsto f(x) = (f_1(x), \dots, f_n(x)) \end{aligned}$$

est différentiable (en  $x$ ) si et seulement si les fonctions  $f_1, \dots, f_n$  le sont. Si c'est le cas, sa différentielle est donnée par

$$df_x(h) = (d(f_1)_x(h), \dots, d(f_n)_x(h)).$$

- **Fonctions définies sur un espace produit** : la situation est un peu plus délicate. Si

$$\begin{aligned} f : U \subset E_1 \times \cdots \times E_n &\rightarrow F \\ x = (x_1, \dots, x_n) &\mapsto f(x) \end{aligned}$$

est différentiable, alors les **applications partielles**  $x_i \mapsto (x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n)$  sont différentiables, et si on note  $d_i f$  leurs différentielles, on a

$$df_{(x_1, \dots, x_n)}(h_1, \dots, h_n) = \sum_{i=1}^n d_i f_{(x_1, \dots, x_n)}(h_i).$$

Mais la différentiabilité des applications partielles n'implique pas nécessairement la différentiabilité de  $f$  ! On verra au chapitre II (voir aussi le théorème I.2 lorsque  $E_1 = \cdots = E_n = \mathbb{R}$ ) que des applications partielles continûment différentiables impliquent que  $f$  est continûment différentiable.

### 3 Premières propriétés

**Proposition I.1** Si  $f : U \subset E \rightarrow F$  et  $g : V \subset E \rightarrow F$  sont différentiables respectivement sur des ouverts  $U$  et  $V$  d'un même espace  $E$ , alors leur somme  $f + g$  est différentiable sur  $U \cap V$  et

$$d(f + g) = df + dg,$$

c'est-à-dire que pour tout  $x \in U \cap V$  et pour tout  $h \in E$ ,

$$d(f + g)_x(h) = df_x(h) + dg_x(h).$$

De plus, quel que soit  $\lambda \in \mathbb{R}$ , la fonction  $\lambda f$  est différentiable et

$$d(\lambda f) = \lambda df.$$

La démonstration découle directement de la définition (et de l'inégalité triangulaire pour ce qui est de la somme). Cette proposition signifie que

- l'ensemble des fonctions différentiables sur  $U$  et à valeurs dans  $F$  est un espace vectoriel,
- la différentiation  $d$  est une application linéaire de cet espace vectoriel dans  $\mathcal{L}(E; F)$ .

Le résultat suivant est moins évident à démontrer, mais il est extrêmement important : on le désigne souvent par **règle de dérivation des fonctions composées** (*chain rule* en anglais). Dans l'énoncé,  $V$  est un ouvert de  $F$  et  $G$  est un espace de Banach.

**Théorème I.1** *Si  $f : U \subset E \rightarrow F$  est différentiable en un point  $x \in U$  et si  $g : V \subset F \rightarrow G$  est différentiable en  $f(x) \in V$ , alors la fonction composée  $g \circ f : U \subset E \rightarrow G$  est différentiable en  $x$  et*

$$d(g \circ f)_x(h) = dg_{f(x)}(df_x(h)) \quad \text{quels que soient } x \in U \text{ et } h \in E.$$

Autrement dit,

$$(3.4) \quad d(g \circ f)_x = dg_{f(x)} \circ df_x \quad \text{quel que soit } x \in U.$$

**Dém.** Notons

$$\varphi(h) = f(x+h) - f(x) - df_x(h) \quad \text{pour tout } h \in E,$$

$$\gamma(k) = g(y+k) - g(y) - dg_y(k) \quad \text{pour tout } k \in F, \quad \text{avec } y = f(x).$$

Par définition de la différentiabilité de  $f$  et de  $g$ ,

$$\lim_{h \xrightarrow{E} 0} \frac{\|\varphi(h)\|_F}{\|h\|_E} = 0 \quad \text{et} \quad \lim_{k \xrightarrow{F} 0} \frac{\|\gamma(k)\|_G}{\|k\|_F} = 0,$$

ce que l'on peut aussi exprimer sous la forme

$$\|\gamma(k)\|_G = \varepsilon(k) \|k\|_F \quad \text{avec} \quad \lim_{k \rightarrow 0_F} \varepsilon(k) = 0.$$

(Il suffit de poser  $\varepsilon(0_F) = 0$  et  $\varepsilon(k) = \|\gamma(k)\|_G / \|k\|_F$  pour  $k \neq 0_F$ .)

On a

$$\begin{aligned} g(f(x+h)) - g(f(x)) &= g(f(x) + df_x(h) + \varphi(h)) - g(f(x)) = \\ &= dg_y(df_x(h) + \varphi(h)) + \gamma(df_x(h) + \varphi(h)). \end{aligned}$$

Par conséquent, pour  $h \neq 0_E$ ,

$$\begin{aligned} \frac{\|g(f(x+h)) - g(f(x)) - dg_y(df_x(h))\|_G}{\|h\|_E} &= \frac{\|dg_y(\varphi(h)) + \gamma(df_x(h) + \varphi(h))\|_G}{\|h\|_E} \\ &\leq \|dg_y\| \frac{\|\varphi(h)\|_F}{\|h\|_E} + \left( \|df_x\| + \frac{\|\varphi(h)\|_F}{\|h\|_E} \right) \varepsilon(df_x(h) + \varphi(h)). \end{aligned}$$

Les deux morceaux ci-dessus tendent bien vers 0 lorsque  $h$  tend vers  $0_E$  puisque c'est le cas de  $\|\varphi(h)\|_F / \|h\|_E$ , que  $df_x(h) + \varphi(h)$  tend vers  $0_F$  et que  $\varepsilon(k)$  tend vers 0 lorsque  $k$  tend vers  $0_F$ .  $\square$

Les exemples donnés au paragraphe précédent et les propriétés énoncées dans la proposition et le théorème ci-dessus permettent de **calculer** beaucoup de différentielles (sans nécessairement revenir à la définition).

Voyons un exemple de calcul, dans lequel on suit la démarche générale : 1) supposer que la différentielle existe et calculer un candidat pour cette différentielle ; 2) injecter ce candidat dans la définition et vérifier que la fonction est bien différentiable. (On se limite ci-dessous à l'étape 1 ; l'étape 2 est laissée en exercice.)

Considérons l'ensemble  $U$  des  $u \in E := \mathcal{L}(X; Y)$  (où  $X$  et  $Y$  sont des  $\mathbb{R}$ -espaces de Banach) qui sont inversibles (et qu'on appelle des **isomorphismes**) :

$$\text{Isom}(X; Y) := \{ u \in \mathcal{L}(X; Y) ; \exists v \in \mathcal{L}(Y; X), v \circ u = \text{Id}_X \text{ et } u \circ v = \text{Id}_Y \}.$$

C'est un ouvert de  $E$ . Soit alors la fonction

$$f : U \rightarrow F := \mathcal{L}(Y; X) \\ u \mapsto u^{-1}.$$

L'application

$$\phi : E \times F \rightarrow \mathcal{L}(X; X) \\ (u, v) \mapsto v \circ u$$

est bilinéaire continue, et on a par définition  $\phi(u, f(u)) = \text{Id}_X$  (l'application identité sur  $X$ ). Par suite, en supposant *a priori* que  $f$  est différentiable, on a pour tout  $u \in U$  et  $h \in E$ ,

$$\phi(h, f(u)) + \phi(u, df_u(h)) = 0.$$

Autrement dit,

$$u^{-1} \circ h = -df_u(h) \circ u, \text{ c'est-à-dire } df_u(h) = -u^{-1} \circ h \circ u^{-1}.$$

**Exercice.** Vérifier, en reportant dans la définition le « candidat » trouvé ci-dessus pour sa différentielle, que  $f$  est différentiable.

## 4 En dimension finie

On suppose dans ce paragraphe  $E = \mathbb{R}^p$  et  $F = \mathbb{R}^q$ , avec  $p, q \in \mathbb{N}^*$ .

**Notations.** Pour un vecteur  $x \in \mathbb{R}^p$ , resp.  $y \in \mathbb{R}^q$ , on notera  $(x_1, \dots, x_p)$ , resp.  $(y_1, \dots, y_q)$  ses composantes. (Attention à ne pas confondre avec les  $n$ -uplets de vecteurs considérés au paragraphe 2.) La « matrice du vecteur »  $x$ , resp.  $y$ , sera notée  $X$ , resp.  $Y$ . Ce sont les matrices colonnes :

$$X = \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} \in \mathcal{M}_{p,1}(\mathbb{R}), \quad Y = \begin{pmatrix} y_1 \\ \vdots \\ y_q \end{pmatrix} \in \mathcal{M}_{q,1}(\mathbb{R}).$$

On note  $e_1, \dots, e_p$  les vecteurs de la **base canonique** de  $\mathbb{R}^p$  : par définition, le vecteur  $e_i$  (pour  $i \in \{1, \dots, p\}$ ) a toutes ses composantes nulles sauf la  $i$ -ème qui vaut 1.

**Dérivées partielles.** Pour tout  $x \in U$  de composantes  $(x_1, \dots, x_p)$  et pour tout  $i \in \{1, \dots, p\}$ , l'ensemble  $V_i(x) := \{t \in \mathbb{R}; (x_1, \dots, x_{i-1}, t, x_{i+1}, \dots, x_p) \in U\}$  est un voisinage ouvert de  $x_i$ .

Supposons  $f : U \subset \mathbb{R}^p \rightarrow F$  différentiable. Alors l'**application partielle**

$$\begin{aligned} g_i : V_i(x) &\rightarrow F \\ t &\mapsto f(x_1, \dots, x_{i-1}, t, x_{i+1}, \dots, x_p) = f(x + (t - x_i) e_i) \end{aligned}$$

est dérivable en  $x_i$  et

$$g'_i(x_i) = df_x(e_i)$$

(dérivée de  $f$  dans la direction  $e_i$  au point  $x$ ). Il est d'usage de noter cette dérivée

$$\frac{\partial f}{\partial x_i}(x)$$

ou même simplement  $\partial_{x_i} f(x)$ . On appelle **dérivées partielles** de  $f$  les fonctions

$$\begin{aligned} \partial_{x_i} f : U &\rightarrow F \\ x &\mapsto \frac{\partial f}{\partial x_i}(x) \end{aligned}$$

pour  $i \in \{1, \dots, p\}$ . Par linéarité de  $df_x$ , on voit que pour tout  $h \in \mathbb{R}^p$ , de composantes  $(h_1, \dots, h_p)$ ,

$$df_x(h) = df_x\left(\sum_{i=1}^p h_i e_i\right) = \sum_{i=1}^p h_i df_x(e_i) = \sum_{i=1}^p h_i \frac{\partial f}{\partial x_i}(x).$$

**Notation.** Quel que soit  $i \in \{1, \dots, p\}$ , l'application  $h \in \mathbb{R}^p \mapsto h_i \in \mathbb{R}$  est une forme linéaire continue (c'est-à-dire un élément de  $\mathcal{L}(\mathbb{R}^p; \mathbb{R})$ ). On la note  $dx_i$ . Ainsi, la différentielle de  $f$  au point  $x$  s'écrit

$$df_x = \sum_{i=1}^p \frac{\partial f}{\partial x_i}(x) dx_i.$$

**Attention !** L'existence de dérivées partielles n'est pas suffisante en général pour qu'une fonction soit différentiable.

**Théorème I.2** Une application  $f : U \subset \mathbb{R}^p \rightarrow F = \mathbb{R}^q$  est **continûment différentiable** si et seulement si ses  $p$  **dérivées partielles** existent et sont **continues** sur  $U$ .

**Dém.** Si  $f$  est continûment différentiable, alors pour tout  $i \in \{1, \dots, p\}$  sa dérivée partielle

$$\partial_{x_i} f : x \in U \mapsto \frac{\partial f}{\partial x_i}(x) = df_x(e_i)$$

est continue comme composée des fonctions continues

$$\begin{aligned} df : U &\rightarrow \mathcal{L}(\mathbb{R}^p; F) & \text{et} & \mathcal{L}(\mathbb{R}^p; F) &\rightarrow F \\ x &\mapsto df_x & \ell &\mapsto \ell(e_i). \end{aligned}$$

Réciproquement, supposons que pour tout  $i \in \{1, \dots, p\}$  et pour tout  $x \in U$ , l'application partielle  $t \mapsto f(x_1, \dots, x_{i-1}, t, x_{i+1}, \dots, x_p)$  soit dérivable en  $x_i$ , et que sa dérivée en  $x_i$  définisse une fonction continue sur  $U$ , notée  $\partial_{x_i} f$ . Alors la fonction

$$\begin{aligned} U &\rightarrow \mathcal{L}(\mathbb{R}^p; F) \\ x &\mapsto \sum_{i=1}^p \frac{\partial f}{\partial x_i}(x) dx_i \end{aligned}$$

est continue comme composée des fonctions continues :

$$\begin{aligned} U &\rightarrow F \times \dots \times F & \text{et} & \quad F \times \dots \times F \rightarrow \mathcal{L}(\mathbb{R}^p; F) \\ x &\mapsto \left( \frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_p}(x) \right) & (y^1, \dots, y^p) &\mapsto \sum_{i=1}^p y^i dx_i. \end{aligned}$$

(Dans la notation ci-dessus,  $y^i dx_i \in \mathcal{L}(\mathbb{R}^p; F)$  est défini par  $(y^i dx_i)(h) = h_i y^i$ .) Vérifier que  $f$  est effectivement différentiable de différentielle

$$df_x = \sum \partial_{x_i} f(x) dx_i,$$

demande un peu de travail. On peut raisonner par récurrence. Le cas  $p = 1$  est clair (c'est l'exemple des fonctions de variable réelle). Supposons que ce soit vrai pour les fonctions  $p$  variables réelles, et considérons une fonction sur un ouvert de  $\mathbb{R}^{p+1}$  admettant  $(p+1)$  dérivées partielles continues. Alors

$$\begin{aligned} f(x+h) - f(x) - \sum_{i=1}^{p+1} h_i \partial_{x_i} f(x) = \\ f(x_1 + h_1, \dots, x_p + h_p, x_{p+1} + h_{p+1}) - f(x_1 + h_1, \dots, x_p + h_p, x_{p+1}) - h_{p+1} \partial_{x_{p+1}} f(x) \\ + f(x_1 + h_1, \dots, x_p + h_p, x_{p+1}) - f(x_1, \dots, x_p, x_{p+1}) - \sum_{i=1}^p h_i \partial_{x_i} f(x). \end{aligned}$$

Lorsqu'on divise cette égalité par  $\|(h_1, \dots, h_{p+1})\|$ , le second morceau tend vers 0 avec  $h$  d'après l'hypothèse de récurrence. Il reste à montrer que le premier morceau tend aussi vers 0. On va pour cela appliquer le **théorème des accroissements finis** à la fonction d'une variable réelle

$$g : \xi \mapsto f(x_1 + h_1, \dots, x_p + h_p, \xi) - (\xi - x_{p+1}) \partial_{x_{p+1}} f(x).$$

Cette fonction est dérivable au voisinage de  $x_{p+1}$ , de dérivée :

$$g'(\xi) = \partial_{x_{p+1}} f(x_1 + h_1, \dots, x_p + h_p, \xi) - \partial_{x_{p+1}} f(x).$$

Or, par hypothèse, la fonction  $\partial_{x_{p+1}} f$  est continue au point  $x$ . Soit donc  $\varepsilon > 0$ . Il existe  $\eta > 0$  tel que

$$\max_{i \in \{1, \dots, p+1\}} |(y-x)_i| \leq \eta \quad \Rightarrow \quad \|\partial_{x_{p+1}} f(y) - \partial_{x_{p+1}} f(x)\|_F \leq \varepsilon.$$

Par conséquent, si  $\max_{i \in \{1, \dots, p+1\}} \|h_i\| \leq \eta$  et  $\xi$  appartient au segment d'extrémités  $x_{p+1}$  et  $x_{p+1} + h_{p+1}$ ,  $\|g'(\xi)\|_F \leq \varepsilon$ . Ceci implique (en appliquant le théorème des accroissements finis aux composantes de  $g$ , qui sont des fonctions de  $\mathbb{R}$  dans  $\mathbb{R}$  !),

$$\|g(x_{p+1} + h_{p+1}) - g(x_{p+1})\|_F \leq \varepsilon |h_{p+1}|.$$

Grâce à l'inégalité triangulaire, on en déduit qu'il existe  $\eta_1 \leq \eta$  tel que  $\max_{i \in \{1, \dots, p+1\}} |h_i| \leq \eta_1$  implique

$$\left\| f(x+h) - f(x) - \sum_{i=1}^{p+1} h_i \partial_{x_i} f(x) \right\|_F \leq \varepsilon (|h_1| + \dots + |h_{p+1}|).$$

□

**Matrice jacobienne.** Si une fonction  $f : U \subset \mathbb{R}^p \rightarrow \mathbb{R}^q$ , de composantes  $(f_1, \dots, f_q)$ , est différentiable au point  $x$ , on définit sa matrice jacobienne au point  $x$  comme la matrice de l'application linéaire  $df_x$  dans les bases canoniques de  $\mathbb{R}^p$  et  $\mathbb{R}^q$ . Elle est donnée par

$$Df(x) = \begin{pmatrix} \partial_{x_1} f_1(x) & \cdots & \partial_{x_p} f_1(x) \\ \vdots & & \vdots \\ \partial_{x_1} f_q(x) & \cdots & \partial_{x_p} f_q(x) \end{pmatrix} \in \mathcal{M}_{q,p}(\mathbb{R}).$$

Autrement dit, le coefficient de la matrice jacobienne de  $f$  d'indice  $i \in \{1, \dots, q\}$  en ligne et  $j \in \{1, \dots, p\}$  en colonne est

$$(Df(x))_{i,j} = \partial_{x_j} f_i(x).$$

En particulier, si  $q = 1$ ,  $Df(x)$  est une matrice ligne. De façon générale, les lignes des  $Df(x)$  sont les  $Df_i(x)$ .

**Nabla.** On note parfois  $\nabla f(x)$  (qui se lit « nabla  $f$  de  $x$  ») la matrice transposée de  $Df(x)$  :

$$(\nabla f(x))_{i,j} = \partial_{x_i} f_j(x).$$

En fait, cette notation est surtout utilisée lorsque  $q = 1$ , auquel cas  $\nabla f(x)$  est une matrice colonne, que l'on identifie à un vecteur de  $\mathbb{R}^p$  appelé **gradient** de  $f$  au point  $x$ .

**Opérateurs différentiels classiques.** Comme on vient de le voir, pour une fonction différentiable à **valeurs scalaires**,  $\varphi : U \subset \mathbb{R}^p \rightarrow \mathbb{R}$ , le **gradient** est défini par :

$$\begin{aligned} \text{grad } \varphi : U \subset \mathbb{R}^p &\rightarrow \mathbb{R}^p \\ x &\mapsto (\text{grad } \varphi)(x) := (\partial_{x_1} \varphi(x), \dots, \partial_{x_p} \varphi(x)). \end{aligned}$$

Par ailleurs, pour une fonction différentiable  $f : U \subset \mathbb{R}^p \rightarrow \mathbb{R}^p$  (noter l'égalité des dimensions au départ et à l'arrivée), de composantes  $(f_1, \dots, f_p)$ , on définit sa **divergence** par

$$\begin{aligned} \text{div } f : U \subset \mathbb{R}^p &\rightarrow \mathbb{R} \\ x &\mapsto (\text{div } f)(x) := \text{tr}(Df(x)) = \sum_{i=1}^p \partial_{x_i} f_i(x). \end{aligned}$$

Lorsque  $p = 3$ , on définit aussi le **rotationnel** de  $f$

$$\begin{aligned} \text{rot } f : U \subset \mathbb{R}^3 &\rightarrow \mathbb{R}^3 \\ x &\mapsto (\text{rot } f)(x) \end{aligned}$$

par

$$(\text{rot } f)(x) := (\partial_{x_2} f_3(x) - \partial_{x_3} f_2(x), \partial_{x_3} f_1(x) - \partial_{x_1} f_3(x), \partial_{x_1} f_2(x) - \partial_{x_2} f_1(x)).$$

# Chapitre II

## Théorème des accroissements finis

Ce chapitre est consacré à l'un des résultats fondamentaux du calcul différentiel, et à quelques unes de ses applications.

Pour une fonction de **variable réelle** et à **valeurs réelles**, on connaît la **formule des accroissements finis** (conséquence du théorème de Rolle) : si  $\varphi : I \subset \mathbb{R} \rightarrow \mathbb{R}$  est dérivable sur l'intervalle  $I$ , alors pour tout  $(x, y) \in I \times I$  (avec  $x < y$ ), il existe  $t \in ]x, y[$  tel que

$$f(x) - f(y) = f'(t)(x - y).$$

Par suite, si  $|f'|$  est majorée par une constante  $k > 0$  sur l'intervalle  $I$ , on a l'**inégalité des accroissements finis** :

$$|f(x) - f(y)| \leq k|x - y| \quad \text{quel que soit } (x, y) \in I \times I.$$

**Attention !** Pour une fonction  $f$  à valeurs dans un espace autre que  $\mathbb{R}$ , il n'y a **pas de formule** des accroissements finis : par exemple, la fonction  $f : x \mapsto e^{ix}$  est dérivable, et sa dérivée  $x \mapsto ie^{ix}$  ne s'annule pas, bien qu'il existe des points (nombreux)  $x$  et  $y$  distincts où  $f(x) = f(y)$  ! En revanche, l'**inégalité** des accroissements finis reste vraie.

### 1 Fonctions d'une variable réelle

**Théorème II.1** Soit  $f : I \subset \mathbb{R} \rightarrow F$  une fonction dérivable sur un intervalle ouvert  $I$  et à valeurs dans un  $\mathbb{R}$ -espace de Banach  $F$ . On suppose qu'il existe  $k > 0$  tel que

$$\|f'(t)\|_F \leq k \quad \text{quel que soit } t \in I.$$

Alors

$$(1.1) \quad \|f(x) - f(y)\|_F \leq k|x - y| \quad \text{quel que soit } (x, y) \in I \times I.$$

**Dém.** Pour fixer les idées, on prend  $(x, y) \in I \times I$  avec  $x < y$ . On ne va pas montrer directement (1.1), mais le résultat intermédiaire suivant : pour tout  $\varepsilon > 0$  et pour tout  $t \in [x, y]$ ,

$$(1.2) \quad \|f(t) - f(x)\|_F \leq (k + \varepsilon)(t - x) + \varepsilon.$$

En l'appliquant à  $t = y$  puis en faisant tendre  $\varepsilon$  vers 0, on en déduit en effet (1.1). Pour démontrer le résultat intermédiaire, considérons l'ensemble :

$$\mathcal{O} := \{t \in [x, y]; \|f(t) - f(x)\|_F > (k + \varepsilon)(t - x) + \varepsilon\}.$$

$\mathcal{O}$  est un ensemble ouvert (car  $f$  est continue) dans  $[x, y]$ , dont on veut montrer qu'il est vide.

Supposons que  $\mathcal{O}$  ne soit pas vide, et soit alors  $\underline{t}$  sa borne inférieure. Comme l'inégalité (1.2) est vraie au voisinage de  $x$  (par continuité de l'application

$$t \mapsto \|f(t) - f(x)\|_F - (k + \varepsilon)(t - x),$$

qui vaut évidemment 0 en  $t = x$ ), on a  $\underline{t} > x$ , et donc  $\underline{t} = \inf \mathcal{O}$  ne peut pas être dans l'ouvert  $\mathcal{O}$  (si  $\tau > x$  et si  $\tau$  appartient à  $\mathcal{O}$ ,  $[\underline{t}, \tau] \cap \mathcal{O}$  est non vide, donc  $\tau$  ne minore pas  $\mathcal{O}$  et ne peut par conséquent pas être la borne inférieure de  $\mathcal{O}$ ). Donc

$$\|f(\underline{t}) - f(x)\|_F \leq (k + \varepsilon)(\underline{t} - x) + \varepsilon.$$

D'autre part, d'après la dérivabilité de  $f$  en  $\underline{t}$ , il existe  $\eta > 0$  tel que pour tout  $t \in ]\underline{t}, \underline{t} + \eta]$ ,

$$k \geq \|f'(t)\|_F \geq \frac{\|f(t) - f(\underline{t})\|_F}{|t - \underline{t}|} - \varepsilon,$$

d'où

$$\|f(t) - f(\underline{t})\|_F \leq (k + \varepsilon)(t - \underline{t}).$$

Par l'inégalité triangulaire, on en déduit

$$\|f(t) - f(x)\|_F \leq (k + \varepsilon)(t - x) + \varepsilon$$

pour  $t \in ]\underline{t}, \underline{t} + \eta]$ , c'est-à-dire que  $]\underline{t}, \underline{t} + \eta] \cap \mathcal{O}$  est vide. Or  $\underline{t} < y$  (si ce n'était pas le cas,  $\mathcal{O}$  serait réduit au singleton  $\{y\}$ , qui n'est pas ouvert dans  $[x, y]$  !) donc pour  $\eta$  assez petit,  $\underline{t} + \eta < y$ . Par conséquent,  $]\underline{t}, \underline{t} + \eta] \cap \mathcal{O} = \emptyset$  montre que  $\underline{t} + \eta$  est un minorant de  $\mathcal{O}$  strictement plus grand que  $\underline{t}$  : ceci contredit le fait que  $\underline{t}$  soit la borne inférieure de  $\mathcal{O}$ .  $\square$

**Remarque.** Le résultat s'applique même pour  $x$  et  $y$  au bord de l'intervalle  $I$ , à condition que  $f$  soit **continue sur l'intervalle fermé**  $\bar{I}$  et qu'on ait une estimation de  $f'$  sur l'intervalle ouvert  $I$ . En effet, dans la démonstration, l'ensemble  $\mathcal{O}$  reste un ouvert dont la borne inférieure, s'il est non vide, appartient à l'intervalle ouvert  $]x, y[$ .

La même méthode de démonstration permet de montrer le résultat plus général suivant.

**Théorème II.2** Soit  $f : I \subset \mathbb{R} \rightarrow F$  une fonction dérivable sur un intervalle ouvert  $I$  et à valeurs dans un  $\mathbb{R}$ -espace de Banach  $F$ . On suppose qu'il existe une fonction  $\varphi : I \rightarrow \mathbb{R}$  dérivable telle que

$$\|f'(t)\|_F \leq \varphi'(t) \quad \text{quel que soit } t \in I.$$

Alors

$$(1.3) \quad \|f(x) - f(y)\|_F \leq |\varphi(x) - \varphi(y)| \quad \text{quel que soit } (x, y) \in I \times I.$$

## 2 Théorème général

Comme au chapitre I,  $E$  et  $F$  désignent des  $\mathbb{R}$ -espace de Banach.

**Théorème II.3** Soit  $f : U \subset E \rightarrow F$  une fonction différentiable sur un ouvert **convexe**  $U$ . On suppose qu'il existe  $k > 0$  tel que

$$\| \mathrm{d}f_u \| \leq k \quad \text{quel que soit } u \in U.$$

Alors

$$(2.4) \quad \| f(x) - f(y) \|_F \leq k \| x - y \|_E \quad \text{quel que soit } (x, y) \in U \times U.$$

**Dém.** Fixons  $(x, y) \in U \times U$ . D'après les hypothèses, le segment d'extrémités  $x$  et  $y$  est inclus dans  $U$ , et la fonction d'une variable réelle

$$\begin{aligned} g : [0, 1] &\rightarrow F \\ t &\mapsto f(x + t(y - x)) \end{aligned}$$

est dérivable, de dérivée

$$g'(t) = \mathrm{d}f_{(x+t(y-x))}(y - x),$$

satisfaisant la majoration  $\|g'(t)\|_F \leq k \|y - x\|_E$  pour tout  $t \in [0, 1]$ . D'après le théorème II.1, on en déduit

$$\| f(y) - f(x) \|_F = \| g(1) - g(0) \|_F \leq k \| y - x \|_E.$$

□

**Remarque.** La démonstration donne en fait l'inégalité plus fine

$$\| f(y) - f(x) \|_F \leq \sup_{t \in [0, 1]} \| \mathrm{d}f_{(x+t(y-x))} \| \| y - x \|_E.$$

### 3 Applications

Le théorème des accroissements finis a de nombreuses applications. Les plus « fondamentales » sont

- la caractérisation des fonctions de différentielle nulle sur les ouverts *connexes*,
- la caractérisation des fonctions de classe  $\mathcal{C}^1$  sur un produit cartésien.

**Théorème II.4** Soit  $f : U \subset E \rightarrow F$  une fonction différentiable sur un ouvert **connexe**  $U$ , telle que  $\mathrm{d}f_x \equiv 0$  pour tout  $x \in U$ . Alors  $f$  est **constante**.

**Dém.** Quel que soit  $x \in U$ , il existe  $r > 0$  tel que la boule ouverte de centre  $x$  et de rayon  $r$ ,  $B(x; r)$  soit incluse dans  $U$ . Cette boule est convexe, et puisque  $\mathrm{d}f \equiv 0$ , le théorème des accroissements finis montre que  $f(y) = f(x)$  quel que soit  $y \in B(x; r)$  : cela signifie que  $f$  est **localement constante**. Comme  $U$  est connexe, on en déduit facilement que  $f$  est constante. En effet, fixons  $x \in U$ . L'ensemble  $f^{-1}(\{f(x)\})$  est non vide puisqu'il contient  $x$ , et fermé par continuité de  $f$ . D'après ce qui précède, cet ensemble est aussi ouvert. Puisque  $U$  est connexe, on a donc  $f^{-1}(\{f(x)\}) = U$ . Autrement dit,  $f(y) = f(x)$  pour tout  $y \in U$ . □

**Théorème II.5** Soient  $E_1, \dots, E_n$  des espaces de Banach et  $E = E_1 \times \dots \times E_n$  muni (par exemple) de la norme

$$\|(x_1, \dots, x_n)\|_E = \|x_1\|_{E_1} + \dots + \|x_n\|_{E_n}.$$

Une fonction

$$\begin{aligned} f : U \subset E &\rightarrow F \\ x = (x_1, \dots, x_n) &\mapsto f(x) \end{aligned}$$

est continûment différentiable si et seulement si pour tout  $i \in \{1, \dots, n\}$  et pour tout  $x = (x_1, \dots, x_n) \in U$ , l'application partielle

$$y_i \mapsto (x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_n)$$

est différentiable en  $y_i = x_i$  et sa différentielle définit une fonction (appelée **différentielle partielle**) continue de  $U$  dans  $\mathcal{L}(E_i; F)$ .

**Dém.** Il y a un sens facile. Si  $f$  est continûment différentiable, quels que soient  $i \in \{1, \dots, n\}$  et  $x = (x_1, \dots, x_n) \in U$ , l'application partielle  $g : y_i \mapsto (x_1, \dots, x_{i-1}, y_i, x_{i+1}, \dots, x_n)$  est différentiable en  $y_i = x_i$ , de différentielle définie par

$$dg_{x_i}(h_i) = df_x(0, \dots, 0, h_i, 0, \dots, 0)$$

pour tout  $h_i \in E_i$ . Notons  $d_i f_x \in \mathcal{L}(E_i; F)$  l'application définie comme ci-dessus par  $d_i f_x(h_i) = df_x(0, \dots, 0, h_i, 0, \dots, 0)$ , et

$$\begin{aligned} d_i f : U &\rightarrow \mathcal{L}(E_i; F) \\ x &\mapsto d_i f_x \end{aligned}$$

Alors  $d_i f$  est continue sur  $U$ , comme composée des fonctions continues

$$\begin{aligned} df : U &\rightarrow \mathcal{L}(E; F) & \text{et} & & \mathcal{L}(E; F) &\rightarrow \mathcal{L}(E_i; F) \\ x &\mapsto df_x & & & \ell &\mapsto (\ell(0, \dots, 0, h_i, 0, \dots, 0)). \end{aligned}$$

Pour tout  $x \in U$ , on a par linéarité de  $df_x$  :

$$df_x(h) = \sum_{i=1}^n d_i f_x(h_i)$$

pour tout  $h = (h_1, \dots, h_n) \in E$ .

La réciproque nécessite le théorème des accroissements finis. Supposons que pour tout  $i \in \{1, \dots, n\}$  la différentielle partielle  $d_i f$  existe et soit continue. Alors l'application

$$\begin{aligned} U &\rightarrow \mathcal{L}(E; F) \\ x &\mapsto \left( (h_1, \dots, h_n) \mapsto \sum_{i=1}^n d_i f_x(h_i) \right) \end{aligned}$$

est bien continue comme composée des fonctions continues

$$\begin{aligned} U &\rightarrow \mathcal{L}(E_1; F) \times \dots \times \mathcal{L}(E_n; F) \\ x &\mapsto (d_1 f_x, \dots, d_n f_x) \end{aligned}$$

(fonction à valeurs dans un produit dont chaque composante est continue) et

$$\begin{aligned} \mathcal{L}(E_1; F) \times \cdots \times \mathcal{L}(E_n; F) &\rightarrow \mathcal{L}(E; F) \\ (\ell^1, \dots, \ell^n) &\mapsto ((h_1, \dots, h_n) \mapsto \sum_{i=1}^n \ell^i(h_i)) \end{aligned}$$

(application linéaire continue). Il s'agit de montrer que  $f$  est différentiable, de différentielle précisément définie par

$$df_x(h) = \sum_{i=1}^n d_i f_x(h_i)$$

pour tout  $x \in U$  et pour tout  $h = (h_1, \dots, h_n) \in E$ . On peut raisonner par récurrence sur  $n$ . Le cas  $n = 1$  est évident : si  $E = E_1$ ,  $df_x(h) = d_1 f_x(h_1)$  pour tout  $x \in U$  et pour tout  $h = h_1 \in E$ . Supposons  $n \geq 2$  et le résultat démontré pour les fonctions définies sur un ouvert d'espace produit de  $(n - 1)$  espaces. Soit  $x \in U$  et  $r > 0$  tel que

$$\max_{i \in \{1, \dots, n\}} \|(y - x)_i\|_{E_i} \leq r \quad \Rightarrow \quad y = (y_1, \dots, y_n) \in U.$$

Pour tout  $h = (h_1, \dots, h_n) \in E$  tel que  $\max_{i \in \{1, \dots, n\}} \|h_i\|_{E_i} \leq r$  on a

$$\begin{aligned} f(x + h) - f(x) - \sum_{i=1}^n d_i f_x(h_i) = \\ f(x_1 + h_1, \dots, x_{n-1} + h_{n-1}, x_n + h_n) - f(x_1 + h_1, \dots, x_{n-1} + h_{n-1}, x_n) - d_n f_x(h_n) \\ + f(x_1 + h_1, \dots, x_{n-1} + h_{n-1}, x_n) - f(x_1, \dots, x_{n-1}, x_n) - \sum_{i=1}^{n-1} d_i f_x(h_i). \end{aligned}$$

Lorsqu'on divise cette égalité par  $\|(h_1, \dots, h_n)\|$ , le second morceau tend vers 0 avec  $h$  d'après l'hypothèse de récurrence. Il reste à montrer que le premier morceau tend aussi vers 0. C'est ici que va intervenir le **théorème des accroissements finis**. Par hypothèse, la fonction

$$g : y_n \mapsto f(x_1 + h_1, \dots, x_{n-1} + h_{n-1}, y_n) - d_n f_x(y_n - x_n)$$

est différentiable dans la boule ouverte de rayon  $r$ , centrée en  $x_n$  dans l'espace  $E_n$ , de différentielle définie par

$$dg_{y_n}(h_n) = d_n f_{(x_1+h_1, \dots, x_{n-1}+h_{n-1}, y_n)}(h_n) - d_n f_x(h_n)$$

et donc de norme

$$\|dg_{y_n}\| \leq \|d_n f_{(x_1+h_1, \dots, x_{n-1}+h_{n-1}, y_n)} - d_n f_x\|.$$

Or, par hypothèse, la différentielle partielle  $d_n f$  est continue au point  $x = (x_1, \dots, x_n)$ . Soit donc  $\varepsilon > 0$ . Il existe  $\eta \in ]0, r]$  tel que

$$\max_{i \in \{1, \dots, n\}} \|(y - x)_i\|_{E_i} \leq \eta \quad \Rightarrow \quad \|d_n f_{(y_1, \dots, y_n)} - d_n f_{(x_1, \dots, x_n)}\| \leq \varepsilon.$$

Par conséquent, si  $\max_{i \in \{1, \dots, n\}} \|h_i\|_{E_i} \leq \eta$ ,

$$\|dg_{y_n}\| \leq \varepsilon \quad \text{pour} \quad \|y_n - x_n\|_{E_n} \leq \eta,$$

et le théorème des accroissements finis (appliqué à  $g$  dans la boule de centre  $x_n$  et de rayon  $\eta$ , qui est bien convexe) implique

$$\|g(x_n + h_n) - g(x_n)\|_F \leq \varepsilon \|h_n\|_{E_n}.$$

Grâce à l'inégalité triangulaire, on en déduit qu'il existe  $\eta_1 \leq \eta$  tel que  $\max_{i \in \{1, \dots, n\}} \|h_i\|_{E_i} \leq \eta_1$  implique

$$\left\| f(x + h) - f(x) - \sum_{i=1}^n d_i f_x(h_i) \right\|_F \leq \varepsilon (\|h_1\|_{E_1} + \dots + \|h_n\|_{E_n}).$$

□

Comme l'auront remarqué les lecteurs perspicaces, la démonstration ci-dessus est calquée sur celle du théorème I.2 dans le chapitre I, que l'on peut voir comme un cas particulier du théorème II.5 avec  $E_1 = \dots = E_n = \mathbb{R}$ .

# Chapitre III

## Difféomorphismes

### 1 Introduction

Soient  $U$  et  $V$  des ouverts (non vides) d'espaces de Banach  $E$  et  $F$  respectivement.

**Définition III.1** On dit qu'une application  $f : U \rightarrow V$  est un **difféomorphisme** (de  $U$  sur  $V$ ) si et seulement si

- i).  $f$  est une **bijection**,
- ii).  $f$  est de classe  $\mathcal{C}^1$ , c'est-à-dire **continûment différentiable** sur  $U$ ,
- iii).  $f^{-1}$  est de classe  $\mathcal{C}^1$  sur  $V$ .

Attention, le point iii) est important.

**Exemple** (avec  $E = F = \mathbb{R}$  !)

La fonction trigonométrique  $\tan$  est un difféomorphisme de  $] - \pi/2, \pi/2 [$  sur  $\mathbb{R}$ .

**Contre-exemple** (avec  $E = F = \mathbb{R}$  !)

La fonction polynômiale  $x \mapsto x^3$  n'est *pas* un difféomorphisme de  $\mathbb{R}$  dans  $\mathbb{R}$ , bien que ce soit une bijection, continûment différentiable : sa réciproque n'est en effet pas différentiable en  $y = 0$ .

**Proposition III.1** Si  $f : U \rightarrow V$  est un difféomorphisme, sa différentielle est en tout point de  $U$  un isomorphisme de  $E$  sur  $F$ , et la différentielle de la fonction réciproque  $f^{-1}$  est liée à celle de  $f$  par la formule :

$$d(f^{-1})_y = (df_{f^{-1}(y)})^{-1} \quad \text{pour tout } y \in V .$$

**Dém.** Pour simplifier l'écriture, notons  $g = f^{-1}$ . Par définition, on a

$$g \circ f = \text{Id}_U \quad \text{et} \quad f \circ g = \text{Id}_V ,$$

d'où en appliquant la règle de dérivation des fonctions composées (théorème I.1)

$$dg_y \circ df_x = \text{Id}_E \quad \text{et} \quad df_x \circ dg_y = \text{Id}_F \quad \text{pour tout } x \in U \quad \text{et} \quad y = f(x) .$$

□

**Corollaire III.1** S'il existe un difféomorphisme d'un ouvert de  $E$  sur un ouvert de  $F$ , les deux espaces sont **isomorphes**. En particulier, si l'un d'eux est de dimension finie, l'autre aussi et sa dimension est la même.

## 2 Théorème d'inversion locale

Le théorème d'inversion locale est une sorte de réciproque de la proposition III.1.

**Théorème III.1** *Si  $f : U \rightarrow V$  est de classe  $\mathcal{C}^1$ , si  $a \in U$  est tel que  $df_a$  soit un isomorphisme de  $E$  sur  $F$ , il existe un voisinage ouvert  $U_a$  de  $a$  dans  $U$  et un voisinage ouvert  $V_b$  de  $b = f(a)$  dans  $V$  tel que la restriction de  $f$  à  $U_a$  soit un difféomorphisme de  $U_a$  sur  $V_b$ .*

La démonstration utilise essentiellement trois ingrédients :

- 1) le fait que l'ensemble  $\text{Isom}(E; F)$  des isomorphismes de  $E$  sur  $F$  soit un ouvert et que l'application  $u \in \text{Isom}(E; F) \mapsto u^{-1} \in \text{Isom}(F; E)$  soit continue,
- 2) le théorème des accroissements finis,
- 3) le théorème du point fixe de Banach-Picard.

Le point 1) a déjà été mentionné au chapitre I. Il est laissé en exercice. Pour le point 2) on renvoie au chapitre II. Pour le point 3), rappelons le

**Théorème III.2** *Si  $C$  est un fermé non vide d'un espace de Banach  $E$  et si  $h : C \rightarrow C$  est contractante, c'est-à-dire qu'il existe  $k \in ]0, 1[$  tel que*

$$\|h(x) - h(x')\|_E \leq k \|x - x'\|_E \quad \text{quels que soient } x \text{ et } x' \in C,$$

alors il existe un unique  $x_0 \in C$  tel que  $h(x_0) = x_0$ .

**Attention !** Dans ce théorème, toutes les hypothèses sont cruciales : l'ensemble  $C$  doit être complet (on a besoin du critère de Cauchy) et stable par  $h$  (on a besoin d'itérer  $h$ ), et la constante  $k$  doit être strictement inférieure à 1 (pour pouvoir appliquer le critère de Cauchy à la suite des itérés  $h^n(x)$ ).

**Démonstration du théorème III.1** On procède en quatre étapes :

**(0) Réduction du problème.** Pour simplifier les calculs, on peut sans perte de généralité se ramener au cas où

$$(2.1) \quad E = F, \quad a = b = 0 \quad \text{et} \quad df_a = \text{Id}_E.$$

C'est possible en remplaçant l'ouvert  $U$  par l'ensemble  $\tilde{U}$  des  $x$  tels que  $a + x$  appartient à  $U$ , et  $f$  par  $\tilde{f}$  définie par

$$\tilde{f}(x) = (df_a)^{-1} (f(a + x) - f(a)).$$

Si l'on démontre le résultat pour  $\tilde{f}$ , on en déduira immédiatement le résultat pour  $f$ . Désormais, on enlève les tildas et l'on se place dans le cas (2.1).

**(1) Existence de la fonction réciproque locale.** Considérons l'application

$$g : x \in U \mapsto g(x) := x - f(x).$$

D'après les hypothèses et (2.1),  $g$  est de classe  $\mathcal{C}^1$  et telle que

$$g(0) = 0, \quad dg_0 = 0.$$

Donc il existe  $r > 0$  tel que pour  $\|x\| \leq 2r$ ,  $x \in U$  et

$$\|dg_x\| \leq 1/2.$$

D'après le théorème des accroissements finis appliqué à  $g$  dans la boule  $B(0; 2r)$  de centre 0 et de rayon  $2r$ ,

$$\|g(x)\| = \|g(x) - g(0)\| \leq \frac{1}{2} \|x\| < r$$

pour tout  $x \in B(0; 2r)$ . Ceci signifie  $g(B(0; 2r)) \subset B(0; r)$ . Par continuité de  $g$ , on a aussi l'inclusion avec les boules fermées :  $g(\overline{B}(0; 2r)) \subset \overline{B}(0; r)$ .

Nous allons montrer que pour tout  $y \in B(0; r)$  il existe un unique  $x \in B(0; 2r)$  tel que  $y = f(x)$ . Pour cela, fixons  $y \in B(0; r)$  et considérons l'application

$$h : x \in \overline{B}(0; 2r) \mapsto h(x) := y + g(x).$$

On aura  $y = f(x)$  si et seulement si  $h(x) = x$ . C'est ici bien sûr qu'intervient le théorème du point fixe. Vérifions d'abord que la boule fermée  $\overline{B}(0; 2r)$  est stable par  $h$  :

$$\text{si } x \in \overline{B}(0; 2r), \|h(x)\| \leq \|y\| + \|g(x)\| < r + r = 2r.$$

Vérifions ensuite que  $h$  est contractante :

$$\text{si } x \text{ et } x' \in B(0; 2r), \|h(x) - h(x')\| = \|g(x) - g(x')\| \leq \frac{1}{2} \|x - x'\|$$

en appliquant une nouvelle fois le théorème des accroissements finis à  $g$  dans  $B(0; 2r)$ . Par continuité de  $g$  (et donc de  $h$ ) à nouveau, l'inégalité ci-dessus est même vraie dans la boule fermée  $\overline{B}(0; 2r)$ . Donc le théorème III.2 s'applique bien à  $h$  dans  $\overline{B}(0; 2r)$ . On observe de plus que son point fixe  $x$  appartient en fait à la boule ouverte  $B(0; 2r)$  : puisque  $y \in B(0; r)$ ,

$$\|x\| = \|h(x)\| \leq \|y\| + \|g(x)\| < r + \|g(x)\| \leq 2r.$$

Par conséquent, la restriction de  $f$  à  $U_0 := B(0; 2r) \cap f^{-1}(B(0; r))$ , qui est un voisinage ouvert de 0, est une bijection de  $U_0$  sur  $B(0; r)$ . Par abus de notation on note simplement  $f^{-1}$  sa réciproque.

**(2) La fonction  $f^{-1}$  est Lipschitzienne** : Soient  $y = f(x)$  et  $y' = f(x') \in B(0; r)$ , avec  $x, x' \in B(0; 2r)$ . Alors

$$\|x - x'\| = \|y + g(x) - y' - g(x')\| \leq \|y - y'\| + \frac{1}{2} \|x - x'\|,$$

d'où  $\|x - x'\| \leq 2\|y - y'\|$ .

**(3) La fonction  $f^{-1}$  est de classe  $\mathcal{C}^1$**  : Par continuité de  $df$  et puisque  $\text{Isom}(E; E)$  est un ouvert, pour  $r$  assez petit,  $df_x \in \text{Isom}(E; E)$  pour tout  $x \in B(0; 2r)$  ; de plus, par continuité de  $u \in \text{Isom}(E; E) \mapsto u^{-1}$ , l'application

$$\begin{aligned} B(0; 2r) &\rightarrow \text{Isom}(E; E) \\ x &\mapsto (df_x)^{-1} \end{aligned}$$

est continue. Comme  $df_0 = \text{Id}_E$ , quitte à diminuer encore  $r$  on peut supposer :

$$\| (df_x)^{-1} \| \leq 2 \quad \text{pour tout } x \in B(0; 2r).$$

Cette estimation et la constante de Lipschitz obtenue au (2) permettent de montrer que le candidat  $(df_x)^{-1}$  est effectivement la différentielle de  $f^{-1}$  au point  $y = f(x)$ , quel que soit  $x \in B(0; 2r)$ . On procède comme d'habitude, en revenant à la définition. Il faut montrer que le rapport

$$\frac{\|f^{-1}(y+k) - f^{-1}(y) - (df_x)^{-1}(k)\|}{\|k\|}$$

tend vers 0 avec  $k$ . Pour cela on écrit  $x+h = f^{-1}(y+k)$  (pour  $k \in B(0; r - \|y\|)$ ). D'où

$$\begin{aligned} \|f^{-1}(y+k) - f^{-1}(y) - (df_x)^{-1}(k)\| &= \|x+h - x - (df_x)^{-1}(f(x+h) - f(x))\| \\ &\leq \|(df_x)^{-1}\| \|f(x+h) - f(x) - (df_x)(h)\|, \end{aligned}$$

d'où encore

$$\frac{\|f^{-1}(y+k) - f^{-1}(y) - (df_x)^{-1}(k)\|}{\|k\|} \leq 2 \frac{\|f(x+h) - f(x) - (df_x)(h)\|}{\|h\|} \frac{\|h\|}{\|k\|}.$$

Or d'après le (2),  $\|h\|/\|k\| \leq 2$ . On conclut en appliquant la définition de la différentiabilité à  $f$ .  $\square$

Le théorème d'inversion locale est fondamental en analyse. Il permet notamment de caractériser les difféomorphismes : le résultat suivant inclut la réciproque de la proposition III.1.

**Corollaire III.2** *Soit  $f : U \rightarrow F$  une application de classe  $\mathcal{C}^1$  avec  $U$  un ouvert non vide. C'est un difféomorphisme (de  $U$  sur  $f(U)$ ) si et seulement si elle est injective et sa différentielle est en tout point de  $U$  un isomorphisme de  $E$  sur  $F$ .*

**Dém.** La partie directe découle de la définition et de la proposition III.1. Réciproquement, supposons  $f$  injective et telle que  $df_x \in \text{Isom}(E; F)$  pour tout  $x \in U$ . Alors  $f$  est bien sûr bijective de  $U$  sur  $f(U)$ . De plus, le théorème d'inversion locale montre que  $f(U)$  est un ouvert. En effet, pour tout  $y = f(x) \in f(U)$ , il existe un voisinage ouvert de  $y$  qui s'écrit  $V_y = f(U_x)$ , où  $U_x$  est un voisinage ouvert de  $x$  dans  $U$ , et qui est par conséquent inclus dans  $f(U)$ . En fait, on montre de la même façon que l'image par  $f$  de tout ouvert inclus dans  $U$  est un ouvert. (On dit alors que  $f$  est une application **ouverte**.) Par suite, l'application réciproque  $f^{-1} : f(U) \rightarrow U$  est continue. Il reste à montrer qu'elle est continûment différentiable. Or l'application

$$\begin{aligned} f(U) &\rightarrow \mathcal{L}(F; E) \\ y &\mapsto (df_{f^{-1}(y)})^{-1}. \end{aligned}$$

est continue comme composée d'applications continues, et de plus, on sait d'après l'étape (3) de la démonstration du théorème d'inversion locale que c'est (localement) la différentielle de  $f^{-1}$ .  $\square$

En dimension finie, le résultat peut s'énoncer ainsi.

**Corollaire III.3** *Soit  $U$  un ouvert de  $\mathbb{R}^p$  et  $f : U \rightarrow \mathbb{R}^p$  injective et de classe  $\mathcal{C}^1$ . Alors  $f$  est un difféomorphisme si et seulement si le déterminant de sa matrice jacobienne (que l'on appelle simplement le **jacobien** de  $f$ ) ne s'annule pas sur  $U$ .*

### 3 Théorème des fonctions implicites

Parmi les conséquences fondamentales du théorème d'inversion locale, on trouve un résultat tout aussi important, connu sous le nom de théorème des fonctions implicites. Il concerne la résolution d'équations non-linéaires de la forme

$$f(x, y) = 0,$$

et doit son nom au fait que, sous les hypothèses que l'on va préciser, on peut en tirer  $y$  comme fonction de  $x$  : on dit alors que  $f(x, y) = 0$  définit *implicitement*  $y$ , ou encore  $y$  comme **fonction implicite** de  $x$ .

Dans l'énoncé qui suit,  $E$ ,  $F$  et  $G$  sont trois espaces de Banach.

**Théorème III.3** *Soit  $U$  un ouvert de  $E \times F$  et  $f : U \rightarrow G$  une fonction de classe  $\mathcal{C}^1$ . On suppose qu'il existe  $(a, b) \in U$  tel que  $f(a, b) = 0_G$  et la différentielle partielle de  $f$  par rapport à  $y$ ,  $d_2f$  est telle que  $d_2f_{(a,b)}$  soit un isomorphisme de  $F$  sur  $G$ . Alors il existe un voisinage ouvert  $U_{(a,b)}$  de  $(a, b)$  dans  $U$ , un voisinage ouvert  $W_a$  de  $a$  dans  $E$  et une fonction de classe  $\mathcal{C}^1$*

$$\varphi : W_a \rightarrow F$$

telle que :

$$((x, y) \in U_{(a,b)} \text{ et } f(x, y) = 0_G) \quad \Leftrightarrow \quad y = \varphi(x).$$

**Dém.** Comme annoncé, on va appliquer le théorème d'inversion locale. Pour cela, on considère la fonction

$$g : U \rightarrow E \times G \\ (x, y) \mapsto (x, f(x, y)).$$

Elle est de classe  $\mathcal{C}^1$ , et pour tout  $(x, y) \in U$ , pour tout  $(h, k) \in E \times F$ , on a

$$dg_{(x,y)}(h, k) = (h, d_1f_{(x,y)}(h) + d_2f_{(x,y)}(k)),$$

où  $d_1f$  désigne la différentielle partielle de  $f$  par rapport à  $x$ . Vérifions que  $dg_{(a,b)}$  est un isomorphisme de  $E \times F$  sur  $E \times G$  : d'après l'hypothèse sur la différentielle partielle  $d_2f$ , on a pour tout  $(h', k') \in E \times G$ ,

$$(h, k) \in E \times F \text{ et } dg_{(a,b)}(h, k) = (h', k')$$

équivalent à

$$h = h' \text{ et } k = (d_2f_{(a,b)})^{-1} (k' - d_1f_{(a,b)}(h')).$$

Donc  $dg_{(a,b)}$  est bien un isomorphisme, d'inverse

$$(dg_{(a,b)})^{-1} : E \times G \rightarrow E \times F \\ (h', k') \mapsto (h', (d_2f_{(a,b)})^{-1} (k' - d_1f_{(a,b)}(h'))).$$

Par conséquent,  $g$  est un difféomorphisme d'un voisinage ouvert  $U_{(a,b)}$  de  $(a, b)$  sur voisinage ouvert de  $(a, 0)$ , que l'on peut supposer de la forme  $W_a \times Z_0$  où  $W_a$  est un voisinage ouvert de  $a$  et  $Z_0$  est un voisinage ouvert de  $0_G$ . La réciproque de  $g$  est nécessairement de la forme

$$g^{-1}(x, z) = (x, \phi(x, z)),$$

avec  $\phi$  de classe  $\mathcal{C}^1$  sur  $W_a \times Z_0$ . Autrement dit, on a

$$((x, y) \in U_{(a,b)} \text{ et } f(x, y) = z) \quad \Leftrightarrow \quad ((x, z) \in W_a \times Z_0 \text{ et } y = \phi(x, z)).$$

En particulier,

$$((x, y) \in U_{(a,b)} \text{ et } f(x, y) = 0) \quad \Leftrightarrow \quad (x \in W_a \text{ et } y = \varphi(x)),$$

où l'on a noté  $\varphi(x) = \phi(x, 0)$ . □

**Proposition III.2** *Sous les hypothèses du théorème III.3, quitte à réduire  $W_a$  on a*

$$d\varphi_x(h) = - (d_2 f_{(x, \varphi(x))})^{-1} d_1 f_{(x, \varphi(x))}(h)$$

pour tout  $x \in W_a$  et pour tout  $h \in E$ .

**Dém.** L'image réciproque de l'ouvert  $\text{Isom}(F; G)$  par l'application continue  $d_2 f$  est un ouvert et il contient  $(a, b)$ . Donc, quitte à réduire  $U_{(a,b)}$  et donc aussi  $W_a$ , on peut supposer

$$d_2 f_{(x, \varphi(x))} \in \text{Isom}(F; G) \quad \text{pour tout } x \in W_a.$$

On obtient le résultat en différentiant l'identité (valable dans  $W_a$ ) :

$$f(x, \varphi(x)) \equiv 0. \quad \square$$

Signalons pour finir un résultat d'analyse fonctionnelle, que l'on admettra, et qui permet de simplifier la vérification des hypothèses des théorèmes d'inversion locale ou des fonctions implicites.

**Théorème III.4** *Si  $E$  et  $F$  sont des espaces de Banach, si  $u$  est une application linéaire continue et bijective de  $E$  sur  $F$ , alors sa réciproque est continue.*

(Évidemment, en dimension finie ce théorème est sans intérêt, puisque toutes les applications linéaires sont continues !) Autrement dit, pour vérifier que  $u$  est un isomorphisme de  $E$  sur  $F$ , il "suffit" de vérifier que  $u$  est **linéaire continue et bijective**.

# Chapitre IV

## Différentielles d'ordre supérieur

### 1 Différentielle seconde

**Définition IV.1** Une fonction  $f$  définie sur un ouvert (non vide)  $U$  d'un  $\mathbb{R}$ -espace de Banach  $E$  et à valeurs dans un  $\mathbb{R}$ -espace de Banach  $F$  est dite **deux fois différentiable** en  $x \in U$  si elle est différentiable dans un voisinage ouvert  $U_x$  de  $x$  et si sa différentielle  $df : U_x \rightarrow \mathcal{L}(E; F)$  est différentiable en  $x$ . On dit que  $f$  est deux fois différentiable dans  $U$  si elle est différentiable en tout point de  $U$ .

Par définition, la différentielle de  $df$  en  $x$ ,  $d(df)_x$  est une application linéaire continue de  $E$  dans  $\mathcal{L}(E; F)$ . Elle s'identifie naturellement avec une application bilinéaire continue sur  $E \times E$ , en vertu de la

**Proposition IV.1** Soient  $E$ ,  $F$  et  $G$  des espaces de Banach. Alors les espaces  $\mathcal{L}(E; \mathcal{L}(F; G))$  et  $\mathcal{L}(E, F; G)$  munis des normes usuelles :

$$\|\ell\|_{\mathcal{L}(E; \mathcal{L}(F; G))} = \sup\{\|\ell(h)\|_{\mathcal{L}(F; G)}, \|h\|_E \leq 1\},$$

$$\|\phi\|_{\mathcal{L}(E, F; G)} = \sup\{\|\phi(h, k)\|_G; \|h\|_E \leq 1, \|k\|_F \leq 1\}$$

sont isométriques.

**Dém.** L'isométrie « naturelle » est définie comme suit. À  $\phi \in \mathcal{L}(E, F; G)$ , on associe  $\ell$  définie par

$$\ell(h) : k \mapsto \phi(h, k).$$

Pour tout  $h \in E$ , on a  $\ell(h) \in \mathcal{L}(F; G)$  avec

$$\|\ell(h)\|_{\mathcal{L}(F; G)} = \sup_{k \in F \setminus \{0\}} \frac{\|\phi(h, k)\|_G}{\|k\|_F} \leq \|\phi\|_{\mathcal{L}(E, F; G)} \|h\|_E,$$

et de plus  $\ell \in \mathcal{L}(E; \mathcal{L}(F; G))$  avec

$$\|\ell\|_{\mathcal{L}(E; \mathcal{L}(F; G))} = \sup_{h \in E \setminus \{0\}} \sup_{k \in F \setminus \{0\}} \frac{\|\phi(h, k)\|_G}{\|h\|_E \|k\|_F} = \|\phi\|_{\mathcal{L}(E, F; G)}.$$

Inversement, à  $\ell \in \mathcal{L}(E; \mathcal{L}(F; G))$  on associe

$$\phi : (h, k) \mapsto \ell(h)(k).$$

C'est une application bilinéaire continue, de norme égale à celle de  $\ell$ . On obtient bien ainsi une bijection linéaire et préservant la norme (donc bicontinue) de  $\mathcal{L}(E, F; G)$  sur  $\mathcal{L}(E; \mathcal{L}(F; G))$   $\square$

**Définition IV.2** La **différentielle seconde** d'une fonction  $f : U \subset E \rightarrow F$  deux fois différentiable est l'application

$$\begin{aligned} d^2 f : U &\rightarrow \mathcal{L}(E, E; F) \\ x &\mapsto d^2 f_x; \end{aligned}$$

définie par

$$d^2 f_x(h, k) = d(df)_x(h)(k) \quad \text{pour tout } (h, k) \in E \times E.$$

**Remarque.** On peut interpréter cette définition de la façon suivante, qu'on utilise en pratique pour calculer  $d^2 f$ . Si  $f$  est deux fois différentiable sur  $U$ , alors quel que soit  $k \in E$ , l'application

$$\begin{aligned} g : U &\rightarrow F \\ x &\mapsto df_x(k) \end{aligned}$$

est différentiable et

$$dg_x(h) = d^2 f_x(h, k).$$

En effet,  $g$  est la composée de  $df$  et de l'application linéaire continue

$$\begin{aligned} \mathcal{L}(E; F) &\rightarrow F \\ \ell &\mapsto \ell(k). \end{aligned}$$

Une autre interprétation utilise ce qu'on a vu sur la différentielle (première). Comme

$$d(df)_x(h) = \left. \frac{d(df_{x+th})}{dt} \right|_{t=0} \quad \text{et} \quad df_{x+th}(k) = \left. \frac{df(x+th+sk)}{ds} \right|_{s=0},$$

on a

$$d^2 f_x(h, k) = \left. \frac{d}{dt} \frac{d}{ds} f(x+th+sk) \right|_{s=0} \Big|_{t=0}.$$

**Théorème IV.1 (Schwarz)** Si  $f : U \subset E \rightarrow F$  deux fois différentiable en  $x$  alors  $d^2 f_x$  est une application bilinéaire **symétrique**.

La démonstration de ce théorème repose sur le

**Lemme IV.1** À une fonction  $f : U \subset E \rightarrow F$  et à un point  $x \in U$  on associe la fonction (définie sur un voisinage ouvert de  $(0, 0)$ ) :

$$A : (h, k) \mapsto A(h, k) := f(x+h+k) - f(x+h) - f(x+k) + f(x).$$

Si  $f$  est deux fois différentiable en  $x$  alors on a

$$\lim_{(h,k) \rightarrow (0,0)} \frac{\|A(h, k) - d^2 f_x(h, k)\|}{\|h\|^2 + \|k\|^2} = 0.$$

Admettons provisoirement ce lemme, et démontrons le théorème IV.1. La fonction  $A$  est clairement symétrique :  $A(h, k) = A(k, h)$ . Par suite, pour tout  $(h, k) \in E \times E$ ,

$$d^2 f_x(h, k) - d^2 f_x(k, h) = d^2 f_x(h, k) - A(h, k) + A(k, h) - d^2 f_x(k, h).$$

Soit  $\varepsilon > 0$ . Le lemme IV.1 montre donc qu'il existe  $\eta > 0$  tel que  $\|h\| \leq \eta$  et  $\|k\| \leq \eta$  entraîne

$$\|d^2 f_x(h, k) - d^2 f_x(k, h)\| \leq 2\varepsilon (\|h\|^2 + \|k\|^2).$$

Ceci implique, par la bilinéarité de  $d^2 f_x$ ,

$$d^2 f_x(h, k) = d^2 f_x(k, h) \quad \text{pour tout } (h, k) \in E \times E.$$

En effet, pour tout  $(h, k) \in E \times E$ , il existe  $\lambda > 0$  tel que  $\|\lambda h\| \leq \eta$  et  $\|\lambda k\| \leq \eta$ , d'où

$$\|d^2 f_x(\lambda h, \lambda k) - d^2 f_x(\lambda k, \lambda h)\| \leq 2\varepsilon \lambda^2 (\|h\|^2 + \|k\|^2)$$

et donc, en simplifiant par  $\lambda^2$ ,

$$\|d^2 f_x(h, k) - d^2 f_x(k, h)\| \leq 2\varepsilon (\|h\|^2 + \|k\|^2).$$

Cette inégalité étant vraie quel que soit  $\varepsilon > 0$ , on en déduit  $\|d^2 f_x(h, k) - d^2 f_x(k, h)\| = 0$ .  $\square$

**Démonstration du lemme IV.1.** L'ingrédient essentiel est le théorème des accroissements finis. Notons

$$B(h, k) = A(h, k) - d^2 f_x(h, k) = f(x+h+k) - f(x+h) - f(x+k) + f(x) - d^2 f_x(h, k).$$

La fonction  $B$  ainsi définie est différentiable par rapport à sa seconde variable au voisinage de  $(0, 0)$ , et pour tout  $k' \in E$ ,

$$\begin{aligned} d_2 B_{(h,k)}(k') &= df_{(x+h+k)}(k') - df_{(x+k)}(k') - d^2 f_x(h, k') = \\ &df_{(x+h+k)}(k') - df_x(k') - d^2 f_x(h+k, k') + df_x(k') - df_{(x+k)}(k') + d^2 f_x(k, k'). \end{aligned}$$

Soit  $\varepsilon > 0$ . Par définition de  $d^2 f$ , il existe  $\eta > 0$  tel que pour tout  $h' \in E$  tel que  $\|h'\| \leq \eta$  et pour tout  $k' \in E$ ,

$$\|df_{x+h'}(k') - df_x(k') - d^2 f_x(h', k')\| \leq \varepsilon \|h'\| \|k'\|.$$

Pour tout  $(h, k) \in E \times E$  tel que  $\|h\| \leq \eta/2$  et  $\|k\| \leq \eta/2$ , on peut appliquer cette inégalité à  $h' = h+k$  et à  $h' = k$ . On en déduit par l'inégalité triangulaire que

$$\|d_2 B_{(h,k)}(k')\| \leq 2\varepsilon (\|h\| + \|k\|) \|k'\| \quad \text{pour tout } k' \in E,$$

c'est-à-dire

$$\|d_2 B_{(h,k)}\| \leq 2\varepsilon (\|h\| + \|k\|).$$

Le théorème des accroissements finis appliqué à  $y \mapsto B(h, y)$  sur le segment  $[0, k]$  (d'après la remarque qui suit le théorème II.3) montre alors que

$$\|B(h, k)\| = \|B(h, k) - B(h, 0)\| \leq 2\varepsilon (\|h\| + \|k\|) \|k\| \leq 4\varepsilon (\|h\|^2 + \|k\|^2)$$

si  $\|h\| \leq \eta/2$  et  $\|k\| \leq \eta/2$ .  $\square$

**Exemples de différentielles secondes.**

- Une application **affine**  $f : x \mapsto \ell(x) + b$  avec  $\ell \in \mathcal{L}(E; F)$  et  $b \in F$  est deux fois différentiable et sa différentielle seconde est identiquement nulle.
- Une application **quadratique**  $f : x \mapsto \phi(x, x)$  avec  $\phi \in \mathcal{L}(E, E; F)$  est deux fois différentiable et sa différentielle seconde est constante (égale à  $2\phi$  si  $\phi$  est symétrique).

**En dimension finie.** Supposons ici que  $E = \mathbb{R}^p$  et soit  $(e_1, \dots, e_p)$  la base « canonique ». Si  $f$  est deux fois différentiable sur un ouvert  $U \subset \mathbb{R}^p$ , alors on a pour tout  $x \in U$ , pour tous  $i, j \in \{1, \dots, p\}$ ,

$$d_x^2 f(e_i, e_j) = \frac{\partial}{\partial x_i} \frac{\partial f}{\partial x_j}(x).$$

Le théorème de Schwarz montre que « les dérivées partielles croisées sont égales » c'est-à-dire

$$\frac{\partial}{\partial x_i} \frac{\partial f}{\partial x_j} = \frac{\partial}{\partial x_j} \frac{\partial f}{\partial x_i}$$

pour tous  $i, j \in \{1, \dots, p\}$ . Ces dérivées sont en général notées

$$\frac{\partial^2 f}{\partial x_i \partial x_j}.$$

Par bilinéarité, si  $h$  et  $k$  sont deux vecteurs de  $\mathbb{R}^p$  de composantes  $(h_1, \dots, h_p)$  et  $(k_1, \dots, k_p)$  respectivement,

$$d_x^2 f(h, k) = \sum_{i=1}^p \sum_{j=1}^p h_i k_j \frac{\partial^2 f}{\partial x_i \partial x_j}(x).$$

## 2 Différentielle d'ordre $n$

Pour les entiers  $n \geq 2$ , on définit par récurrence les notions suivantes, qui généralisent de façon évidente le cas  $n = 2$ .

**Définition IV.3** Soit une fonction  $f$  définie sur un ouvert (non vide)  $U$  d'un  $\mathbb{R}$ -espace de Banach  $E$  et à valeurs dans un  $\mathbb{R}$ -espace de Banach  $F$ , et  $n$  un entier au moins égal à 2. On dit qu'elle est

- **$n$  fois différentiable** en  $x \in U$  si elle est différentiable dans un voisinage ouvert  $U_x$  de  $x$  et si sa différentielle  $df : U_x \rightarrow \mathcal{L}(E; F)$  est  $(n - 1)$  fois différentiable en  $x$ .
- $n$  fois différentiable dans  $U$  si elle est  $n$  fois différentiable en tout point de  $U$ .
- de classe  $\mathcal{C}^n$  si et seulement si sa différentielle est de classe  $\mathcal{C}^{n-1}$ .
- de classe  $\mathcal{C}^\infty$  si elle est de classe  $\mathcal{C}^n$  pour tout  $n \geq 1$ .

**Exemple.** Les applications linéaires continues et plus généralement les applications  $k$ -linéaires continues sont de classe  $\mathcal{C}^\infty$ .

En effet, le cas  $k = 1$  est clair : si  $u \in \mathcal{L}(E; F)$ , elle est de classe  $\mathcal{C}^1$  et sa différentielle est constante (partout égale à  $u$ ), donc de classe  $\mathcal{C}^\infty$  (les différentielles successives sont toutes identiquement nulles). Voyons le cas  $k = 2$ . Si  $\phi \in \mathcal{L}(E_1, E_2; F)$  alors

$$d\phi_{(x_1, x_2)}(h_1, h_2) = \phi(x_1, h_2) + \phi(h_1, x_2),$$

$$d(d\phi)_{(x_1, x_2)}(h_1, h_2)(k_1, k_2) = \phi(k_1, h_2) + \phi(h_1, k_2),$$

c'est-à-dire que  $d(d\phi)$  est constante. Donc les différentielles suivantes sont toutes identiquement nulles. Le cas général se traite par récurrence ; il est laissé en exercice.

**Remarque.** Comme on peut s'y attendre, la notion de fonction  $n$  fois différentiable est plus forte que celle de fonction  $(n - 1)$  fois différentiable, et la notion de fonction de classe  $\mathcal{C}^n$  est plus forte que celle de fonction de classe  $\mathcal{C}^{n-1}$ . Autrement dit, si  $f$  est  $n$  fois différentiable, elle est *a fortiori*  $(n - 1)$  fois différentiable ; si  $f$  est de classe  $\mathcal{C}^n$ , elle est *a fortiori* de classe  $\mathcal{C}^{n-1}$ . La preuve repose sur une récurrence facile.

Dans l'énoncé suivant,  $E, F$  et  $G$  sont des espaces de Banach,  $U$  est un ouvert de  $E$  et  $V$  est un ouvert de  $F$  contenant  $U$ .

**Théorème IV.2** Si  $f : U \subset E \rightarrow F$  est  $n$  fois différentiable en  $x \in U$  et  $g : V \subset F \rightarrow G$  est  $n$  fois différentiable en  $y = f(x) \in V$ , alors  $g \circ f$  est  $n$  fois différentiable en  $x$ . Si  $f$  est de classe  $\mathcal{C}^n$  sur  $U$  et  $g$  est de classe  $\mathcal{C}^n$  sur  $V$  alors  $g \circ f$  est de classe  $\mathcal{C}^n$ .

**Dém.** Le cas  $n = 1$  résulte du théorème I.1 et du fait que l'application

$$d(g \circ f) : x \mapsto dg_{f(x)} \circ df_x$$

est continue si  $df$  et  $dg$  sont continues, comme composée de

$$x \mapsto (dg_{f(x)}, df_x) \in \mathcal{L}(F; G) \times \mathcal{L}(E; F) \quad \text{et} \quad (v, u) \mapsto v \circ u.$$

(Noter que la seconde application est bilinéaire continue.)

Le cas général se traite par récurrence sur  $n$ . Supposons les propriétés démontrées jusqu'à l'entier  $n - 1$ . Si  $f$  et  $g$  sont  $n$  fois différentiables en  $x$  et  $f(x)$  respectivement, alors  $df$  est  $(n - 1)$  fois différentiable en  $x$ , de même que  $x \mapsto dg_{f(x)}$  comme composée de  $dg$  et  $f$  (en appliquant l'hypothèse de récurrence). En observant que l'application bilinéaire continue

$$(v, u) \in \mathcal{L}(F; G) \times \mathcal{L}(E; F) \mapsto v \circ u$$

est de classe  $\mathcal{C}^\infty$  et en appliquant à nouveau l'hypothèse de récurrence, on en déduit que  $d(g \circ f)$  est  $(n - 1)$  fois différentiable en  $x$ . Ceci signifie précisément que  $g \circ f$  est  $n$  fois différentiable en  $x$ . Le raisonnement est identique pour montrer que  $g \circ f$  est de classe  $\mathcal{C}^n$  si  $f$  et  $g$  le sont.  $\square$

**Théorème IV.3** Si  $f$  est un difféomorphisme de  $U$  sur  $V$  et si  $f$  est de classe  $\mathcal{C}^n$  alors  $f^{-1}$  est aussi de classe  $\mathcal{C}^n$ .

**Dém.** L'énoncé est vide si  $n = 1$ . Le cas  $n \geq 2$  se traite par récurrence, en utilisant la formule

$$d(f^{-1})_y = (df_{f^{-1}(y)})^{-1}$$

et le

**Lemme IV.2** L'application

$$\begin{array}{ccc} \text{Isom}(E; F) & \rightarrow & \text{Isom}(F; E) \\ u & \mapsto & u^{-1} \end{array}$$

est de classe  $\mathcal{C}^\infty$ .

Supposons que la réciproque de tout difféomorphisme de classe  $\mathcal{C}^{n-1}$  est de classe  $\mathcal{C}^{n-1}$ . Soit  $f$  un difféomorphisme de classe  $\mathcal{C}^n$ . Alors par définition,  $df$  est de classe  $\mathcal{C}^{n-1}$ , et d'après l'hypothèse de récurrence  $f^{-1}$  est de classe  $\mathcal{C}^{n-1}$  (au moins). Donc le théorème IV.2 montre que l'application  $y \mapsto df_{f^{-1}(y)}$  est de classe  $\mathcal{C}^{n-1}$ , et grâce au lemme IV.2, le même théorème montre que  $y \mapsto (df_{f^{-1}(y)})^{-1}$  est de classe  $\mathcal{C}^{n-1}$ .  $\square$

**Dém. du lemme IV.2** On sait déjà que l'application  $\Theta : u \mapsto u^{-1}$  est différentiable et

$$d\Theta_u(h) = -\Theta(u) \circ h \circ \Theta(u) \quad \text{c'est-à-dire} \quad d\Theta_u = \phi(\Theta(u), \Theta(u)),$$

où  $\phi$  est l'application bilinéaire continue

$$\begin{aligned} \phi : \mathcal{L}(F; E) \times \mathcal{L}(F; E) &\rightarrow \mathcal{L}(\mathcal{L}(E; F); \mathcal{L}(F; E)) \\ (v, w) &\mapsto (h \mapsto -v \circ h \circ w) \end{aligned}$$

L'application  $\phi$  est en fait de classe  $\mathcal{C}^\infty$ . Le théorème IV.2 montre donc  $\Theta$  est de classe  $\mathcal{C}^\infty$  (récurrence immédiate).  $\square$

On va maintenant définir la différentielle d'ordre  $n$  de façon à généraliser la notion de différentielle seconde. Si  $f$  est  $n$  fois différentiable en  $x$ , sa différentielle d'ordre  $n$  au point  $x$  sera une application  $n$ -linéaire symétrique.

**Définition IV.4** Soient  $E$  et  $F$  des  $\mathbb{R}$ -espaces de Banach, et  $\mathcal{L}_n(E; F)$  l'espace des applications  $n$ -linéaires continues sur  $E^n$ . Une application  $\phi \in \mathcal{L}_n(E; F)$  est dite **symétrique** si pour toute permutation  $\sigma \in \Sigma_n$  et pour tout  $n$ -uplet  $(x_1, \dots, x_n) \in E^n$ ,

$$\phi(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = \phi(x_1, \dots, x_n).$$

On notera  $\mathcal{L}_n^s(E; F)$  l'espace des applications  $n$ -linéaires continues et symétriques sur  $E^n$ .

**Remarque.** Pour vérifier qu'une application  $n$ -linéaire est symétrique, il suffit de vérifier son invariance par les **transpositions** qui échangent  $i$  et  $i+1$  avec  $i \leq n-1$  et par celle qui échange 1 et  $n$ .

**Théorème IV.4** Une fonction  $f : U \subset E \rightarrow F$  est  $n$  fois différentiable au point  $x \in U$  si et seulement s'il existe un voisinage ouvert  $U_x$  de  $x$  dans  $U$ , des fonctions  $d^p f : U_x \rightarrow \mathcal{L}_p^s(E; F)$  pour  $p \leq n-1$  et  $d^n f_x \in \mathcal{L}_n^s(E; F)$  telles que  $d^1 f = df$  dans  $U_x$ , pour tout  $p \leq n-2$ ,  $d^p f$  est différentiable sur  $U_x$ , avec pour tout  $y \in U_x$  et pour tout  $(h_1, \dots, h_{p+1}) \in E^{p+1}$  :

$$d^{p+1} f_y(h_1, \dots, h_{p+1}) = d_{p+1} g_y^{[p]}(h_{p+1}) \quad \text{où} \quad g^{[p]}(h_1, \dots, h_p, y) := d^p f_y(h_1, \dots, h_p),$$

et enfin  $d^{n-1} f$  est différentiable en  $x$  et

$$d^n f_x(h_1, \dots, h_n) = d_n g_x^{[n-1]}(h_n) \quad \text{où} \quad g^{[n-1]}(h_1, \dots, h_{n-1}, y) := d^{n-1} f_y(h_1, \dots, h_{n-1}).$$

**Dém.** Le cas  $n=1$  est évident. Le cas  $n=2$  résulte de la définition IV.2 de la différentielle seconde et du théorème de Schwarz IV.1. Le reste de la démonstration se fait bien sûr par récurrence.

Par souci de clarté, voyons le cas  $n=3$ . Il y a deux sens à démontrer. Supposons que  $f$  soit 3 fois différentiable en  $x$  : d'après la définition IV.3, il existe un voisinage ouvert  $U_x$  de  $x$  telle que  $df$  est définie sur  $U_x$  et deux fois différentiable en  $x$ .

En particulier,  $df$  est différentiable sur  $U_x$ , ce qui montre que  $f$  est deux fois différentiable sur  $U_x$  (d'après la définition IV.2), et sa différentielle seconde  $d^2 f$  est une application de  $U_x$  dans  $\mathcal{L}_2^s(E; F)$  d'après le théorème de Schwarz IV.1, telle que (voir la remarque suivant la définition IV.2) :

$$d^2 f_y(h_1, h_2) = d_2 g_y^{[1]}(h_2) \quad \text{où} \quad g^{[1]}(h_1, y) = df_y(h_1).$$

On aura montré les propriétés voulues pour  $n = 3$  si l'on prouve que  $d^2f$  est différentiable en  $x$  et que

$$d^3f_x(h_1, h_2, h_3) = d_3g_x(h_3) \quad \text{avec} \quad g^{[2]}(h_1, h_2, y) = d^2f_y(h_1, h_2)$$

définit bien  $d^3f_x \in \mathcal{L}_3^s(E; F)$ .

Or, pour tout  $y \in U_x$ , on a :

$$d^2f_y = \mathcal{I}(d(df)_y),$$

où  $\mathcal{I}$  désigne l'isomorphisme naturel :

$$\begin{aligned} \mathcal{I} : \mathcal{L}(E; \mathcal{L}(E; F)) &\rightarrow \mathcal{L}_2(E; F) \\ \ell &\mapsto ((h, k) \mapsto \ell(h)(k)). \end{aligned}$$

Comme  $df$  est deux fois différentiable en  $x$  (par hypothèse), on en déduit par composition que  $d^2f$  est différentiable en  $x$  et (en se rappelant que  $\mathcal{I}$  est linéaire donc de différentielle constamment égale à  $\mathcal{I}$ ) :

$$d(d^2f)_x = \mathcal{I} \circ d(d(df))_x.$$

D'un point de vue un peu différent, on a pour tout  $y \in U_x$  :

$$g^{[2]}(h_1, h_2, y) = d^2f_y(h_1, h_2) = \mathcal{E}_{(h_1, h_2)}(d(df)_y),$$

où

$$\begin{aligned} \mathcal{E}_{(h_1, h_2)} : \mathcal{L}(E; \mathcal{L}(E; F)) &\rightarrow F \\ \ell &\mapsto \ell(h_1)(h_2). \end{aligned}$$

Par dérivation de fonctions composées on a donc

$$d_3g_x^{[2]}(h_3) = (d(d(df)))_x(h_3)(h_1)(h_2).$$

Dans cette dernière expression, on peut échanger  $h_1$  et  $h_2$  d'après le théorème de Schwarz appliqué à  $f$  (et qui montre que  $g(y) = \mathcal{E}_{(h_2, h_1)}(d(df)_y)$ ), et on peut échanger  $h_3$  et  $h_1$  d'après le théorème de Schwarz appliqué à  $df$ .

Inversement, si  $f$  satisfait les propriétés énoncées pour  $n = 3$ , alors  $f$  est deux fois différentiable,  $df$  est différentiable sur  $U_x$ , et  $d^2f$  est différentiable en  $x$ . Ceci implique que  $df$  est deux fois différentiable en  $x$ , car en notant  $\mathcal{J} = \mathcal{I}^{-1}$  l'isomorphisme :

$$\begin{aligned} \mathcal{J} : \mathcal{L}_2(E; F) &\rightarrow \mathcal{L}(E; \mathcal{L}(E; F)) \\ \phi &\mapsto (h \mapsto \phi(h, \cdot)). \end{aligned}$$

on a  $d(df)_y = \mathcal{J}(d^2f_y)$  pour tout  $y \in U_x$  et donc  $d(df)$  est différentiable comme fonction composée de fonctions différentiables.

Ceci termine le cas  $n = 3$ . Le passage de  $n$  à  $n + 1$  se traite de façon analogue, il est laissé en exercice.  $\square$



# Chapitre V

## Formules de Taylor

### 1 Formule de Taylor avec reste intégral

Les formules de Taylor de type **global** (avec reste intégral ou de Lagrange) sont fondées sur un calcul simple concernant les fonctions d'une variable réelle.

**Lemme V.1** Soit  $I$  un intervalle ouvert de  $\mathbb{R}$ ,  $F$  un espace de Banach et  $g : I \rightarrow F$  une fonction  $(n + 1)$  fois dérivable. On note  $g^{(p)}$  ses dérivées successives,  $p \in \{1, \dots, n + 1\}$ . Alors pour tout  $t \in I$  on a :

$$\frac{d}{dt} \left( g(t) + \sum_{p=1}^n \frac{(1-t)^p}{p!} g^{(p)}(t) \right) = \frac{(1-t)^n}{n!} g^{(n+1)}(t).$$

**Dém.** En dérivant terme à terme, on trouve :

$$\begin{aligned} \frac{d}{dt} \left( g'(t) + \sum_{p=1}^n \frac{(1-t)^p}{p!} g^{(p+1)}(t) - \sum_{p=1}^n \frac{(1-t)^{p-1}}{(p-1)!} g^{(p)}(t) \right) \\ = \sum_{p=0}^n \frac{(1-t)^p}{p!} g^{(p+1)}(t) - \sum_{p=0}^{n-1} \frac{(1-t)^p}{p!} g^{(p+1)}(t) \end{aligned}$$

par un simple changement d'indice. Tous les termes se compensent deux à deux sauf le dernier de la première somme.  $\square$

En fait, on va obtenir la formule de Taylor avec reste intégral grâce à une “version intégrale” du lemme V.1. Avant cela, il faut s'accorder sur la notion d'intégrale pour les fonctions à valeurs dans  $F$ . Comme on n'intégrera que des fonctions continues, la notion d'intégrale au sens de Riemann nous suffira. Résumons-en les propriétés importantes.

**Proposition V.1** Soient  $a, b \in \mathbb{R}$ ,  $a \leq b$ , et  $F$  un espace de Banach. Alors l'intégrale de Riemann sur le segment  $[a, b]$  définit une application linéaire continue sur l'espace  $\mathcal{C}([a, b]; F)$  des fonctions continues sur  $[a, b]$  et à valeurs dans  $F$ , muni de la norme sup. Pour tout  $g \in \mathcal{C}([a, b]; F)$ , l'intégrale de Riemann de  $g$  sur le segment  $[a, b]$ , notée  $\int_a^b g(t) dt$ , vérifie l'inégalité :

$$\left\| \int_a^b g(t) dt \right\| \leq \int_a^b \|g(t)\| dt \leq (b-a) \max_{t \in [a, b]} \|g(t)\|.$$

Quels que soient  $x, y \in [a, b]$ , on a par définition

$$\int_x^y g(t) dt = - \int_y^x g(t) dt$$

(et donc  $\int_x^x g(t) dt = 0$ ), et

$$\int_x^y g(t) dt = \int_x^z g(t) dt + \int_z^y g(t) dt$$

pour tout  $z \in [a, b]$ . De plus, l'application  $y \mapsto \int_x^y g(t) dt$  est dérivable et sa dérivée est  $g$ . Inversement, pour toute primitive  $G$  de  $g$ , on a

$$G(y) - G(x) = \int_x^y g(t) dt.$$

**Remarque.** Si  $F$  est de dimension finie,  $F = \mathbb{R}^q$  pour simplifier,  $\int_a^b g(t) dt$  est simplement le vecteur dont les composantes sont  $\int_a^b g_i(t) dt$  où les  $g_i$  pour  $i \in \{1, \dots, n\}$  sont les composantes de  $g$  dans la base « canonique ».

On a ainsi une conséquence immédiate du lemme V.1 :

**Corollaire V.1** Si  $I$  est un intervalle ouvert de  $\mathbb{R}$  contenant  $[0, 1]$ ,  $F$  un espace de Banach et  $g : I \rightarrow F$  une fonction de classe  $\mathcal{C}^{n+1}$ , alors

$$(1.1) \quad g(1) - g(0) - \sum_{p=1}^n \frac{1}{p!} g^{(p)}(0) = \int_0^1 \frac{(1-t)^n}{n!} g^{(n+1)}(t) dt.$$

Le corollaire V.1 a lui-même un corollaire pour les fonctions de classe  $\mathcal{C}^{n+1}$  sur un ouvert d'un espace de Banach.

**Notation.** Pour tout  $h \in E$  et  $n \in \mathbb{N}^*$ , on désigne par  $h^{[n]}$  le  $n$ -uplet de vecteurs tous égaux à  $h$ .

**Théorème V.1** Si  $U$  est un ouvert d'un espace de Banach  $E$ , si  $F$  est un espace de Banach et  $f : U \rightarrow F$  est une fonction de classe  $\mathcal{C}^{n+1}$ , alors, pour tout  $(x, h) \in U \times E$  tel que le segment  $[x, x+h]$  soit inclus dans  $U$ ,

$$(1.2) \quad f(x+h) = f(x) + \sum_{p=1}^n \frac{1}{p!} d^p f_x(h^{[p]}) + \int_0^1 \frac{(1-t)^n}{n!} d^{n+1} f_{x+th}(h^{[n+1]}) dt.$$

**Dém.** Appliquer la formule (1.1) à la fonction

$$g : t \mapsto g(t) := f(x + th).$$

□

La formule (1.2) est ce qu'on appelle la *formule de Taylor avec reste intégral* à l'ordre  $n+1$ . En particulier, on a à l'ordre 1 :

$$f(x+h) = f(x) + \int_0^1 df_{x+th}(h) dt.$$

On peut remarquer que cette formule implique l'inégalité des accroissements finis, sous une hypothèse cependant plus forte ( $f$  de classe  $\mathcal{C}^1$ ) que dans le théorème des accroissements finis vu au chapitre II. (On verra ci-après une généralisation à l'ordre  $n + 1$  de l'inégalité des accroissements finis, connue sous le nom de formule de Taylor-Lagrange, qui ne demande que  $f$  soit  $(n + 1)$  différentiable.) La formule à l'ordre 2 (c'est-à-dire (1.2) avec  $n = 1$ ) s'écrit quant à elle :

$$f(x + h) = f(x) + df_x(h) + \int_0^1 (1 - t) d^2 f_{x+th}(h, h) dt.$$

## 2 Formule de Taylor-Lagrange

Une autre façon d'exploiter le lemme V.1 est contenue dans la

**Proposition V.2** *Si  $I$  est un intervalle ouvert de  $\mathbb{R}$  contenant  $[0, 1]$ ,  $F$  un espace de Banach et  $g : I \rightarrow F$  une fonction  $(n + 1)$  fois différentiable telle que*

$$\|g^{(n+1)}(t)\| \leq M, \quad \text{pour tout } t \in [0, 1],$$

alors

$$(2.3) \quad \left\| g(1) - g(0) - \sum_{p=1}^n \frac{1}{p!} g^{(p)}(0) \right\| \leq \frac{M}{(n+1)!}.$$

**Dém.** Appliquer le théorème II.2 (version raffinée du théorème des accroissements finis) à

$$f(t) := g(t) + \sum_{p=1}^n \frac{(1-t)^p}{p!} g^{(p)}(t) \quad \text{et} \quad \varphi(t) := -M \frac{(1-t)^{n+1}}{(n+1)!}.$$

Le lemme V.1 montre en effet que

$$\|f'(t)\| \leq \varphi'(t) \quad \text{pour tout } t.$$

□

Cette proposition se traduit pour les fonctions  $(n + 1)$  fois différentiables sur un ouvert d'espace de Banach par le

**Théorème V.2** *Si  $U$  est un ouvert d'un espace de Banach  $E$ , si  $F$  est un espace de Banach, si  $(x, h) \in U \times E$  est tel que le segment  $[x, x + h]$  soit inclus dans  $U$ , et si  $f : U \rightarrow F$  est une fonction  $(n + 1)$  fois différentiable telle que*

$$\max_{y \in [x, x+h]} \|d^{n+1} f_y\|_{\mathcal{L}_{n+1}(E;F)} \leq M,$$

alors

$$(2.4) \quad \left\| f(x + h) - f(x) - \sum_{p=1}^n \frac{1}{p!} d^p f_x(h^{[p]}) \right\| \leq \frac{M}{(n+1)!} \|h\|^{n+1}.$$

**Dém.** Appliquer la proposition 2.3 à la fonction

$$g : t \mapsto g(t) := f(x + th).$$

□

L'inégalité (2.4), qui généralise l'inégalité des accroissements finis, est connue sous le nom de formule de *Taylor avec reste de Lagrange*, le reste en question étant seulement connu à travers une majoration, contrairement au reste intégral, qui est *exact*. Lorsque  $f$  est de classe  $\mathcal{C}^{n+1}$ , c'est une conséquence immédiate de la formule de Taylor avec reste intégral.

### 3 Formule de Taylor-Young

Voyons une dernière version de la formule de Taylor, valable sous des hypothèses encore moins fortes, et qui pour cette raison donne un résultat *local* seulement.

**Théorème V.3** *Si  $U$  est un ouvert d'un espace de Banach  $E$ , si  $F$  est un espace de Banach, si  $f : U \rightarrow F$  est une fonction  $n$  fois différentiable en  $x \in U$  alors*

$$(3.5) \quad \left\| f(x+h) - f(x) - \sum_{p=1}^n \frac{1}{p!} d^p f_x(h^{[p]}) \right\| = o(\|h\|^n).$$

Dans cet énoncé la notation de Landau  $o$  (qu'on lit « petit  $o$  de ») signifie que le membre de gauche divisé par  $\|h\|^n$  tend vers 0 lorsque  $h$  tend vers 0. Il s'agit donc d'un résultat **local**, qui donne des renseignements sur le comportement de  $f$  au voisinage de  $x$  seulement.

**Dém.** Il n'y a rien à faire dans le cas  $n = 1$  : c'est la définition de la différentiabilité, reformulée à l'aide de  $o$ . On procède ensuite par récurrence. Soit  $n \geq 2$ . On suppose le théorème démontré pour les fonctions  $(n-1)$  fois différentiables en un point. Si  $f$  est  $n$  fois différentiable en  $x$ , considérons la fonction

$$F : h \mapsto f(x+h) - f(x) - \sum_{p=1}^n \frac{1}{p!} d^p f_x(h^{[p]}).$$

Alors  $F$  est différentiable au voisinage de  $0_E$  et

$$\begin{aligned} dF_h(k) &= df_{x+h}(k) - df_x(k) - \sum_{p=2}^n \frac{p}{p!} d^p f_x(h^{[p-1]}, k) \\ &= df_{x+h}(k) - df_x(k) - \sum_{p=1}^{n-1} \frac{1}{p!} d^p (df)_x(h^{[p]})(k) \end{aligned}$$

quel que soit  $k \in E$ . L'hypothèse de récurrence appliquée à la différentielle de  $f$  montre que

$$\|dF_h\| = o(\|h\|^{n-1}).$$

Ceci signifie que pour tout  $\varepsilon > 0$ , il existe  $\eta > 0$  tel que  $\|h\| \leq \eta$  implique

$$\|dF_h\| \leq \varepsilon \|h\|^{n-1}.$$

Ainsi  $\|dF_y\| \leq \varepsilon \|h\|^{n-1}$  pour tout  $y$  appartenant au segment  $[0, h]$ . Grâce au théorème des accroissements finis on en déduit :

$$\|F(h)\| = \|F(h) - F(0)\| \leq \varepsilon \|h\| \|h\|^{n-1},$$

c'est-à-dire  $\|F(h)\| = o(\|h\|^n)$ . □

**Remarque.** Lorsque  $f$  a un cran de plus de régularité, et plus précisément lorsqu'elle est de classe  $\mathcal{C}^{n+1}$ , la formule de Taylor-Young (3.5) est une conséquence immédiate de la formule de Taylor-Lagrange (2.4); dans ce cas le  $o(\|h\|^n)$  est carrément un  $O(\|h\|^{n+1})$  (qu'on lit « grand o de » ...).

La série (formelle)

$$\sum_p \frac{1}{p!} d^p f_x(h^{[p]})$$

est souvent appelée **série de Taylor** de  $f$  au point  $x$  : attention, rien ne dit qu'elle converge *a priori* ! (C'est le cas si  $f$  est analytique.)

La somme finie

$$f(x) + \sum_{p=1}^n \frac{1}{p!} d^p f_x(h^{[p]})$$

est le **développement limité d'ordre  $n$**  de  $f$  au point  $x$ .



# Chapitre VI

## Extrema

Dans ce chapitre on considère exclusivement des fonctions à **valeurs réelles**, et on s'intéresse à leurs **extrema**, c'est-à-dire à leurs **minima** et **maxima**. On parlera en fait seulement de minima pour simplifier : les maxima d'une fonction  $f$  peuvent en effet être vus comme les minima de  $-f$ .

### 1 Extrema libres

**Définition VI.1** Si  $f$  est une fonction définie sur une partie  $D$  d'un espace de Banach  $E$  et à valeurs réelles, un point  $a \in U$  est un **minimum local** de  $f$  s'il existe un voisinage  $V_a$  de  $a$ , ouvert dans  $D$ , tel que

$$f(x) \geq f(a) \quad \text{pour tout } x \in V_a.$$

On dira que  $a$  est un **minimum global** de  $f$  si

$$f(x) \geq f(a) \quad \text{pour tout } x \in U.$$

Un minimum est dit **strict** si l'inégalité est stricte, c'est-à-dire  $f(x) > f(a)$ , pour  $x \neq a$ .

L'objectif ici est de dégager des conditions nécessaires et/ou suffisantes pour avoir un minimum local selon le degré de différentiabilité de  $f$ .

Commençons par rappeler ce que l'on sait dans le cas  $E = \mathbb{R}$ .

**Proposition VI.1** Soit  $g$  une fonction définie sur un intervalle ouvert  $I$  de  $\mathbb{R}$  et à valeurs dans  $\mathbb{R}$ , dérivable en  $a \in I$ . Si  $a$  est un minimum local de  $g$  alors  $g'(a) = 0$ . Si de plus  $g$  est deux fois dérivable en  $a$ , alors  $g''(a) \geq 0$ . « Inversement » si  $b \in I$  est tel que  $g'(b) = 0$  et  $g''(b) > 0$  alors  $b$  est un minimum local de  $g$ .

**Dém.** Par définition de la dérivabilité,

$$g(t) - g(a) - (t - a)g'(a) = \varepsilon(t)(t - a)$$

avec  $\lim_{t \rightarrow a} \varepsilon(t) = 0$ . Si  $g'(a) \neq 0$ , supposons par exemple  $g'(a) > 0$ , alors il existe  $\eta > 0$  tel que pour  $|t - a| \leq \eta$ ,  $|\varepsilon(t)| \leq \frac{1}{2}g'(a)$ , d'où

$$g(t) - g(a) = (g'(a) - \varepsilon(t))(t - a) \leq \frac{1}{2}g'(a)(t - a) < 0$$

pour  $a - \eta \leq t < a$ . Donc  $a$  ne peut pas être un minimum local.

Si  $g$  est deux fois dérivable, supposons que  $a$  soit un minimum local de  $g$ . Alors  $g'(a) = 0$  d'après ce qui précède. Supposons  $g''(a) < 0$ . Comme d'après la formule de Taylor-Young

$$g(t) - g(a) - \frac{1}{2}(t-a)^2 g''(a) = \varepsilon(t) (t-a)^2$$

avec  $\lim_{t \rightarrow a} \varepsilon(t) = 0$ , il existe  $\eta > 0$  tel que pour  $|t-a| \leq \eta$ ,  $|\varepsilon(t)| \leq -\frac{1}{4}g''(a)$ , d'où

$$g(t) - g(a) = (\frac{1}{2}g''(a) - \varepsilon(t)) (t-a)^2 \leq \frac{1}{4}g''(a) (t-a)^2 < 0$$

pour  $|t-a| \leq \eta$ ,  $t \neq a$ . Donc  $a$  ne peut pas être un minimum local.

Enfin, si  $g'(b) = 0$  et  $g''(b) > 0$  alors

$$g(t) - g(b) - \frac{1}{2}(t-b)^2 g''(b) = \varepsilon(t) (t-b)^2$$

avec  $\lim_{t \rightarrow b} \varepsilon(t) = 0$ , et donc il existe  $\eta > 0$  tel que pour  $|t-b| \leq \eta$ ,  $|\varepsilon(t)| \leq \frac{1}{4}g''(b)$ , d'où

$$g(t) - g(b) = (\frac{1}{2}g''(b) - \varepsilon(t)) (t-b)^2 \geq \frac{1}{4}g''(b) (t-b)^2 \geq 0$$

pour  $|t-b| \leq \eta$ . □

**Remarque VI.1** Attention, les conditions  $g'(a) = 0$  et  $g''(a) \geq 0$  ne sont évidemment pas suffisantes (ex :  $g(t) = t^3$  en  $t = 0$ ) et la condition  $g''(b) > 0$  n'est pas nécessaire (ex :  $g(t) = t^4$  en  $t = 0$ )!

Les conditions de la proposition VI.1 s'étendent aux fonctions définies sur un ouvert d'espace de Banach.

**Théorème VI.1** Soit  $f$  une fonction définie sur un ouvert  $U$  d'un espace de Banach  $E$  et à valeurs réelles, différentiable en  $a \in U$ . Si  $a$  est un minimum local de  $f$  alors  $df_a = 0$ . Si de plus  $f$  est deux fois différentiable en  $a$ , alors  $d^2f_a(h, h) \geq 0$  pour tout  $h \in E$ . « Inversement » si  $b \in U$  est tel que  $df_b = 0$  et il existe  $C > 0$  avec  $d^2f_b(h, h) \geq C \|h\|^2$  pour tout  $h \in E$  alors  $b$  est un minimum local de  $f$ .

**Dém.** Les conditions nécessaires sont des conséquences immédiates de la proposition VI.1. En effet, si  $a$  est un minimum local de  $f$  alors, quel que soit  $h \in E$ ,  $0$  est un minimum local de la fonction d'une variable réelle  $g : t \mapsto g(t) := f(a + th)$ . Or  $g'(a) = df_a(h)$  et  $g''(a) = d^2f_a(h, h)$ . Pour les conditions suffisantes, on applique la formule de Taylor-Young à  $f$ . On a en effet

$$f(b+h) - f(b) - \frac{1}{2}df_b(h, h) = \varepsilon(h) \|h\|^2$$

avec  $\lim_{h \rightarrow 0} \varepsilon(h) = 0$ . Donc il existe  $\eta > 0$  tel que pour  $\|h\| \leq \eta$ ,  $|\varepsilon(h)| \leq \frac{C}{4}$ , d'où

$$f(b+h) - f(b) \geq \frac{C}{4} \|h\|^2 \geq 0$$

pour  $\|h\| \leq \eta$ . □

**Remarque VI.2** En dimension finie, l'existence de  $C > 0$  tel que  $d^2f_b(h, h) \geq C \|h\|^2$  pour tout vecteur  $h \in E$  équivaut à  $d^2f_b(h, h) > 0$  quel que soit  $h \neq 0_E$ . Il suffit en effet de remarquer que la fonction continue  $h \mapsto d^2f_b(h, h)$  atteint son minimum sur la sphère unité (qui est compacte si  $E$  est de dimension finie). Par bilinéarité de  $d^2f_b$  on en déduit l'inégalité voulue avec  $C := \min_{\|h\|=1} d^2f_b(h, h)$ . De plus, pour avoir  $d^2f_b(h, h) > 0$  quel que soit  $h \neq 0_E$ , il faut et il suffit que la **matrice hessienne** de  $f$  en  $b$ , c'est-à-dire la matrice symétrique réelle de coefficients  $\frac{\partial^2 f}{\partial x_j \partial x_j}(b)$  ait ses valeurs propres toutes strictement positives. On sait en effet que toute matrice symétrique réelle est diagonalisable sur  $\mathbb{R}$  : une démonstration de cette propriété utilise précisément la notion d'extremum lié...

## 2 Extrema liés

**Définition VI.2** Si  $f$  et  $g_1, \dots, g_p$  sont des fonctions définies sur un ouvert  $U$  d'un espace de Banach  $E$  et à valeurs réelles, un point  $a \in U$  tel que  $g_1(a) = 0 \dots g_p(a) = 0$  est un **minimum local de  $f$  sous les contraintes**  $g_1, \dots, g_p$  s'il existe un voisinage ouvert  $V_a$  de  $a$  tel que

$$f(x) \geq f(a) \quad \text{pour tout } x \in V_a \text{ tel que } g_1(x) = 0 \dots g_p(x) = 0.$$

On va obtenir ici une condition nécessaire pour qu'un point soit un minimum local sous contraintes lorsque les fonctions  $f$  et  $g_1, \dots, g_p$  sont continûment différentiables.

On dira que les contraintes  $g_1, \dots, g_p$  sont **indépendantes** au point  $a \in U$  si la famille de formes linéaires continues  $\{(dg_1)_a, \dots, (dg_p)_a\}$  est libre.

**Théorème VI.2** Soient  $f$  et  $g_1, \dots, g_p$  des fonctions de classe  $\mathcal{C}^1$  sur un ouvert  $U$  d'un espace de Banach  $E$  et à valeurs réelles. Soit  $a \in U$  tel que  $g_1(a) = 0, \dots, g_p(a) = 0$  et les contraintes  $g_1, \dots, g_p$  sont indépendantes au point  $a$ . Si  $a$  est un minimum local de  $f$  sous les contraintes  $g_1, \dots, g_p$  alors il existe des réels  $\lambda_1, \dots, \lambda_p$  tels que

$$df_a = \lambda_1 (dg_1)_a + \dots + (dg_p)_a.$$

Dans cet énoncé les nombres  $\lambda_1, \dots, \lambda_p$  sont appelés des **multiplicateurs de Lagrange**.

**Dém.** Grâce au théorème des fonctions implicites, on va se ramener au cas d'un minimum libre. On peut supposer sans perte de généralité  $a = 0$ . (Il suffit de considérer la fonction  $x \mapsto f(x - a)$  au lieu de  $f$ .) Notons pour simplifier  $\psi_i := (dg_i)_0$  pour tout  $i \in \{1, \dots, p\}$ . Soit

$$G = \left( \text{Vect}(\psi_1, \dots, \psi_p) \right)^\perp = \{h \in E; \psi_i(h) = 0, i \in \{1, \dots, p\}\}.$$

Puisque la famille  $(\psi_1, \dots, \psi_p)$  est libre, il existe une famille de  $p$  vecteurs  $h_1, \dots, h_p \in E$  indépendants tels que  $\psi_i(h_j) = \delta_i^j$  (symbole de Kronecker, valant 1 si  $i = j$  et 0 sinon). Alors le sous-espace  $F = \text{Vect}(h_1, \dots, h_p)$  est tel que  $G \oplus F = E$ , c'est-à-dire que pour tout  $x \in E$  il existe un unique couple  $(z, y) \in G \times F$  tel que  $x = z + y$ . On a donc un isomorphisme

$$\begin{aligned} \mathcal{J} : G \times F &\rightarrow E \\ (z, y) &\mapsto x = z + y. \end{aligned}$$

(La continuité de la réciproque découle du théorème général III.4, mais aussi de la formule explicite :  $y = \sum_{i=1}^p \psi_i(x) h_i$ .) Considérons alors la fonction

$$\begin{aligned} G \times F &\rightarrow \mathbb{R}^p \\ (z, y) &\mapsto (g_1(z + y), \dots, g_p(z + y)). \end{aligned}$$

C'est une fonction de classe  $\mathcal{C}^1$  (comme fonction composée de fonctions de classe  $\mathcal{C}^1$ ) et sa différentielle partielle par rapport à  $y$  au point  $(0, 0)$  est un isomorphisme de  $F$  sur  $\mathbb{R}^p$ , par hypothèse sur les fonctions  $g_i$  : en effet, dans la base  $(h_1, \dots, h_p)$  de  $F$  et la base canonique de  $\mathbb{R}^p$ , sa matrice jacobienne est la matrice de coefficients  $\psi_i(h_j) = \delta_i^j$ , c'est-à-dire la matrice identité !

Donc le théorème des fonctions implicites montre qu'il existe un voisinage  $V_0$  de 0 dans  $U$  (image par  $\mathcal{J}$  d'un voisinage de  $(0, 0)$  dans  $G \times F$ ), et une application  $\varphi$  définie sur un voisinage  $W_0$  de 0 dans  $G$  tels que

$$g_1(x) = 0, \dots, g_p(x) = 0 \quad x \in V_0 \quad \Leftrightarrow \quad x = z + \varphi(z).$$

Par conséquent, 0 est un minimum local de  $f$  sous les contraintes  $g_1, \dots, g_p$  si et seulement si 0 est un minimum local de la fonction  $g : z \mapsto g(z) := f(z + \varphi(z))$ . Une condition nécessaire est donc  $dg_0 = 0$ , c'est-à-dire

$$df_0(k + d\varphi_0(k)) = 0 \quad \text{pour tout } k \in G.$$

Or par construction de  $\varphi$  on a précisément  $d\varphi_0(k) = 0$  pour tout  $k \in G$ . En effet, comme  $\varphi$  est à valeurs dans l'espace vectoriel  $F = \text{Vect}(h_1, \dots, h_p)$ , il suffit de montrer que  $\psi_i(d\varphi_0(k)) = 0$  quel que soit  $i \in \{1, \dots, p\}$  : ceci se déduit par différentiation de la fonction

$$z \mapsto g_i(z + \varphi(z)),$$

qui est identiquement nulle, en utilisant le fait que  $(dg_i)_0(k) = \psi_i(k) = 0$  pour  $k \in G$  (par définition de  $G$  !).

On a donc montré que  $df_0(k) = 0$  pour tout  $k \in G$ . Autrement dit, en notant pour simplifier  $\psi = df_0$ , on a

$$\left(\text{Vect}(\psi_1, \dots, \psi_p)\right)^\perp \subset \left(\text{Vect}(\psi)\right)^\perp.$$

On en déduit, grâce au lemme algébrique classique rappelé ci-après :

$$\text{Vect}(\psi) \subset \text{Vect}(\psi_1, \dots, \psi_p)$$

c'est-à-dire que  $\psi$  est effectivement une combinaison linéaire des  $\psi_i$ . □

**Lemme VI.1** Soient  $\psi, \psi_1, \dots, \psi_p$ , des formes linéaires sur un espace vectoriel  $E$ . Si

$$\bigcap_{i=1}^p \text{Ker } \psi_i \subset \text{Ker } \psi$$

alors  $\psi$  est combinaison linéaire des  $\psi_i$ .

La démonstration est laissée en exercice.

Jusqu'à présent, nous avons considéré des problèmes d'extremum essentiellement sur des ouverts : les conditions nécessaires d'extremum local (dans la proposition VI.1 et les théorèmes VI.1, VI.2) sont fausses lorsque  $U$  n'est pas un ouvert. Nous allons maintenant considérer des problèmes d'extremum sur des sous-ensembles convexes de  $E$  (la convexité étant une propriété géométrique et non topologique).

### 3 Convexité et minima

**Définition VI.3** Un sous-ensemble  $C$  d'un  $\mathbb{R}$ -espace vectoriel  $E$  est dit **convexe** si pour tous  $x, y \in C$ , pour tout  $\theta \in [0, 1]$ ,  $\theta x + (1 - \theta)y \in C$ . Une fonction  $f$  définie sur un convexe  $C$  et à valeurs dans  $\mathbb{R}$  est dite **convexe** si pour tous  $x, y \in C$ , pour tout  $\theta \in [0, 1]$ ,

$$f(\theta x + (1 - \theta)y) \leq \theta f(x) + (1 - \theta)f(y).$$

Elle est dite **strictement convexe** si l'inégalité ci-dessus est stricte lorsque  $x \neq y$  et  $\theta \in ]0, 1[$ .

**Théorème VI.3** Soit  $f : U \rightarrow \mathbb{R}$  une fonction différentiable sur un ouvert  $U$  d'un  $\mathbb{R}$ -espace de Banach  $E$  et soit  $C$  un sous-ensemble convexe de  $U$ . Alors  $f|_C$  est convexe si et seulement si, pour tous  $x, y \in C$ ,

$$f(y) \geq f(x) + df_x(y - x).$$

Elle est strictement convexe si l'inégalité ci-dessus est stricte pour  $x \neq y$ . En supposant en outre que  $f$  est deux fois différentiable,  $f|_C$  est convexe si et seulement si pour tous  $x, y \in C$ ,

$$d^2 f_x(y - x, y - x) \geq 0.$$

Elle est strictement convexe si l'inégalité ci-dessus est stricte pour  $x \neq y$ .

**Dém.** Supposons  $f$  convexe. Soient  $x, y \in C$  et  $\theta \in ]0, 1[$ . On a

$$\frac{f(x + \theta(y - x)) - f(x)}{\theta} \leq f(y) - f(x),$$

d'où

$$df_x(y - x) \leq f(y) - f(x)$$

en faisant tendre  $\theta$  vers 0. Si  $f$  est strictement convexe, on a une inégalité stricte pour  $x \neq y$  et  $\theta \in ]0, 1[$ , mais elle devient large dans le passage à la limite. Pour démontrer qu'effectivement

$$df_x(y - x) < f(y) - f(x),$$

on observe que pour tout  $\omega > 0$ ,

$$x + \theta(y - x) = \frac{\omega - \theta}{\omega} x + \frac{\omega}{\omega} (x + \omega(y - x)),$$

d'où, pour  $0 < \theta < \omega < 1$ ,

$$\frac{f(x + \theta(y - x)) - f(x)}{\theta} < \frac{f(x + \omega(y - x)) - f(x)}{\omega} < f(y) - f(x).$$

On obtient l'inégalité stricte souhaitée en gardant  $\omega$  fixé et en faisant tendre  $\theta$  vers 0. Réciproquement, si on a

$$f(y) \geq f(x) + df_x(y - x)$$

quels que soient  $x$  et  $y \in C$ , on obtient l'inégalité de convexité en prenant la combinaison convexe des inégalités

$$f(x) \geq f(x + \theta(y - x)) - \theta df_{x+\theta(y-x)}(y - x),$$

$$f(y) \geq f(x + \theta(y - x)) + (1 - \theta) df_{x+\theta(y-x)}(y - x).$$

Pour la caractérisation de la convexité en terme de différentielles secondes, on peut considérer, à  $x$  fixé, la fonction

$$g : y \mapsto f(y) - df_x(y).$$

La différence avec  $f$  étant une fonction affine,  $g$  est convexe si et seulement si  $f$  l'est, et  $d^2 f_x = d^2 g_x$ . Or, si  $f$  est convexe, la première partie montre que  $x$  est un minimum (global) de  $g|_C$ . En

appliquant à  $\theta \in [0, 1] \mapsto g(x + \theta(y - x))$  la formule de Taylor-Young exactement comme dans la démonstration de la proposition VI.1, on en déduit que nécessairement

$$d^2 f_x(y - x, y - x) \geq 0.$$

Inversement, supposons que l'on ait cette inégalité quels que soient  $x$  et  $y$  dans  $C$ . Alors d'après la formule des accroissements finis appliquée entre 0 et 1 à la fonction

$$\begin{aligned} [0, 1] &\rightarrow \mathbb{R} \\ \theta &\mapsto f(x + \theta(y - x)) + (1 - \theta) df_{x+\theta(y-x)}(y - x), \end{aligned}$$

il existe  $\theta \in ]0, 1[$  tel que

$$\begin{aligned} f(y) - f(x) - df_x(y - x) &= (1 - \theta) d^2_{x+\theta(y-x)} f(y - x, y - x) \\ &= \frac{1}{1 - \theta} d^2_{x+\theta(y-x)} f(y - (x + \theta(y - x)), y - (x + \theta(y - x))) \geq 0. \end{aligned}$$

Donc  $f$  est convexe d'après la première partie. □

**Théorème VI.4** Soit  $f : U \rightarrow \mathbb{R}$  une fonction définie sur un ouvert  $U$  d'un  $\mathbb{R}$ -espace de Banach  $E$  et soit  $C$  un sous-ensemble convexe de  $U$ .

- i). si  $f|_C$  est convexe et admet un minimum local dans  $C$ , c'est un minimum global ;
- ii). si  $f|_C$  est strictement convexe alors elle admet au plus un minimum, et c'est un minimum strict ;
- iii). si  $f$  est différentiable, une condition nécessaire pour qu'un point  $a \in C$  soit un minimum de  $f|_C$  est

$$df_a(y - a) \geq 0$$

pour tout  $y \in C$ . Si de plus  $f|_C$  est convexe, cette condition est également suffisante.

**Dém.**

- i). supposons  $f|_C$  convexe et admettant un minimum local en  $a \in C$ , et soit  $x \in C$ . Pour tout  $\theta \in [0, 1]$ ,

$$f(a + \theta(x - a)) - f(a) \leq \theta (f(x) - f(a)),$$

et le membre de gauche est positif ou nul pour  $\theta > 0$  assez petit. Par conséquent,  $f(x) - f(a) \geq 0$ .

- ii). si  $f|_C$  est strictement convexe, on obtient comme ci-dessus l'inégalité stricte  $f(x) - f(a) > 0$  pour  $x \neq a$ . Un minimum strict est toujours unique.
- iii). supposons  $f$  différentiable et admettant un minimum en  $a \in C$ . Soit  $x \in C$  : il existe une fonction  $\varepsilon : \theta \mapsto \varepsilon(\theta)$  tendant vers 0 en 0, telle que

$$f(a + \theta(x - a)) - f(a) = \theta df_a(x - a) + \theta \varepsilon(\theta).$$

Si on avait  $df_a(x - a) < 0$ , on aurait  $f(a + \theta(x - a)) < f(a)$  pour  $\theta > 0$  assez petit. Inversement, si  $f|_C$  est convexe et si l'on a l'inégalité

$$df_a(y - a) \geq 0$$

pour tout  $y \in C$ , alors d'après la première partie du théorème VI.3,  $a$  est un minimum de  $f$ .

□

Un exemple important de problème de minimisation sur un ensemble convexe est fourni par ce que l'on appelle des *contraintes-inégalités*

$$g_1(a) \leq 0, \dots, g_p(a) \leq 0,$$

où les fonctions  $g_j$  sont convexes. Un exemple important de fonction convexe à minimiser est celui des fonction(elle)s **quadratiques**, c'est-à-dire de la forme

$$x \mapsto f(x) = \phi(x, x) + \ell(x),$$

où  $\phi$  est une forme bilinéaire continue positive et  $\ell$  est une forme linéaire continue.

## 4 Introduction au calcul des variations

Considérons l'espace  $E = \mathcal{C}^1([0, 1]; \mathbb{R}^n)$  muni de la norme définie par

$$\|u\| := \max(\|u\|_\infty, \|u'\|_\infty), \quad \|u\|_\infty := \max_{t \in [0, 1]} \|u(t)\|_{\mathbb{R}^n}.$$

Soient  $a, b \in \mathbb{R}^n$ . Considérons l'ensemble

$$C := \{u \in E; u(0) = a, u(1) = b\}$$

(évidemment convexe) et une fonction(elle) de la forme

$$A : u \in E \mapsto \int_0^1 L(u(t), u'(t)) dt,$$

où  $L \in \mathcal{C}^1(\mathbb{R}^n \times \mathbb{R}^n; \mathbb{R})$ . On montre sans peine que  $A$  est différentiable sur  $E$ . D'après le théorème VI.4, si  $f$  admet un minimum  $u$  sur  $C$ , alors

$$df_u(h) \geq 0$$

pour tout  $h \in C - u$ , c'est-à-dire pour tout  $h \in E$  tel que  $h(0) = h(1) = 0$ . Par suite, l'inégalité est en fait une égalité : une condition nécessaire pour que  $u$  soit un minimum de  $f$  sur  $C$  est par conséquent

$$\int_0^1 \sum_{i=1}^n \frac{\partial L}{\partial q_i}(u(t), u'(t)) h_i(t) + \int_0^1 \sum_{i=1}^n \frac{\partial L}{\partial p_i}(u(t), u'(t)) h'_i(t) dt = 0,$$

quel que soit  $h \in E$  tel que  $h(0) = h(1) = 0$ . (On a noté  $q_i$  et  $p_i$  les composantes des arguments de  $L$ .) En intégrant par parties le deuxième morceau, on peut réécrire l'égalité ci-dessus sous la forme

$$\int_0^1 \sum_{i=1}^n \left( \frac{\partial L}{\partial q_i}(u(t), u'(t)) - \frac{d}{dt} \left( \frac{\partial L}{\partial p_i}(u(t), u'(t)) \right) \right) h_i(t) dt = 0.$$

Pour qu'elle soit satisfaite quelle que soit la fonction  $h$ , il faut et il suffit, d'après ce que l'on appelle parfois le *lemme fondamental du calcul intégral*, que

$$\forall t \in [0, 1], \quad \frac{\partial L}{\partial q_i}(u(t), u'(t)) - \frac{d}{dt} \left( \frac{\partial L}{\partial p_i}(u(t), u'(t)) \right) = 0,$$

c'est-à-dire que  $u$  soit solution de l'équation différentielle

$$\frac{d}{dt} \left( \frac{\partial L}{\partial p_i}(u, u') \right) = \frac{\partial L}{\partial q_i}(u, u').$$

C'est l'**équation d'Euler-Lagrange** associée à la fonction  $L$ . Si  $L$  est convexe, alors  $A$  aussi (par linéarité de l'intégrale), et par conséquent si  $u \in C$  est solution de l'équation d'Euler-Lagrange, c'est un minimum de  $A$ . Lorsque  $L$  n'est pas convexe, l'équation d'Euler-Lagrange est loin d'être suffisante pour minimiser  $A$ .

# Chapitre VII

## Équations différentielles

La théorie des équations différentielles est très vaste. Nous allons aborder ici seulement quelques points : rappels sur la résolution explicite d'équations simples (essentiellement linéaires) ; théorie de l'existence et de l'unicité des solutions locales (théorème de Cauchy-Lipschitz) ; solutions maximales ; analyse de quelques propriétés qualitatives (portraits de phase).

**Définition VII.1** On appelle **équation différentielle d'ordre  $n$**  une équation de la forme

$$(0.1) \quad u^{(n)} = f(t, u, u', \dots, u^{(n-1)}),$$

où  $f$  est une application définie sur  $I \times U \times U_1 \times \dots \times U_{n-1}$ ,  $I$  étant un intervalle de  $\mathbb{R}$  et  $U, U_1, \dots, U_{n-1}$  des ouverts d'un espace de Banach  $E$ . On appelle **solution** de (0.1) une application  $u$  de classe  $\mathcal{C}^n$  sur un intervalle  $J \subset I$ , telle que pour tout  $t \in J$ ,  $u(t) \in U$ ,  $u'(t) \in U_1, \dots, u^{(n-1)}(t) \in U_{n-1}$  et

$$u^{(n)}(t) = f(t, u(t), u'(t), \dots, u^{(n-1)}(t)).$$

**Cas particuliers :**

- Lorsque  $E = \mathbb{R}$ , l'équation différentielle est dite **scalaire**.
- Lorsque  $f$  est linéaire par rapport à  $(u, u', \dots, u^{(n-1)})$ , l'équation différentielle est dite **linéaire homogène** (à coefficients constants si de plus  $f$  ne dépend pas de  $t$ ). Si  $f$  est de la forme

$$f(t, u, u', \dots, u^{(n-1)}) = g(t, u, u', \dots, u^{(n-1)}) + b(t)$$

avec  $g$  linéaire par rapport à  $(u, u', \dots, u^{(n-1)})$ , l'équation différentielle est dite **linéaire avec terme source**.

- Lorsque  $f$  est indépendante de  $t$ , l'équation différentielle est dite **autonome**.

### 1 Résolution explicite

Il n'y a pas de méthode générale pour calculer des solutions de (0.1). On ne sait résoudre explicitement que certains types d'équations.

#### 1.1 Équations linéaires scalaires d'ordre 1

Étant données deux fonctions  $a$  et  $b \in \mathcal{C}(I; \mathbb{R})$ , l'équation

$$u' = a(t)u + b(t)$$

admet pour solutions les fonctions de la forme

$$u(t) = u_0 e^{\int_{t_0}^t a(\tau) d\tau} + \int_{t_0}^t e^{\int_s^t a(\tau) d\tau} b(s) ds,$$

avec  $t_0 \in I$ ,  $u_0 \in \mathbb{R}$ .

## 1.2 Équations linéaires scalaires d'ordre 2 à coefficients constants

Pour  $p, q \in \mathbb{R}$ , les solutions de l'équation

$$u'' + pu' + qu = 0$$

sont de trois formes possibles, suivant les propriétés de l'**équation caractéristique** :

$$z^2 + pz + q = 0.$$

Les solutions de l'équation différentielle sont de la forme :

- $u(t) = u_1 e^{at} \cos bt + u_2 e^{at} \sin bt$  si l'équation caractéristique a ses racines complexes conjuguées ( $a \pm ib$ );
- $u(t) = u_1 e^{a_1 t} + u_2 e^{a_2 t} \sin bt$  si l'équation caractéristique a deux racines réelles distinctes ( $a_1$  et  $a_2$ );
- $u(t) = (u_1 + t u_2) e^{at}$  si l'équation caractéristique a une racine (réelle) double ( $a$ ).

## 1.3 Équations linéaires à coefficients constants

### Cas homogène

Si  $A$  est une application linéaire continue sur  $E$ , les solutions de

$$u' = Au$$

s'expriment à l'aide de l'**exponentielle** :

$$u(t) = e^{tA} u_0 = \sum_{n=0}^{+\infty} \frac{t^n}{n!} A^n u_0.$$

**En dimension finie**, on peut préciser la forme des solutions en fonction des **valeurs propres** de  $A$ . Supposons en effet que  $A \in \mathcal{L}(E)$  admette comme ensemble des valeurs propres distinctes deux à deux  $\{\lambda_1, \dots, \lambda_k\}$ . On sait (cf cours d'algèbre linéaire) que  $E$  se décompose en somme directe des **sous espaces caractéristiques**

$$E_j = \text{Ker} (A - \lambda_j I)^{s_j},$$

où  $s_j$  est défini comme le plus grand entier  $s$  tel que  $\text{Ker} (A - \lambda_j I)^{s-1}$  soit strictement inclus dans  $\text{Ker} (A - \lambda_j I)^s$ . Par suite, tout élément  $u_0$  de  $E$  s'écrit de façon unique  $u_0 = \sum_{j=1}^k u_{0,j}$  avec  $u_{0,j} \in E_j$ . D'où

$$e^{tA} u_0 = \sum_{j=1}^k e^{tA} u_{0,j} = \sum_{j=1}^k e^{t\lambda_j} e^{t(A - \lambda_j I)} u_{0,j} = \sum_{j=1}^k e^{t\lambda_j} \sum_{n \leq s_j - 1} \frac{t^n}{n!} (A - \lambda_j I)^n u_{0,j}$$

puisque par définition de  $E_j$ ,  $(A - \lambda_j I)_{|E_j}^n = 0$  pour  $n \geq s_j$ . Autrement dit, les solutions de l'équation différentielle  $u' = Au$  sont de la forme

$$u(t) = \sum_{j=1}^k e^{t\lambda_j} P_j(t) u_{0,j},$$

où  $P_j$  est une application polynomiale à valeurs dans  $\mathcal{L}(E)$ . Noter que  $P_j$  est en fait constante si  $s_j = 1$ , et de façon générale,  $P_j$  est de degré  $s_j - 1$ , l'entier  $s_j$  étant lui-même inférieur ou égal à la dimension de  $E_j$ , c'est-à-dire à la multiplicité (algébrique) de  $\lambda_j$ .

L'équation différentielle scalaire du second ordre peut être vue comme un cas particulier, en la réécrivant dans  $\mathbb{R}^2$  :

$$\begin{pmatrix} u \\ v \end{pmatrix}' = \begin{pmatrix} 0 & 1 \\ -q & -p \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix}.$$

Les valeurs propres de la matrice  $2 \times 2$  ci-dessus sont précisément les racines de l'équation caractéristique  $z^2 + pz + q = 0$ .

### Cas avec terme source

Pour résoudre l'équation

$$u' = Au + b(t)$$

on peut appliquer ce que l'on appelle la méthode de « variation de la constante ». Celle-ci consiste à remplacer, dans la solution générale de l'équation homogène la constante  $u_0$  par une fonction : il s'avère alors que  $u_0$  se calcule par **quadrature**, c'est-à-dire au prix du calcul d'une primitive. En effet,  $u(t) = e^{tA} u_0(t)$  fournit une solution de l'équation avec terme source  $b$  si et seulement si

$$u_0'(t) = e^{-tA} b(t).$$

On en déduit la formule générale (formule de **Duhamel**) :

$$u(t) = e^{tA} u_{0,0} + \int_{t_0}^t e^{(t-s)A} b(s) ds.$$

On observe facilement que  $t \mapsto \int_{t_0}^t e^{(t-s)A} b(s) ds$  est une solution particulière de l'équation avec terme source. Autrement dit, la solution générale de l'équation avec terme source est la somme de la solution générale de l'équation homogène et d'une solution particulière de l'équation avec terme source. Attention : ce **principe de superposition** est réservé aux équations **linéaires** ! Cependant, la formule de Duhamel est très utile, même pour les équations non-linéaires, car on peut toujours l'appliquer à un « terme source » qui dépend de la solution  $u$  elle-même : ceci fournit alors une équation implicite en  $u$ ...

## 2 Lemme de Gronwall

Avec la formule de Duhamel, le lemme de Gronwall est l'un des outils fondamentaux dans la théorie des équations différentielles. Il en existe plusieurs versions mais l'idée est toujours d'obtenir une estimation pour une fonction  $u$  à valeurs dans  $\mathbb{R}$  qui satisfait une inégalité implicite mais linéaire.

## 2.1 Inéquations différentielles

Supposons qu'une fonction  $u \in \mathcal{C}^1(I; \mathbb{R})$  vérifie

$$u'(t) \leq a(t)u(t) + b(t),$$

avec  $a$  et  $b \in \mathcal{C}(I; \mathbb{R})$ . Alors en s'inspirant de la résolution de l'équation différentielle  $u' = a(t)u + b(t)$ , on multiplie l'inégalité par le nombre strictement positif  $e^{-\int_{t_0}^t a(\tau) d\tau}$  et on en déduit

$$\frac{d}{dt} \left( e^{-\int_{t_0}^t a(\tau) d\tau} u(t) \right) \leq e^{-\int_{t_0}^t a(\tau) d\tau} b(t),$$

d'où en intégrant :

$$e^{-\int_{t_0}^t a(\tau) d\tau} u(t) - u(t_0) \leq \int_{t_0}^t e^{-\int_{t_0}^s a(\tau) d\tau} b(s) ds,$$

et finalement :

$$u(t) \leq u(t_0) e^{\int_{t_0}^t a(\tau) d\tau} + \int_{t_0}^t e^{\int_s^t a(\tau) d\tau} b(s) ds.$$

## 2.2 Inéquations intégrales

On peut remarquer que l'inéquation différentielle  $u'(t) \leq a(t)u(t) + b(t)$  implique l'inéquation intégrale :

$$u(t) \leq u_0 + \int_{t_0}^t a(s)u(s) + b(s) ds,$$

mais pas l'inverse. Or en pratique, on a souvent affaire à des inéquations déjà sous forme intégrale, qu'il est **hors de question** de dériver ! Le lemme de Gronwall permet justement de contourner ce problème.

**Lemme VII.1 (Gronwall)** Soient  $a$  et  $c \in \mathcal{C}([0, T]; \mathbb{R})$ , où  $T > 0$ . On suppose de plus la fonction  $a$  à valeurs positives. Si  $u \in \mathcal{C}^1([0, T]; \mathbb{R})$  vérifie

$$u(t) \leq c(t) + \int_0^t a(s)u(s) ds$$

quel que soit  $t \in [0, T]$ , alors

$$u(t) \leq c(t) + \int_0^t c(s) a(s) e^{\int_0^s a(\tau) d\tau} ds.$$

**Remarque VII.1** Il ne s'agit pas de retenir cette formule (souvent appelée inégalité de Gronwall) par cœur mais plutôt de savoir la retrouver très vite ; suivant le contexte, avec  $c$  constante par exemple, elle peut se simplifier considérablement ; d'autre part, on peut écrire des formules analogues avec un point  $t_0$  quelconque à la place de 0. Mais attention aux bornes des intégrales que l'on manipule : mettre des valeurs absolues au bon endroit !

**Dém.** L'astuce pour démontrer le lemme de Gronwall consiste à démontrer l'inégalité voulue non pas pour  $u$  directement mais pour le second membre de l'inégalité dont on dispose. Plus précisément, posons

$$v(t) = \int_0^t a(\tau) u(\tau) \, d\tau.$$

Alors on a

$$v'(t) = a(t) u(t) \leq a(t) (c(t) + v(t))$$

d'après les hypothèses : c'est ici qu'intervient l'hypothèse  $a \geq 0$ . Ainsi  $v$  satisfait une inéquation différentielle comme au paragraphe précédent, avec  $b(t) = a(t) c(t)$ . En observant que  $v(0) = 0$ , on en déduit facilement que

$$v(t) \leq \int_0^t a(\tau) c(\tau) e^{\int_\tau^t a(s) \, ds} \, d\tau, \quad \text{pour tout } t \in [0, T].$$

On conclut en injectant cette majoration de  $v(t)$  dans l'inégalité de départ.  $\square$

### 3 Théorème de Cauchy-Lipschitz

L'essentiel de la théorie des équations différentielles est fondé sur l'existence et l'unicité locale des solutions pour ce que l'on appelle le problème de Cauchy, c'est-à-dire l'équation différentielle assortie d'une « condition initiale ». Dans le cas général d'une équation d'ordre  $n$ , une condition initiale est la donnée de la fonction inconnue  $u$  et de toutes ses dérivées jusqu'à l'ordre  $n - 1$  en un point  $t_0$ . En fait, on peut toujours se ramener au cas  $n = 1$ , quitte à agrandir l'espace  $E$  (comme on l'a vu pour l'équation linéaire d'ordre 2) et en considérant comme nouvelle fonction inconnue  $(u, u', \dots, u^{(n-1)})$ .

Désormais, on se place donc dans le cas  $n = 1$ , et on se donne une fonction  $f : I \times U \rightarrow E$ , où  $I$  est un intervalle ouvert de  $\mathbb{R}$  et  $U$  un ouvert d'un espace de Banach  $E$ . Il faut bien sûr faire des hypothèses de régularité sur  $f$ . On supposera au minimum  $f$  continue. Ceci permet de démontrer (en faisant appel au théorème d'Ascoli, ce qui sort du cadre de ce cours) l'existence si  $E$  est de dimension finie mais en tous cas pas l'unicité des solutions. Pour assurer l'unicité, on demandera à  $f$  d'être localement Lipschitzienne par rapport à sa variable  $x \in U$ .

**Théorème VII.1 (Cauchy-Lipschitz)** Soit  $f \in \mathcal{C}(I \times U; E)$ , où  $I$  est un intervalle ouvert de  $\mathbb{R}$  et  $U$  un ouvert d'un espace de Banach  $E$ . On suppose de plus qu'il existe un voisinage de  $(t_0, u_0)$  dans  $I \times U$  et  $L > 0$  tel que pour tous  $(t, x)$  et  $(t, y)$  dans ce voisinage

$$\|f(t, x) - f(t, y)\| \leq L \|x - y\|.$$

**Existence :** Il existe  $\tau > 0$  et  $u \in \mathcal{C}^1([t_0 - \tau, t_0 + \tau]; U)$  solution du problème de Cauchy :

$$(3.2) \quad \begin{cases} u' = f(t, u) \\ u(t_0) = u_0. \end{cases}$$

**Unicité :** Si  $v$  est une autre solution, elle coïncide avec  $u$  sur un intervalle d'intérieur non vide inclus dans  $[t_0 - \tau, t_0 + \tau]$ .

**Régularité :** Si de plus  $f$  est de classe  $\mathcal{C}^r$ ,  $r \geq 1$ , alors  $u$  est de classe  $\mathcal{C}^{r+1}$ .

**Remarque VII.2** Dès que  $f$  est de classe  $\mathcal{C}^1$  elle est effectivement localement Lipschitzienne au sens de l'énoncé. En effet, d'après le théorème des accroissements finis,

$$\|f(t, x) - f(t, y)\| \leq L \|x - y\| \quad \text{où } L := \max_{\theta \in [0,1]} \|\|d_2 f_{\theta x + (1-\theta)y}\|\|$$

est fini pour  $x$  et  $y$  voisins (une fonction continue étant toujours localement bornée).

**Dém.** Pour simplifier, on va démontrer les résultats en remplaçant l'intervalle centré  $[t_0 - \tau, t_0 + \tau]$  par  $[t_0, t_+]$  avec  $t_+ = t_0 + \tau > t_0$  (le cas de l'intervalle  $[t_0 - \tau, t_0]$  s'en déduisant par le changement de variable  $t \mapsto (t_0 - t)$ ) : ceci évite quelques valeurs absolues dans les calculs.

On commence par choisir un voisinage de  $(t_0, u_0)$  de la forme

$$C(t_+, R) = [t_0, t_+] \times \bar{B}(u_0, R) \subset I \times U$$

avec  $t_+ > t_0$  et  $R > 0$ , où  $\bar{B}(u_0, R)$  désigne la boule fermée de centre  $u_0$  et de rayon  $R$ , dans lequel  $f$  est bornée, disons par une constante  $M$  :

$$\|f(t, x)\| \leq M \quad \text{quel que soit } (t, x) \in [t_0, t_+] \times \bar{B}(u_0, R),$$

et aussi Lipschitzienne :

$$\|f(t, x) - f(t, y)\| \leq L \|x - y\| \quad \text{pour tous } t \in [t_0, t_+], x \text{ et } y \in \bar{B}(u_0, R).$$

On va obtenir une solution du problème de Cauchy (3.2) telle que  $(t, u(t)) \in C(t_+; R)$ , quitte à diminuer  $t_+$ . Pour cela on observe qu'il suffit de résoudre le problème intégral :

$$u(t) = u_0 + \int_{t_0}^t f(\tau, u(\tau)) \, d\tau,$$

ce qui peut se faire en invoquant un théorème de point fixe ou bien en construisant « à la main » une solution comme limite de la suite  $(u^n)_{n \in \mathbb{N}}$  définie par

$$u^0 = u_0, \quad u^{n+1}(t) = u_0 + \int_{t_0}^t f(\tau, u^n(\tau)) \, d\tau.$$

Il faut d'abord vérifier que ce schéma itératif fournit bien une suite  $(u^n)_{n \in \mathbb{N}}$  de fonctions continues sur  $[t_0, t_+]$  et à valeurs dans  $\bar{B}(u_0, R)$ , du moins pour  $t_+$  assez proche de  $t_0$ . On raisonne bien sûr par récurrence : la fonction  $u^0$  constante est trivialement continue et à valeurs dans  $\bar{B}(u_0, R)$ . Supposons qu'on ait construit  $u^n$  continue sur  $[t_0, t_+]$  et à valeurs dans  $\bar{B}(u_0, R)$ . Alors  $u^{n+1}$  est continue par construction et de plus :

$$\|u^{n+1}(t) - u_0\| \leq \int_{t_0}^t \|f(\tau, u^n(\tau))\| \, d\tau \leq (t_+ - t_0) M.$$

Donc il suffit que

$$(t_+ - t_0) M \leq R$$

pour que  $u^{n+1}$  soit aussi à valeurs dans  $\bar{B}(u_0, R)$ . Désormais, on suppose que  $t_+$  vérifie cette contrainte.

L'étape essentielle est de montrer que la suite  $(u^n)_{n \in \mathbb{N}}$  est de Cauchy et donc convergente dans l'espace de Banach  $\mathcal{C}([t_0, t_+]; E)$  (muni de la norme uniforme). Or on montre facilement par récurrence que

$$\|u^{n+1}(t) - u^n(t)\| \leq \frac{L^n (t - t_0)^n}{n!} \sup_{[t_0, t_+]} \|u^1 - u^0\| \quad \text{pour } t \in [t_0, t_+].$$

Par suite,

$$\sup_{[t_0, t_+]} \|u^{n+p}(t) - u^n(t)\| \leq \sum_{k=n}^{n+p-1} \frac{L^k (t_+ - t_0)^k}{k!} \sup_{[t_0, t_+]} \|u^1 - u^0\|,$$

ce qui tend vers 0 lorsque  $n$  tend vers  $+\infty$  puisque la série  $\sum \frac{L^n (t_+ - t_0)^n}{n!}$  est convergente. Ceci signifie précisément que  $(u^n)$  est de Cauchy dans  $\mathcal{C}([t_0, t_+]; E)$ .

Soit donc  $u$  la limite de cette suite. On vérifie sans peine que  $u$  est solution de notre problème : par passage à la limite dans l'égalité

$$u^{n+1}(t) = u_0 + \int_{t_0}^t f(\tau, u^n(\tau)) \, d\tau$$

on voit que

$$u(t) = u_0 + \int_{t_0}^t f(\tau, u(\tau)) \, d\tau.$$

Comme  $u$  est continue ainsi que  $f$ , le second membre de cette égalité est de classe  $\mathcal{C}^1$  et donc  $u$  aussi. (Plus généralement, si  $f$  est de classe  $\mathcal{C}^r$ ,  $r \geq 0$ ,  $u$  est de classe  $\mathcal{C}^{r+1}$ .) Et  $u$  vérifie :

$$u'(t) = f(t, u(t)) \quad \text{pour tout } t \in [t_0, t_+].$$

Ceci achève la preuve de l'existence d'une solution.

La preuve de l'unicité locale est très facile grâce au lemme de Gronwall. En effet, si  $v$  est une autre solution du même problème de Cauchy, elle est à valeurs dans  $\bar{B}(u_0, R)$  sur un intervalle  $[t_0, t_1]$ ,  $0 < t_1 \leq t_+$ . Et on a

$$u(t) - v(t) = \int_{t_0}^t (f(\tau, u(\tau)) - f(\tau, v(\tau))) \, d\tau.$$

et donc

$$\|u(t) - v(t)\| \leq L \int_{t_0}^t \|u(\tau) - v(\tau)\| \, d\tau.$$

Par conséquent, une version simplissime du lemme de Gronwall implique :

$$\|u(t) - v(t)\| \leq e^{Lt} \|u(t_0) - v(t_0)\| = 0$$

pour tout  $t \in [t_0, t_1]$ . □

Désormais, on suppose que  $f : I \times U \rightarrow E$  est continue et est localement Lipschitzienne en tout point  $(t_0, u_0) \in I \times U$ , c'est-à-dire qu'il existe un voisinage de  $(t_0, u_0)$  dans  $I \times U$  et  $L > 0$  tel que pour tous  $(t, x)$  et  $(t, y)$  dans ce voisinage

$$\|f(t, x) - f(t, y)\| \leq L \|x - y\|.$$

Ceci nous permettra d'appliquer le théorème VII.1 en tout point  $(t_0, u_0) \in I \times U$ .

## 4 Solutions maximales

L'unicité locale dans le théorème VII.1 permet de démontrer le résultat suivant.

**Lemme VII.2** *Si  $u_1 \in \mathcal{C}^1(J_1; U)$  et  $u_2 \in \mathcal{C}^1(J_2; U)$  sont deux solutions de l'équation différentielle  $u' = f(t, u)$  sur des intervalles  $J_1$  et  $J_2$  respectivement, et s'il existe  $t_0 \in J_1 \cap J_2$  tel que  $u_1(t_0) = u_2(t_0)$ , alors*

$$u_1(t) = u_2(t), \quad \text{pour tout } t \in J_1 \cap J_2.$$

**Dém.** C'est une application classique de la notion de **connexité**. On remarque que  $J_1 \cap J_2$  est un intervalle non vide par hypothèse et on considère l'ensemble :

$$A = \{t \in J_1 \cap J_2; u_1(t) = u_2(t)\}.$$

D'après l'unicité locale des solutions,  $A$  est un ouvert. De plus, il est clairement fermé par continuité de  $u_1$  et  $u_2$ . Donc  $A$  est égal à  $J_1 \cap J_2$ .  $\square$

Ce lemme montre qu'il existe un plus grand intervalle  $J$  sur lequel le problème de Cauchy (3.2) admet une solution, et que cette solution est unique. Cette solution est appelée **solution maximale** : par définition, on ne peut pas la prolonger à  $I \setminus J$ . Lorsque  $J = I$  on dit que cette solution est **globale**.

La question naturelle est ensuite de savoir à quelle(s) condition(s) une solution maximale est globale.

**Théorème VII.2** *Soit  $u \in \mathcal{C}^1(J; U)$  une solution maximale de  $u' = f(t, u)$ . Notons  $b$  la borne supérieure de  $I$  et  $\beta$  la borne supérieure de  $J$ . Alors ou bien  $\beta = b$  ou bien «  $u$  sort de tout compact » de  $U$ , c'est-à-dire que pour tout compact  $K \subset U$ , il existe  $\eta < \beta$  tel que*

$$u(t) \in U \setminus K, \quad \text{pour } t \geq \eta \text{ avec } t \in J.$$

Et on a le résultat analogue pour les bornes inférieures.

**Dém.** Supposons  $\beta < b$  et raisonnons par l'absurde. S'il existait un compact  $K$  et une suite  $t_n$  tendant vers  $\beta$  telle que  $u(t_n) \in K$  pour tout  $n$ , quitte à en extraire une sous-suite, on pourrait supposer qu'elle converge vers  $\underline{u} \in K$ . Soient  $\underline{\tau} > 0$  et  $R > 0$  tel que  $f$  soit bornée et Lipschitzienne par rapport à  $u$  dans  $[\beta - 2\underline{\tau}, \beta + 2\underline{\tau}] \times \bar{B}(\underline{u}, 2R)$ . Alors en vertu du lemme VII.3 ci-après, il existe  $\tau \leq \underline{\tau}$  tel que pour tout  $(t_0, u_0) \in [\beta - \tau, \beta + \tau] \times \bar{B}(\underline{u}, R)$  la solution maximale du problème de Cauchy (3.2) est définie sur un intervalle contenant  $[t_0 - \tau, t_0 + \tau]$ . Or, pour  $n$  assez grand,  $(t_n, u(t_n)) \in [\beta - \tau, \beta + \tau] \times \bar{B}(\underline{u}, R)$ , et  $t_n + \tau > \beta$ . Ceci contredit le fait que  $u$  soit une solution maximale.  $\square$

**Lemme VII.3** *Supposons que  $f$  soit continue, bornée et Lipschitzienne par rapport à  $x$  dans  $[\underline{t} - 2\underline{\tau}, \underline{t} + 2\underline{\tau}] \times \bar{B}(\underline{u}, 2R)$  pour  $\underline{\tau} > 0$  et  $R > 0$ . Alors il existe  $\tau \in ]0, \underline{\tau}]$  tel que pour tout  $(t_0, u_0) \in [\underline{t} - \tau, \underline{t} + \tau] \times \bar{B}(\underline{u}, R)$ , la solution maximale du problème de Cauchy (3.2) soit définie sur un intervalle contenant  $[t_0 - \tau, t_0 + \tau]$ .*

**Dém.** Il suffit de reprendre la démonstration du théorème VII.1 : en effet, la méthode suivie permet de construire une solution dans l'intervalle  $[t_0 - \tau, t_0 + \tau]$  qui soit à valeurs dans  $\bar{B}(u_0, R)$  (donc *a fortiori* à valeurs dans  $\bar{B}(\underline{u}, 2R)$  par l'inégalité triangulaire), pour

$$\tau := \min\left(\underline{\tau}, \frac{R}{M}\right),$$

où  $M$  est une borne de  $f$  sur  $[\underline{t} - 2\tau, \underline{t} + 2\tau] \times \bar{B}(\underline{u}, 2R)$ .  $\square$

Si l'espace ambiant  $E$  est de dimension finie, et si  $U = E$ , le théorème VII.2 signifie que pour une solution maximale non globale  $u \in \mathcal{C}^1(J; E)$  avec  $\beta := \sup J < \sup I$ ,

$$\lim_{t \rightarrow \beta} \|u(t)\| = +\infty.$$

**Remarque VII.3** *Ce phénomène est indépendant de la régularité de la fonction  $f$  ! Il se produit par exemple pour l'équation de Riccati, où  $E = \mathbb{R}$  et  $f(t, u) = u^2$  (fonction évidemment de classe  $\mathcal{C}^\infty$ ) : seule la solution nulle est globale. En effet, d'après l'unicité, une solution qui s'annule est identiquement nulle ; et une solution qui ne s'annule pas vérifie nécessairement*

$$\frac{u'(t)}{u(t)^2} = 1 \quad \text{d'où} \quad -\frac{1}{u(t)} + \frac{1}{u(t_0)} = t - t_0 \quad \text{et donc} \quad \lim_{t \rightarrow t_0 + 1/u(t_0)} |u(t)| = +\infty.$$

Il y a cependant quelques cas remarquables dans lesquels toutes les solutions maximales sont globales :

- Lorsque  $U = E$  et  $f$  est **globalement Lipschitzienne**, c'est-à-dire qu'il existe  $L > 0$  tel que pour tous  $(t, x)$  et  $(t, y)$  dans  $I \times E$ ,

$$\|f(t, x) - f(t, y)\| \leq L \|x - y\|,$$

on peut reprendre la démonstration du théorème VII.1 sans avoir besoin d'assurer que la suite  $u^n$  prend ses valeurs dans une boule  $\bar{B}(u_0; R)$  : ceci supprime la restriction sur le temps d'existence.

- En particulier, lorsque  $f$  est **linéaire** par rapport à  $x$ , c'est-à-dire pour une équation différentielle de la forme  $u' = A(t)u$  où  $A(t) \in \mathcal{L}(E)$ , toutes les solutions maximales sont globales.
- Lorsque  $U = E$  de **dimension finie** et  $f$  est **uniformément bornée**, le théorème VII.2 montre que toutes les solutions maximales sont globales. En effet, si  $u'(t) = f(t, u(t))$  pour tout  $t \in J$  alors

$$\|u(t)\| \leq (\beta - t_0) M$$

avec  $\beta = \sup J$ ,  $t_0 \in J$  fixé et  $M$  une borne pour  $\|f\|$ .

## 5 Analyse qualitative

Notons que  $u \in \mathcal{C}^1(J; U)$  est solution maximale d'une équation différentielle  $u' = f(t, u)$  si et seulement si  $s \in J \mapsto (s, u(s)) \in J \times E$  est solution maximale du système autonome

$$\begin{cases} \dot{t} = 1 \\ \dot{u} = f(t, u). \end{cases}$$

D'un point de vue théorique, on ne perd donc rien à se restreindre aux équations autonomes. On ne sait pas résoudre explicitement une équation différentielle non-linéaire générale  $u' = f(u)$ . Cependant, on peut mettre en évidence des propriétés dites qualitatives de cette équation en se plaçant dans l'**espace des phases**  $E$  et en considérant les **trajectoires**, c'est-à-dire les courbes

$$\{u(t); t \in J \text{ et } u \in \mathcal{C}^1(J) \text{ solution de } u' = f(u)\}.$$

Le terme « trajectoire » appartient au vocabulaire de la cinématique. On parle aussi d'**orbite** (dans le langage de l'algèbre, à cause d'une propriété de groupe qui sort du cadre de ce cours). D'après l'unicité des solutions, les orbites ne s'intersectent pas.

On appelle **portrait de phase** la partition de l'espace des phases en orbites. On montre que le portrait de phase ne dépend que de la direction des vecteurs  $f(x) \in E$  et non de leur norme.

En dimension finie, l'étude du portrait de phase commence par la recherche du signe des composantes de  $f(x)$ . Les points où toutes ces composantes sont nulles sont des **points fixes**, c'est-à-dire des orbites réduites à un singleton correspondant à des **solutions stationnaires** (c'est-à-dire constantes) de  $u' = f(u)$ . La nature de ces points fixes, selon la position des valeurs propres de la jacobienne de  $f$  par rapport à l'axe imaginaire, est un autre élément important : un point fixe  $x_0$  est dit **hyperbolique** si  $df_{x_0}$  n'a pas de valeur propre imaginaire pure. On montre (théorème de Hartman-Grobman) que le portrait de phase de  $u' = f(u)$  au voisinage d'un point fixe hyperbolique  $x_0$  est **topologiquement** équivalent au portrait de phase de l'équation linéaire  $u' = df_{x_0}(u)$  au voisinage de 0.

On cherche ensuite d'autres orbites remarquables, et par exemple des **cycles**, c'est-à-dire des courbes fermées simples correspondant à des **solutions périodiques** : c'est en général un problème non trivial de trouver des cycles !

# Index

- Banach
  - espace de, 3
- base canonique, 9
- classe  $C^1$ 
  - fonction de, 4
- classe  $C^n$ 
  - fonction de, 28
- dérivée
  - dans une direction, 7
  - partielle, 10
- dérivation des fonctions composées, 8
- développement limité, 37
- difféomorphisme, 19
- différentiable, 4
  - continûment, 4
  - en un point, 3
- différentielle
  - d'une fonction, 4
  - d'une fonction en un point, 4
  - seconde, 26
- divergence, 12
- Duhamel
  - formule de, 45
- gradient, 12
- isomorphisme, 9
- jacobien, 22
- jacobienn
  - matrice, 12
- multi-linéaire
  - application, 4
- norme
  - d'une application linéaire continue, 3
  - dans  $\mathbb{R}^n$ , 3
  - de matrice, 3
- orbite, 52
- ouverte
  - application, 22
- partielle
  - application, 5, 7, 10, 16
  - dérivée, 10
  - différentielle, 16
- phase
  - portrait, 52
- phases
  - espace des, 51
- rotationnel, 12
- solution
  - globale, 50
  - maximale, 50
- symétrique, 30
- Taylor
  - reste de Lagrange, 35
  - reste intégral, 34
  - série de, 37
  - Young, 36
- trajectoire, 51
- transposition, 30



# Table des matières

<b>I</b>	<b>Différentielle d'une fonction</b>	<b>3</b>
1	Introduction . . . . .	3
2	Premiers exemples . . . . .	4
3	Premières propriétés . . . . .	7
4	En dimension finie . . . . .	9
<b>II</b>	<b>Théorème des accroissements finis</b>	<b>13</b>
1	Fonctions d'une variable réelle . . . . .	13
2	Théorème général . . . . .	14
3	Applications . . . . .	15
<b>III</b>	<b>Difféomorphismes</b>	<b>19</b>
1	Introduction . . . . .	19
2	Théorème d'inversion locale . . . . .	20
3	Théorème des fonctions implicites . . . . .	23
<b>IV</b>	<b>Différentielles d'ordre supérieur</b>	<b>25</b>
1	Différentielle seconde . . . . .	25
2	Différentielle d'ordre $n$ . . . . .	28
<b>V</b>	<b>Formules de Taylor</b>	<b>33</b>
1	Formule de Taylor avec reste intégral . . . . .	33
2	Formule de Taylor-Lagrange . . . . .	35
3	Formule de Taylor-Young . . . . .	36
<b>VI</b>	<b>Extrema</b>	<b>39</b>
1	Extrema libres . . . . .	39
2	Extrema liés . . . . .	41
3	Convexité et minima . . . . .	42
4	Introduction au calcul des variations . . . . .	45
<b>VII</b>	<b>Équations différentielles</b>	<b>47</b>
1	Résolution explicite . . . . .	47
1.1	Équations linéaires scalaires d'ordre 1 . . . . .	47
1.2	Équations linéaires scalaires d'ordre 2 à coefficients constants . . . . .	48
1.3	Équations linéaires à coefficients constants . . . . .	48
2	Lemme de Gronwall . . . . .	49
2.1	Inéquations différentielles . . . . .	50

2.2	Inéquations intégrales . . . . .	50
3	Théorème de Cauchy-Lipschitz . . . . .	51
4	Solutions maximales . . . . .	54
5	Analyse qualitative . . . . .	55