EQUIPE D’ANALYSE
URA 754 – CNRS
Université Pierre et Marie Curie - Paris 6

Tour 46 - 0 - Boîte 186 - 4, place Jussieu - 75252 PARIS CEDEX 05
Tél : (33-1) 44 27 53 49 - Télécopie : (33-1) 44 27 25 55 - lana@ccr.jussieu.fr

CALCUL DIFFERENTIEL

Licence B

Fonctions différentiables,
Théorème des fonctions implicites,
Equations différentielles, Optimisation.

Sylvie DELABRIERE

1998-1999

1

APPLICATIONS DIFFÉRENTIABLES ET FORMULE DES ACCROISSEMENTS FINIS

1.1 Définitions et notations.

Définitions 1.1.1. *Espaces vectoriels normés.*

Dans tout ce chapitre, E et F désigneront des *espaces vectoriels normés* sur un corps $\mathbb{K}$, qui sera en général $\mathbb{R}$ et plus exceptionnellement $\mathbb{C}$. Leurs *normes* seront notées $\|.\|_E$ et $\|.\|_F$ ou plus simplement $\|.\|$ s'il n'y a pas de confusion possible.

On dira que deux normes $\|.\|$ et $|||.|||$ sur E sont *équivalentes* s'il existe des constantes $C_1, C_2 > 0$ telles

$$\forall x \in E\ ,\ C_1 \|x\| \leq |||x||| \leq C_2 \|x\|$$

Proposition 1.1.2. *Sur les espaces vectoriels de dimension finie, toutes les normes sont équivalentes et seuls les espaces vectoriels de dimension finie ont cette propriété.*

Soit $n \in \mathbb{N}$. L'espace vectoriel $\mathbb{R}^n$ est toujours muni de sa base canonique, notée $\{e_1, e_2, ..., e_n\}$ ($\{i, j, k\}$ dans le cas $n = 3$) ou encore par ses coordonnées

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ . \\ . \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ . \\ . \\ 0 \end{pmatrix}, ...e_n = \begin{pmatrix} 0 \\ 0 \\ . \\ . \\ 1 \end{pmatrix}$$

Les normes usuelles sur $\mathbb{K}^n$ sont, pour $x = \begin{pmatrix} x_1 \\ x_2 \\ . \\ . \\ x_n \end{pmatrix}$:

$$\|x\|_1 = |x_1| + |x_2| + ... + |x_n|$$

$$\|x\|_\infty = \sup\{|x_1|, |x_2|, ..., |x_n|\}$$
$$\|x\|_2 = (|x_1|^2 + |x_2|^2 + ... + |x_n|^2)^{1/2}$$

La norme $\|x\|_2$ provient du produit scalaire $<,>$ défini par :

$$< x, y >= x_1y_1 + x_2y_2 + ... + x_ny_n$$

Un espace vectoriel muni d'un produit scalaire est appelé un *espace préhilbertien* et la norme associée, définie par $\|x\| = (< x, x >)^{1/2}$ est appelée une *norme hilbertienne.*

Notons que l'on peut vérifier l'équivalence des trois normes ci-dessus de façon élémentaire en écrivant :

$$\|x\|_\infty \leq \|x\|_2 \leq \|x\|_1 \leq n\,\|x\|_\infty$$

Définitions 1.1.3. *Applications linéaires continues.*

L'espace vectoriel des *applications linéaires continues* de E dans F sera noté $\mathscr{L}(E;F)$ ou plus simplement $\mathscr{L}(E)$ si $E = F$.

Si $F = \mathbb{K}$, on notera $E' = \mathscr{L}(E;\mathbb{K})$ et E' sera appelé le *dual de E*. Les éléments de E' seront appelées des *formes linéaires continues* sur E.

Si de plus, $E = \mathbb{K}, \mathscr{L}(E;\mathbb{K}) = \mathscr{L}(\mathbb{K})$ s'identifie à $\mathbb{K}$.

La valeur d'un élément $T \in \mathscr{L}(E;F)$ sur un vecteur $x \in E$ sera notée simplement Tx.

Exemple.

Soit $E = E_1 \times E_2 \times ... \times E_n$, où $E_1, E_2, ..., E_n$ sont des espaces vectoriels normés. Alors tout élément $T \in \mathscr{L}(E,F)$ est défini par n applications linéaires continues $T_i \in \mathscr{L}(E_i,F)$ telles que si $x = \begin{pmatrix} x_1 \\ x_2 \\ . \\ . \\ x_n \end{pmatrix} \in E$, alors $Tx = \sum_{i=1}^n T_i x_i$.

Réciproquement, tout ensemble d'applications $T_i \in \mathscr{L}(E_i,F)$, $i = 1, 2, ...n$, définit un élément $T \in \mathscr{L}(E,F)$ par la formule $Tx = \sum_{i=1}^n T_i x_i$. On en déduit que $\mathscr{L}(E,F)$ et $\prod_{i=1}^n \mathscr{L}(E_i,F)$ sont deux représentations du même espace.

Si les espaces sont de dimension finie, par extension des notations matricielles, si

$$A, A_1, A_2, ..., A_n$$

sont les matrices de

$$T, T_1, T_2, ..., T_n$$

dans une base adaptée à la décomposition de E en $E = E_1 \times E_2 \times ... \times E_n$, (c'est à dire que la base considérée de E est constituée de la réunion de bases de chaque E_i), on notera

$$A = \begin{pmatrix} A_1 & A_2 & . & . & A_n \end{pmatrix}$$

et on aura la relation matricielle

$$AX = \sum_{i=1}^n A_i X_i$$

où X est une matrice colonne représentant les coordonnées d'un vecteur $x \in E$ et les X_i sont les coordonnées de la projection de x sur chaque E_i.

La *norme* d'une application linéaire continue $T \in \mathscr{L}(E;F)$ est définie par

$$\|T\|_{E,F} = \sup\{\|Tx\|_F \ \mid \ \|x\|_E \leq 1\}$$

On vérifie sans difficultés que cette norme vaut aussi :

$$\sup\{\|Tx\|_F \ \mid \ \|x\|_E = 1\} \text{ et } \sup\{\|Tx\|_F \ \mid \ \|x\|_E < 1\}$$

S'il n'y a pas de confusion possible, on notera simplement $\|T\|$ cette norme.

En particulier, la norme du dual E' de E est donnée par

$$\forall \varphi \in E' \ , \ \|\varphi\| = \sup\{|\varphi x| \ \mid \ \|x\| \leq 1\}$$

Proposition 1.1.4. *Une application linéaire est continue si et seulement sa norme est finie.*

Proposition 1.1.5. *Si E et F sont de dimension finie, toutes les applications linéaires de E dans F sont continues.*

Exemple.

Dans le cas où $E = E_1 \times E_2 \times ... \times E_n$, avec $E_1, E_2, ..., E_n$ espaces vectoriels normés, on peut munir E de la norme :

$$\forall x = \begin{pmatrix} x_1 \\ x_2 \\ . \\ . \\ x_n \end{pmatrix} \in E \ , \ \|x\| = \sum_{i=1}^{n} \|x_i\|$$

La norme associée de$\mathscr{L}(E,F)$ est alors définie par

$$\|T\| = \sup_{1 \leq i \leq n} \|T_i\|$$

pour T tel que $Tx = \sum_{i=1}^{n} T_i x_i$.

On dira qu'une application linéaire continue de E dans F est un *isomorphisme* si cette application est bijective et son inverse est continu.

Proposition 1.1.6. *Si E est de dimension finie, une application linéaire T de E dans E est bijective si et seulement si elle est injective si et seulement si elle est surjective.*

C'est une conséquence du théorème de la dimension :

$$dim(E) = dim(TE) + dim(kerT)$$

Une *application affine continue* A de E dans F est définie par un vecteur $b \in F$ et une application linéaire continue $T \in \mathscr{L}(E;F)$ telle que

$$\forall x \in E \ , \ Ax = b + Tx$$

Définitions 1.1.7. *Représentation matricielle en dimension finie.*

Soient $E = \mathbb{K}^n$ et $F = \mathbb{K}^p$. Les coordonnées des vecteurs de $\mathbb{K}^n$ dans la base canonique sont représentés par des matices $n \times 1$, appelées *matrices colonne.*

Une application linéaire $T \in \mathscr{L}(\mathbb{K}^n; \mathbb{K}^p)$ est représentée dans les bases canoniques de $\mathbb{K}^n$ et $\mathbb{K}^p$ par *une matice* $p \times n$, A, à p lignes et n colonnes. Les vecteurs colonne de A sont les coordonnées des images par T, des vecteurs de la base initiale, exprimées dans la base de l'espace d'arrivée.

En particulier, les formes linéaires sur $\mathbb{K}^n$, c'est à dire les applications linéaires de $\mathbb{K}^n$ dans $\mathbb{K}$, sont représentées par des matrices $1 \times n$, c'est à dire par des *matrices ligne.*

En identifiant les matrices colonnes et les matrices lignes à n éléments, ceci nous permet d'identifier algébriquement le dual $(\mathbb{K}^n)' = \mathscr{L}(\mathbb{K}^n; \mathbb{K})$ de $\mathbb{K}^n$ à $\mathbb{K}^n$.

Tout espace vectoriel E de dimension n muni d'une base s'identifie à $\mathbb{K}^n$, par l'intermédiaire des coordonées des vecteurs de E sur la base. Les représentations matricielles ci-dessus restent donc valables dans ce cadre.

Par contre, en tant qu'espace normé, le dual de $\mathbb{K}^n$ ne s'identifie à lui-même que s'il est muni de la norme hilbertienne $\|.\|_2$.
On vérifiera que la norme duale sur $\mathbb{K}^n$ de la norme $\|.\|_1$ est la norme $\|.\|_\infty$ et réciproquement.

Définitions 1.1.8. *Boules ouvertes, boules fermées, voisinages.*

Soit E un espace vectoriel normé.
Soit $a \in E$ et $r > 0$. Une *boule ouverte* $B(a, r)$ de centre a et de rayon r est le sous-ensemble de E défini par

$$B(a, r) = \{x \in E \mid \|x - a\| < r\}$$

On appellera plus simplement ces ensembles des boules.

Une *boule fermée* $\overline{B}(a, r)$ de centre a et de rayon r est le sous-ensemble de E défini par

$$\overline{B}(a, r) = \{x \in E \mid \|x - a\| \leq r\}$$

Soit $a \in E$.Un sous-ensemble $V \subset E$ est un *voisinage* de a dans E s'il existe $r > 0$ tel que la boule $B(a, r)$ de centre a et de rayon r, est incluse dans V.
On notera $\mathscr{V}(a)$ l'ensemble des voisinages de a.

Un *ouvert* de E est un souensemble de E qui est voisinage de chacun de ses points.

Définitions 1.1.9. *Convergence dans les espaces vectoriels normés.*

Une suite $(x_n)_{n\in\mathbb{N}}$ de vecteurs d'un espace vectoriel normés E *converge vers un vecteur* $x \in E$ si

$$\lim_{n\to+\infty} \|x_n - x\| = 0$$

ce qui s'écrit aussi :

$$\forall \varepsilon > 0 \ , \ \exists N \in \mathbb{N} \text{ tel que } n \geq N \Rightarrow \|x_n - x\| \leq \varepsilon$$

On dit qu'une suite $(x_n)_{n\in\mathbb{N}}$ de vecteurs d'un espace vectoriel normés E est *de Cauchy* si

$$\forall \varepsilon > 0 \ , \ \exists N \in \mathbb{N} \text{ tel que } p, q \geq N \Rightarrow \|x_p - x_q\| \leq \varepsilon$$

Proposition 1.1.10. *Toute suite convergente est de Cauchy et la réciproque est fausse en général.*

On dit qu'un espace vectoriel normé est *complet* si toute suite de Cauchy converge.
Les espaces vectoriels normés complets sont appelés *espaces de Banach* et s'ils sont munis d'une norme hilbertienne *espaces de hilbert*.

Proposition 1.1.11. *Les espaces vectoriels normés de dimension finie sont complets.*

Une série $\sum u_n$ de vecteurs d'un espace vectoriel normés E *converge vers un vecteur* $s \in E$ si la suite des sommes partielles $(\sum_{i=0}^{n} u_i)_{n\in\mathbb{N}}$ converge vers s dans E, c'est à dire si

$$\lim_{n\to+\infty} \left\| \sum_{i=0}^{n} u_i - s \right\| = 0$$

ce qui s'écrit aussi :

$$\forall \varepsilon > 0 \ , \ \exists N \in \mathbb{N} \text{ tel que } n \geq N \Rightarrow \left\| \sum_{i=0}^{n} u_i - s \right\| \leq \varepsilon$$

On dit qu'une série $\sum u_n$ de vecteurs d'un espace vectoriel normés E est *normalement convergente* si la série numérique, à termes positifs, $\sum \|u_n\|$ est convergente.

Proposition 1.1.12. *Un espace vectoriel normé est complet si et seulement si toute série normalement convergente est convergente. Par contre, même dans un espace complet, une série convergente n'est pas nécessairement normalement convergente.*

Définitions 1.1.13. *Fonctions continues.*

Une fonction $f : D \subset E \to F$ définie sur une partie D d'un espace vectoriel normé E et à valeurs dans un espace vectoriel normé F est *continue en* $a \in D$ si :

$$\forall \varepsilon > 0, \exists \eta > 0 \text{ tel que } \|x - a\| \leq \eta \Rightarrow \|f(x) - f(a)\| \leq \varepsilon$$

Cette propriété s'exprime aussi en disant que pour toute boule $B(f(a), \varepsilon)$, il existe une boule $B(a, \eta)$ telle que $f(B(a, \eta)) \subset B(f(a), \varepsilon)$
f est continue en $a \in D$ si et seulement si pour toute suite $(x_n)_{n\in\mathbb{N}}$ convergeant vers a dans E, la suite $(f(x_n))_{n\in\mathbb{N}}$ converge vers $f(a)$ dans F.

On dit que f est *continue sur* D si f est continue en tout point $a \in D$. On notera $\mathscr{C}(D; F)$ l'ensemble des fonctions continues de $D \subset E$ dans F.
La somme, le produit et la composition de deux fonctions continues sont continues.

1.2 Applications différentiables

Définition 1.2.1. *1) Soit E et F deux espaces vectoriels normés, $a \in E$ et V un voisinage de a dans E. Une application $f : V \to F$ est différentiable en a s'il existe une application linéaire continue de E dans F, notée $f'(a)$, telle*

$$\forall h \in E \text{ tel que } a + h \in V \ , \ f(a+h) = f(a) + f'(a)h + \|h\| \, \varepsilon(h)$$

où $\lim_{h \to 0} \varepsilon(h) = 0$ *dans F.*

2) L'application linéaire continue de E dans F $f'(a)$ est appelée la dérivée (ou la différentielle de f en a et dans ce cas, elle sera notée $Df(a)$) ou encore dans certains cas l'application linéaire tangente à f en a.

3) L'application affine de E dans F, $x \to f(a)+f'(a)(x-a)$ est appelée l'application affine tangente à f en a.

Remarque 1.2.2. *On notera qu'une application différentiable en un point est en particulier continue en ce point.*

Exemple 1.2.3. *Une application réelle définie sur un voisinage V de $a \in \mathbb{R}$ est différentiable en a si elle est dérivable en a et sa dérivée est l'application linéaire de $\mathbb{R}$ dans $\mathbb{R}$ représentée par le réel $f'(a)$.*

Théorème 1.2.4. *(Dérivation suivant un vecteur).*
Soit $a \in E$ et V un voisinage de a dans E. Si $f : V \to F$ est différentiable en a, alors quel que soit $h \in E$, on a

$$f'(a)h = \lim_{t \to 0} \frac{f(a+th) - f(a)}{t}$$

Démonstration : Soit $h \in E$ donné. On écrit la définition de la différentiabilité de f au point a, en utilisant le vecteur th :

$$f(a+th) = f(a) + tf'(a)h + |t| \, \|h\| \, \varepsilon(th) = f(a) + tf'(a)h + |t| \, \varepsilon_1(t)$$

où on a posé $\varepsilon_1(t) = \|h\| \, \varepsilon(th)$.
Comme $\varepsilon_1(t) \to 0$ quand $t \to 0$, ceci implique bien que

$$f'(a)h = \lim_{t \to 0} \frac{f(a+th) - f(a)}{t}$$

■

Corollaire 1.2.5. *Soit $a \in E$ et V un voisinage de a dans E. Si $f : V \to F$ est différentiable en a, alors, $f'(a)$ est unique.*

Démonstration : L'application linéaire $f'(a)$ est en effet entièrement déterminée par les égalités du théorème 1.2.4. ■

Exemples 1.2.6. *1) Soient $E = \mathbb{R}^2$ et $F = \mathbb{R}$. L'application $f : \mathbb{R}^2 \to \mathbb{R}$ telle que $f(x, y) = xy$ est différentiable en tout point $(x_0, y_0) \in \mathbb{R}^2$ et sa dérivée en (x_0, y_0) est l'application linéaire de $\mathbb{R}^2$ dans $\mathbb{R}$ telle que*

$$f'(x_0, y_0)(h_1, h_2) = y_0 h_1 + x_0 h_2 = \begin{pmatrix} y_0 & x_0 \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}$$

2) Soient $E = F = \mathbb{R}^2$. L'application $f : \mathbb{R}^2 \to \mathbb{R}^2$ telle que $f(x, y) = (x + y, xy)$ est différentiable en tout point $(x_0, y_0) \in \mathbb{R}^2$ et sa dérivée en (x_0, y_0) est l'application linéaire de $\mathbb{R}^2$ dans $\mathbb{R}^2$ telle que

$$f'(x_0, y_0)(h_1, h_2) = (h_1 + h_2, x_0 h_2 + y_0 h_1) = \begin{pmatrix} 1 & 1 \\ y_0 & x_0 \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}$$

Démonstration : *1)* En effet, on peut écrire :

$$f(x_0 + h_1, y_0 + h_2) - f(x_0, y_0) - y_0 h_1 - x_0 h_2 =$$

$$(x_0 + h_1)(y_0 + h_2) - x_0 y_0 - y_0 h_1 - x_0 h_2 = h_1 h_2$$

et on a :

$$|h_1 h_2| \leq \frac{1}{2}(|h_1|^2 + |h_2|^2) = \frac{1}{2} \|(h_1, h_2)\|_2^2$$

2) De la même façon, on écrit :

$$f(x_0 + h_1, y_0 + h_2) - f(x_0, y_0) - (h_1 + h_2, x_0 h_2 + y_0 h_1) =$$

$$(x_0 + h_1 + y_0 + h_2, (x_0 + h_1)(y_0 + h_2)) - (x_0 + y_0, x_0 y_0) - (h_1 + h_2, x_0 h_2 + y_0 h_1) =$$
$$-(0, h_1 h_2)$$

et

$$\|(0, h_1 h_2)\|_2 \leq \frac{1}{2}(|h_1|^2 + |h_2|^2) = \frac{1}{2} \|(h_1, h_2)\|_2^2$$

■

Proposition 1.2.7. *1) Toute application constante sur E est différentiable, de dérivée nulle en tout point de E.*

2) Soient E et F deux espaces vectoriels normés. Alors toute application linéaire continue $T : E \to F$ est différentiable en tout point $x \in E$ et sa dérivée en chaque point est l'application linéaire constante, égale à T.

3) Toute application affine continue définie par $\forall x \in E$, $Ax = b + Tx$, où $b \in F$ et $T \in \mathscr{L}(E; F)$ est différentiable en tout point $x \in E$ et sa dérivée en chaque point est l'application linéaire constante, égale à T.

4) Soient E et F deux espaces vectoriels normés. Alors toute application bilinéaire continue $B : E \times E \to F$ est différentiable en tout point $(x, y) \in E \times E$ et sa dérivée au point $(x, y) \in E \times E$ est l'application linéaire : $h \to B(x, h) + B(h, y)$.

Démonstration : *1)* est évident

2) Soit $x \in E$. On écrit :

$$T(x+h) - T(x) = T(h)$$

D'où $T'(x) = T$ pour tout $x \in E$.
3) se démontre comme *2)*.
4) On écrit :

$$B(x+h, y+k) - B(x,y) = B(x+h, y+k) - B(x+h, y) + B(x+h, y) - B(x,y) =$$

$$B(x+h, k) + B(h, y) = B(x,k) + B(h,y) + B(h,k)$$

Ceci donne le résultat puisque $\|B(h,k)\| \leq \|B\| \, \|h\| \, \|k\|$. ■

Définition 1.2.8. *Le vecteur de F défini par $\lim_{t \to 0} \dfrac{f(a+th) - f(a)}{t}$ quand il existe est appelé dérivée de f suivant le vecteur h.*

On peut donc interpréter le théorème 1.2.4 en disant que toute fonction différentiable en a admet des dérivées suivant tout vecteur de E.

La réciproque de cette propriété est fausse :

Exemple 1.2.9. *Soient $E = \mathbb{R}^2$ et $F = \mathbb{R}$. L'application $f : \mathbb{R}^2 \to \mathbb{R}$ telle que*

$$f(x,y) = \frac{x\,|y|}{(x^2+y^2)^{1/2}}$$

admet des dérivées suivant tout vecteur en $(0,0)$ mais n'est pas différentiable en ce point.

Démonstration : En effet, on peut écrire :

$$f(th, tk) = t\frac{hk}{(h^2+k^2)^{1/2}} = tf(h,k)$$

Donc

$$\frac{f(th,tk) - f(0,0)}{t} = f(h,k)$$

et f admet une dérivée suivant tout vecteur (h,k) qui vaut $f(h,k)$.
Par contre, si f était différentiable en $(0,0)$, il existerait deux réels notés $f'_x(0,0)$ et $f'_y(0,0)$ tels que

$$f'(0,0)(h,k) = hf'_x(0,0) + kf'_x(0,0)$$

Mais d'après la démonstration précédente,

$$f'(0,0)(1,0) = f'_x(0,0) = f(1,0) = 0$$

$$f'(0,0)(0,1) = f'_y(0,0) = f(0,1) = 0$$

Ceci impliquerait que pour tout (h,k), $f(h,k) = 0$ ce qui n'est pas vrai. f n'est donc pas différentiable en $(0,0)$.

■

Proposition 1.2.10. *Si f est différentiable en $a \in E$, elle est aussi différentiable si on remplace les normes de E et F par des normes équivalentes et $f'(a)$ ne change pas lors d'un tel remplacement.*

Démonstration : Il suffit de remarquer que, par définition de l'équivalence des normes, la quantité

$$\frac{\|f(a+h) - f(a) - f'(a)h\|_F}{\|h\|_E}$$

est majorée et minorée par une constante fois la même expression calculée avec les autres normes. ∎

Théorème 1.2.11. *(Composition avec une translation)*
Soient E et F deux espaces vectoriels normés et $a \in E$. Soit $f : V \to F$ une application d'un voisinage V de a dans E, à valeurs dans F. Soit $x_0 \in E$ et f_0 l'application de $x_0 + V$ à valeurs dans F telle que

$$\forall x \in x_0 + V \ , \ f_0(x) = f(x - x_0)$$

Alors, si f est différentiable en a, f_0 est différentiable en $a + x_0$ et

$$f_0'(a + x_0) = f'(a)$$

Démonstration : Si f est différentiable en a, on écrit

$$f_0(a + x_0 + h) - f_0(a + x_0) = f(a + h) - f(a) = f'(a)h + \|h\| \varepsilon(h)$$

avec $\varepsilon \to 0$ quand $h \to 0$.
D'où la différentiabilité de f_0 en $a + x_0$ et l'égalité $f_0'(a + x_0) = f'(a)$. ∎

Théorème 1.2.12. *(Composition avec une application linéaire continue)*
Soient E, F et G trois espaces vectoriels normés et $a \in E$. Soit $f : V \to F$ une application d'un voisinage V de a dans E, à valeurs dans F, différentiable en a.

1) Si $u \in \mathscr{L}(F;G)$, $u \circ f$ est dérivable en a et

$$(u \circ f)'(a) = u \circ f'(a)$$

2) Si $v \in \mathscr{L}(G;E)$ et si $v(b) = a$, $f \circ v$ est dérivable en b et

$$(f \circ v)'(b) = f'(a) \circ v$$

Démonstration : *1)* La relation

$$f(a + h) - f(a) - f'(a)h = \|h\| \varepsilon(h)$$

avec $\lim_{h \to 0} \varepsilon(h) = 0$ entraine :

$$(u \circ f)(a + h) - (u \circ f)(a) - (u \circ f'(a))h = \|h\| \varepsilon_1(h)$$

avec $\varepsilon_1(h) = u \circ \varepsilon(h)$ et donc $\lim_{h \to 0} \varepsilon_1(h) = 0$
2) On a

$$(f \circ v)(b + k) = f(a + v(k)) = f(a) + f'(a)v(k) + \|v(k)\| \varepsilon(v(k)) =$$
$$(f \circ v)(b) + (f'(a) \circ v)(k) + \|k\| \varepsilon_2(k)$$

avec $\|\varepsilon_2(k)\| \leq \|\varepsilon(v(k))\| \, \|v\| \to 0$ quand $k \to 0$. ∎

Corollaire 1.2.13. *1) Si u est un isomorphisme de F dans G, alors une condition nécessaire et suffisante pour que f soit différentiable en a est que $u \circ f$ soit différentiable en a.*

2) Si v est un isomorphisme de G dans E, une condition nécessaire et suffisante pour que f soit différentiable en a est que $f \circ v$ soit différentiable en $v^{-1}(a)$.

Définition 1.2.14. *Soit $f : V \subset E \to F$. Pour tout $\varphi \in F'$, l'application*

$$x \to f_\varphi(x) = < f(x), \varphi > = \varphi\big(f(x)\big) = \varphi \circ f(x)$$

de E dans $\mathbb{K}$ est appelée la fonction scalarisée de f.

Corollaire 1.2.15. *(Scalarisation)*
Soit V un voisinage de a dans E et $f : V \to F$. Si f est différentiable en a, alors pour tout $\varphi \in F'$, l'application $x \to f_\varphi(x) = < f(x), \varphi >$ de E dans $\mathbb{K}$ est différentiable en a et

$$\forall h \in E \ , \ f'_\varphi(a)h = < f'(a)h, \varphi >$$

Démonstration : Il suffit d'appliquer le théorème 1.2.12 à l'application linéaire

$$\varphi \in \mathscr{L}(F; \mathbb{R}) = F' \text{ telle que } \forall x \in E \ , \ \varphi(x) = < x, \varphi >$$

■

1.3 Cas particuliers

Exemple 1.3.1. $E = \mathbb{R}$, $F = \mathbb{R}$.

En identifiant $\mathbb{R}$ et son dual, on retrouve la définition usuelle de la dérivabilité des fonctions d'une variable réelle à valeurs réelles.

Exemple 1.3.2. $E = \mathbb{R}$, $F = \mathbb{R}^p$.

Si f est différentiable en a, sa dérivée au point a, $f'(a)$ est un élément de $\mathscr{L}(\mathbb{R}; \mathbb{R}^p)$. En termes de matrice, les éléments de $\mathscr{L}(\mathbb{R}; \mathbb{R}^p)$ sont représentés par des matrices $p \times 1$, c'est à dire par des matrices colonne à p éléments. On peut donc l'identifier à un élément de $\mathbb{R}^p$.
Posons

$$\forall x \in \mathbb{R} \ , \ f(x) = \begin{pmatrix} f^1(x) \\ f^2(x) \\ . \\ . \\ f^p(x) \end{pmatrix}$$

D'après le théorème de scalarisation 1.2.15 appliqué successivement aux vecteurs de la base canonique de $\mathbb{R}^p$

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ . \\ . \\ 0 \end{pmatrix} \ , \ e_2 = \begin{pmatrix} 0 \\ 1 \\ . \\ . \\ 0 \end{pmatrix} \ , ..., \ e_p = \begin{pmatrix} 0 \\ 0 \\ . \\ . \\ 1 \end{pmatrix}$$

on trouve que pour tout $j = 1, 2, ..., p$, la j-ème coordonnée de $f'(a)$ est la dérivée de la j-ème coordonnée f^j de f en a, ce qui s'écrit encore :

$$f'(a) = \begin{pmatrix} (f^1)'(a) \\ (f^2)'(a) \\ . \\ . \\ (f^p)'(a) \end{pmatrix}$$

Dans le cas où F est un espace vectoriel de dimension p, muni d'une base, le même argument prouve que si f est différentiable en a, ses coordonnées dans cette base sont dérivables et la formule ci-dessus est encore valable.

Exemple 1.3.3. $E = \mathbb{R}^n$, $F = \mathbb{R}$.

Si f est différentiable en a, sa dérivée au point a, $f'(a)$ est un élément de$\mathscr{L}(\mathbb{R}^n; \mathbb{R}) = (\mathbb{R}^n)'$. En termes de matrice, les éléments de$\mathscr{L}(\mathbb{R}^n; \mathbb{R})$ sont représentés par des matrices $1 \times n$, c'est à dire par des matrices ligne à n éléments. On peut donc l'identifier à un élément de $\mathbb{R}^n$.

Soit $\{e_1, e_2, ..., e_n\}$ la base canonique de $\mathbb{R}^n$. Alors, par définition la i-ième composante de $f'(a)$ soit $f'(a)_i$ est égale à la valeur de la forme linéaire $f'(a)$ sur e_i. On a donc l'égalité :

$$f'(a)_i = f'(a)e_i$$

On peut donc dire également que $f'(a)_i$ est la dérivée de f en a selon le vecteur e_i, soit encore :

$$f'(a)_i = \lim_{t \to 0} \frac{f(a + te_i) - f(a)}{t} = \lim_{t \to 0} \frac{f(a_1, ..., a_i + t, ...a_n) - f(a_1, ...a_i, ..., a_n)}{t}$$

Cette dernière expression est la dérivée en a_i de l'application partielle f_i, de $\mathbb{R}$ dans $\mathbb{R}$:

$$f_i : x_i \to f(a_1, ...x_i, ..., a_n)$$

Définition 1.3.4. *$f_i'(a)$ est appelée dérivée partielle de f en a, selon le vecteur e_i. On la notera $f_i'(a)$.*

La représentation matricielle de $f'(a)$ est alors :

$$f'(a) = (\, {f_1}'(a) \quad {f_2}'(a) \quad ... \quad {f_n}'(a)\,)$$

On a aussi démontré la proposition suivante :

Proposition 1.3.5. *Si $f : \mathbb{R}^n \to \mathbb{R}$ est différentiable en a, elle admet des dérivées partielles suivant tous les vecteurs de la base canonique, données par :*

$$f_i'(a) = f'(a)e_i$$

La réciproque de cette proposition est fausse, c'est à dire qu'une fonction $f : \mathbb{R}^n \to \mathbb{R}$ peut avoir des dérivées partielles selon tous les vecteurs de la base canonique en un point a sans être différentiable en ce point.

Exemple 1.3.6. *Soit* $f = \dfrac{x\,|y|}{(x^2+y^2)^{1/2}}$. *D'après l'exemple 1.2.9, f admet des dérivées partielles dans toutes les directions mais n'est pas différentiable à l'origine.*

Cette étude reste valable dans le cas des espaces de dimension finie, munis d'une base fixée comme on l'a remarqué au paragraphe 1.1, rappels 1.1.7

Exemple 1.3.7. *Soit $f : E \to \mathbb{R}$. Si E est en dualité avec lui-même en tant qu'espace normé, c'est à dire si $E = E'$, ce qui est le cas d'un espace de Hilbert, alors il existe un vecteur de E noté grad $f(a)$ tel que*

$$\forall h \in E \ , \ f'(a)h =< h, grad\ f(a) >$$

Exemple 1.3.8. *$E = \mathbb{R}^n$, $F = \mathbb{R}^p$. Si $f : \mathbb{R}^n \to \mathbb{R}^p$ est différentiable en a, alors $f'(a) \in \mathscr{L}(\mathbb{R}^n, \mathbb{R}^p)$. Cette application linéaire est représentée par une matrice $p \times n$ à p lignes et n colonnes, appellée matrice jacobienne de f en a et notée $J_f(a)$.*

Soient f^j , $1 \leq j \leq p$ les coordonnées de l'application f. Alors, d'après le théorème de scalarisation 1.2.15 , on a :

$$(f'(a))^j = (f^j)'(a)$$

En d'autres termes, la j-ième coordonnée de la différentielle de f en a est égale à la différentielle de f^j en a.

D'autre part, comme dans l'exemple précédent, si $\{e_1, e_2, ..., e_n\}$ est la base canonique de $\mathbb{R}^n$, alors, par définition la i-ième composante de $f'(a)$ soit $f'(a)_i$ est égale à la valeur de l'application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^p$ $f'(a)$ sur e_i. On a donc l'égalité :

$$f'(a)_i = f'(a)e_i$$

On peut donc dire également que $f'(a)_i$ est la dérivée de f en a selon le vecteur e_i, soit encore comme précédemment :

$$f'(a)_i = \lim_{t\to 0} \frac{f(a+te_i) - f(a)}{t} = \lim_{t\to 0} \frac{f(a_1, ..., a_i + t, ...a_n) - f(a_1, ...a_i, ..., a_n)}{t}$$

Cette dernière expression est la dérivée en a_i de l'application partielle f_i, de $\mathbb{R}$ dans $\mathbb{R}^p$:

$$x_i \to f(a_1, ...x_i, ..., a_n)$$

On notera donc cette dérivée partielle en a selon le vecteur e_i : $f'_i(a)$.

Matriciellement, ceci s'exprime par la relation suivante :

$$J_f(a) = \begin{pmatrix} f_1^{1'}(a) & f_2^{1'}(a) & ... & f_n^{1'}(a) \\ f_1^{2'}(a) & f_2^{2'}(a) & ... & f_n^{2'}(a) \\ . & . & ... & . \\ . & & & \\ f_1^{p'}(a) & f_2^{p'}(a) & ... & f_n^{p'}(a) \end{pmatrix}$$

On peut aussi dire que les colonnes de $J_f(a)$ sont les dérivées partielles de f selon les vecteurs de la base canonique de $\mathbb{R}^n$ et que les lignes de f sont les dérivées des coordonnées de f dans la base canonique de $\mathbb{R}^p$.

Comme précédemment (cf. 1.1.7), cette représentation s'applique aussi au cas de deux espaces vectoriels normés de dimension finie, munis chacun d'une base fixée.

1.4 Règles de différentiation

Théorème 1.4.1. *Soient E et F deux espaces vectoriels normés, $a \in E$ et V un voisinage de a dans E. Soient f et g deux fonctions définies de V dans F et λ et μ deux scalaires.*
Alors, si f et g sont différentiables en a, $\lambda f + \mu g$ l'est aussi et

$$(\lambda f + \mu g)'(a) = \lambda f'(a) + \mu g'(a)$$

Démonstration : Par hypothèse, on a

$$f(a+h) = f(a) + f'(a)h + \varepsilon_1(h)\|h\| \text{ avec } \varepsilon_1(h) \to 0 \text{ quand } h \to 0$$

$g(a+h) = g(a) + g'(a)h + \varepsilon_2(h)\|h\|$ avec $\varepsilon_2(h) \to 0$ quand $h \to 0$ D'où

$$(\lambda f + \mu g)(a+h) = (\lambda f + \mu g)(a) + \lambda f'(a) + \mu g'(a)h + \lambda\varepsilon_1(h)\|h\| + \mu\varepsilon_2(h)\|h\|$$

En posant $\varepsilon(h) = \lambda\varepsilon_1(h) + \mu\varepsilon_2(h)$, on voit que $\varepsilon(h) \to 0$ quand $h \to 0$ et ceci prouve le théorème 1.4.1. ■

Théorème 1.4.2. *Soient E, F, G trois espaces vectoriels normés, $a \in E$ et V un voisinage de a dans E. Soient f une fonction définie dans V à valeurs dans F et g une fonction définie dans un voisinage W de $f(a)$ dans F, à valeurs dans G.*
Alors, si f est différentiable en a et g est différentiable en $f(a)$, $g\circ f$ est différentiable en a et

$$(g\circ f)'(a) = g'(f(a))\circ f'(a)$$

Démonstration : Par définition de la différentiabilité de f en a, on a :

$$f(a+h) = f(a) + f'(a)h + \varepsilon_1(h)\|h\| \text{ avec } \varepsilon_1(h) \to 0 \text{ quand } h \to 0$$

ou encore en posant $k = f'(a)h + \varepsilon_1(h)\|h\|$ (k est donc un vecteur de F),

$$f(a+h) = f(a) + k$$

En appliquant l'hypothèse de différentiabilité de g en $f(a)$, on peut écrire :

$$(g\circ f)(a+h) = g(f(a+h)) = g(f(a)+k) = g(f(a)) + g'(f(a))k + \varepsilon_2(k)\|k\|$$

avec $\varepsilon_2(k) \to 0$ quand $k \to 0$
soit encore :

$$(g\circ f)(a+h) = g(f(a)) + g'(f(a))(f'(a)h + \varepsilon_1(h)\|h\|) + \varepsilon_2(k)\|k\| =$$

$$(g\circ f)(a) + g'(f(a))\circ f'(a)h + \varepsilon(h)\,\|h\|$$

avec

$$\varepsilon(h)\,\|h\| = g'\big(f(a)\big)\varepsilon_1(h)\,\|h\| + \varepsilon_2(k)\,\|f'(a)h + \varepsilon_1(h)\,\|h\|\|$$

D'où $\|\varepsilon(h)\| \leq \|g'(f(a)\|\,\|\varepsilon_1(h)\| + \|\varepsilon_2(k)\|\,\|\varepsilon_1(h)\|$.
Si $h \to 0$, alors $k \to 0$ et $\varepsilon_2(k) \to 0$, et par suite, $\varepsilon(h) \to 0$.
Ceci prouve le théorème 1.4.2. ■
On remarque qu'on retrouve le cas de composition avec des applications linéaires continues, Théorème 1.2.12.

Corollaire 1.4.3. *Soit f une application bijective de V dans W, différentiable ainsi que son inverse. Alors, en posant $b = f(a)$, $f'(a)$ est bijective et on a :*

$$\left(f^{-1}\right)'(b) = \Big(f'\big(f^{-1}(b)\big)\Big)^{-1}$$

Application 1.4.4. *Si $E = \mathbb{R}^n$, $F = \mathbb{R}^p$, et $G = \mathbb{R}^q$, alors on a la relation suivante entre les matrices jacobiennes de f en a, de g en $f(a)$ et de $g\circ f$ en a :*

$$J_{g\circ f}(a) = J_g(f(a))J_f(a)$$

Théorème 1.4.5. *Soient E et F deux espaces vectoriels normés, $a \in E$ et V un voisinage de a dans E. Soit f une fonction définie dans V, à valeurs dans F, différentiable en a. Soit H un sous-espace vectoriel de E. Alors,*

1) si H contient a, la restriction $f|_H$ de f à $V \cap H$ est différentiable en a et

$$(f|_H)'(a) = f'(a)|_H$$

2) La fonction $\phi : H \to F$ telle que $\phi(h) = f(a+h)$ est différentiable en 0 et

$$\phi'(0) = f'(a)$$

Démonstration : C'est une conséquence immédiate des définitions. ■

Théorème 1.4.6. *Soient E et F deux espaces vectoriels normés, $a \in E$ et V un voisinage de a dans E. Soit f une fonction définie dans V, à valeurs dans F^p. On pose pour tout $x \in V$,*

$$f(x) = \begin{pmatrix} f^1(x) \\ f^2(x) \\ . \\ . \\ . \\ f^p(x) \end{pmatrix}$$

Pour que f soit différentiable en a il faut et il suffit que chaque $f^j, 1 \leq j \leq p$, soit différentiable en a et alors :

$$f'(a) = \begin{pmatrix} (f^1)'(a) \\ (f^2)'(a) \\ . \\ . \\ (f^p)'(a) \end{pmatrix}$$

DÉMONSTRATION : Soit
$$y = \begin{pmatrix} y^1 \\ y^2 \\ . \\ . \\ y^p \end{pmatrix} \in F^p$$
On note p_j la projection de F^p sur F, qui envoie y sur sa j-ième coordonnée, c'est à dire : $p_j(y) = y^j$.
On a alors : $f^j = p_j \circ f$. Par suite, comme p_j est une application linéaire continue, si f est différentiable en a, f^j l'est aussi.

Réciproquement, supposons f^j différentiable en a pour tout j, c'est à dire avec les notations habituelles :
$f^j(a+h) = f^j(a) + (f^j)'(a)h + \varepsilon_j(h)\,\|h\|$, avec $\varepsilon_j(h) \to 0$ quand $h \to 0$
On en déduit ;
$$f(a+h) = f(a) + \begin{pmatrix} (f^1)'(a) \\ (f^2)'(a) \\ . \\ . \\ (f^p)'(a) \end{pmatrix} h + \begin{pmatrix} \varepsilon_1(h) \\ \varepsilon_2(h) \\ . \\ . \\ \varepsilon_p(h) \end{pmatrix} \|h\|$$

avec $\varepsilon(h) = \begin{pmatrix} \varepsilon_1(h) \\ \varepsilon_2(h) \\ . \\ . \\ \varepsilon_p(h) \end{pmatrix} \to 0$ quand $h \to 0$. ∎

1.5 Théorème des Accroissements Finis.

Définition 1.5.1. *Soient E et F deux espaces vectoriels normés.*

1) Une fonction f définie sur un ouvert Ω de E est dite différentiable sur Ω si elle est différentiable en tout point de Ω.

2) La fonction $f' : \Omega \to \mathscr{L}(E,F)$ est appellée la dérivée de f sur Ω.

3) Si f est différentiable sur Ω et si sa dérivée $f' : \Omega \to \mathscr{L}(E,F)$ est continue sur Ω on dira que f est continûment différentiable sur Ω. On dira aussi que f est de classe C^1 sur Ω.

Rappelons le résultat classique pour les fonctions d'une variable scalaire, à valeurs scalaires :

Théorème 1.5.2. *(Théorème des Accroissements Finis scalaire)*
Soit f une fonction continue d'un intervalle $[a,b]$, à valeurs dans $\mathbb{K}$, dérivable sur $]a,b[$. Alors il existe $c \in]a,b[$ tel que :
$$f(b) - f(a) = f'(c)(b-a)$$
On en déduit une première extension pour les fonction d'une variable vectorielle, à valeurs scalaires :

Théorème 1.5.3. *(Théorème des Accroissements Finis pour les fonctions d'une variable vectorielle à valeurs scalaires)*
Soit E un espace vectoriel normé et f une application différentiable d'un ouvert Ω de E, à valeurs dans $\mathbb{K}$. Si le segment $[a, a+h] \subset \Omega$, il existe $\theta \in]0,1[$ tel que :

$$f(a+h) - f(a) = f'(a+\theta h)h$$

ce qui s'écrit aussi : il existe $c \in]a, a+h[$ tel que :

$$f(a+h) - f(a) = f'(c)h$$

Démonstration : On applique le théorème des Accroissements Finis scalaire 1.5.2 à la fonction $F : t \to f(a+th)$: cette fonction est définie sur $[0,1]$, dérivable sur $]0,1[$ et $F'(t) = f'(a+th)h$. Il existe donc $c \in]a, a+h[$ tel que

$$F(1) - F(0) = F'(c)$$

ce qui prouve le résultat. ■

Remarque 1.5.4. *Le résultat précédent est faux si la fonction f est à valeurs dans un espace vectoriel de dimension supérieure ou égale à 2.*

On peut à partir de ces résultats obtenir le théorème des Accroissements Finis pour les fonctions d'une variable vectorielle, à valeurs vectorielles :

Théorème 1.5.5. *(Théorème des Accroissements Finis)*
Soient E et F deux espaces vectoriels normés et f une application différentiable dans un ouvert Ω de E, à valeurs dans F. Si $[a, a+h] \subset \Omega$, on a :

$$\|f(a+h) - f(a)\| \leq \|h\| \sup_{x \in]a, a+h[} \|f'(x)\|$$

Démonstration : Soit $v \in F'$. L'application $v \circ f : \Omega \to \mathbb{K}$ est telle que

$$v \circ f(x) = v(f(x)) = < f(x), v >$$

C'est donc une application différentiable sur Ω, à valeurs scalaires. On peut donc écrire d'après le théorème 1.5.3 :

$$< f(a+h) - f(a), v > = v \circ f(a+h) - v \circ f(a) =$$
$$(v \circ f)'(a+\theta h)(h)$$

avec $\theta \in]0,1[$ (qui dépend de v). D'où

$$< f(a+h) - f(a), v > = v \circ f'(a+\theta h)(h) \leq \|v\| \, \|f'(a+\theta h)\| \, \|h\| \leq$$
$$\|v\| \, \|h\| \sup\{\|f'(a+\theta h)\| \;\mid\; \theta \in]0,1[\}$$

On en déduit donc :

$$\|f(a+h) - f(a)\| = \sup\{|< f(a+h) - f(a), v >| \;\mid\; \|v\| \leq 1\} \leq$$
$$\|h\| \sup\{\|f'(a+\theta h)\| \;\mid\; \theta \in]0,1[\}$$

et ceci prouve le théorème 1.5.5.
On notera que, pour évaluer la norme de $f(a+h) - f(a)$, on a utilisé la propriété suivante :

$$\forall x \in E \ , \ \|x\| = \sup\{|< x, v >| \;\mid\; v \in E' \ , \ \|v\| \leq 1\}$$

Cette propriété est vraie dans tous les espaces de Banach grâce au théorème de Hahn Banach (hors programme). Il existe des démonstrations de ce résultat qui n'utilisent pas le théorème de Hahn Banach mais qui sont plus difficiles. ■

Corollaire 1.5.6. *Sous les mêmes hypothèses,*

1) Si $T \in \mathscr{L}(E,F)$,

$$\|f(a+h) - f(a) - Th\| \leq \|h\| \sup\{\|f'(x) - T\| \mid x \in]a, a+h[\}$$

2)

$$\|f(a+h) - f(a) - f'(a)h\| \leq \|h\| \sup\{\|f'(x) - f'(a)\| \mid x \in]a, a+h[\}$$

3) *Si Ω est connexe et si $f'(x) = 0$ pour tout $x \in \Omega$, alors f est constante sur Ω.*

Pour démontrer la propriété *3)*, on utilisera l'équivalence suivante :

Ω est connexe si et seulement si il existe une ligne polygonale joignant deux points arbitraires de Ω.

Proposition 1.5.7. *Soient E et F deux espaces vectoriels normés et f une application différentiable dans un ouvert Ω de E, à valeurs dans F. Si f' est continue en a, on a, pour tous h et $k \in E$:*

$$f(a+h) - f(a+k) = f'(a)(h-k) + \|h-k\|\, \varepsilon(h-k)$$

avec $\varepsilon(h,k) \to 0$ quand $h, k \to 0$.

Définition 1.5.8. *Dans ce cas, on dit que f est strictement différentiable en a.*

Démonstration : Si l'on définit la fonction $\varepsilon(h,k)$ par la formule de l'énoncé, le corollaire 1.5.6 implique :

$$\|\varepsilon(h,k)\| \leq \sup\{\|f'(x) - f'(a+k)\| \mid x \in]a+h, a+k[\}$$

Or si $h, k \to 0$,

$$x \in]a+h, a+k[\Rightarrow \|x-a\| \to 0$$

Comme par hypothèse, f' est continue en a, la quantité

$$\sup\{\|f'(x) - f'(a+k)\| \mid x \in]a+h, a+k[\}$$

tend vers 0, ce qui implique bien que $\varepsilon(h,k) \to 0$ quand $h, k \to 0$. ■

Une autre propriété importante des applications différentiables à dérivée continue, sur les produits d'espaces vectoriels normés, est liée à la notion de dérivées partielles, qui généralise le cas de $E = \mathbb{R}^n$.

Définition 1.5.9. *Soit $E = E_1 \times E_2 \times ... \times E_n$, où $E_1, E_2, ..., E_n$ sont des espaces vectoriels normés. Soit f une application définie dans un ouvert Ω de E, à valeurs dans un espace vectoriel normé F. Si l'application partielle $x_i \to f(a_1, ..., a_{i-1}, x_i, a_{i+1}, ...a_n)$, de E_i dans F est différentiable en a_i, sa dérivée en a_i sera appelée dérivée partielle de f en a. On la notera $f'_i(a)$ ou $\dfrac{\partial f}{\partial x_i}$.*

Théorème 1.5.10. *Soit $E = E_1 \times E_2 \times ... \times E_n$, où $E_1, E_2, ..., E_n$ sont des espaces vectoriels normés, Ω un ouvert de E et f une application définie dans Ω, à valeurs dans un espace vectoriel normé F. Si f est différentiable en $a \in \Omega$, alors f a des dérivées partielles en a et l'application linéaire $f'(a) \in \mathscr{L}(E,F) = \mathscr{L}(E_1 \times E_2 \times ... \times E_n, F)$ s'écrit*

$$f'(a) = \begin{pmatrix} f'_1(a) & f'_2(a) & . & . & . & f'_n(a) \end{pmatrix}$$

où chaque $f'_i(a)$ est un élément de $\mathscr{L}(E_i, F)$.

Démonstration : Elle est analogue à celle du cas $E = \mathbb{R}^n$, cf. Exemple 1.3.3. ■

Remarque 1.5.11. *On a déja vu (dans le cas où $E = \mathbb{R}^n$), que la réciproque de ce théorème est fausse. Par contre, on a la propriété fondamentale suivante :*

Théorème 1.5.12. *Soient $E = E_1 \times E_2 \times ... \times E_n$ et F deux espaces vectoriels normés et f une application différentiable dans un ouvert Ω de E, à valeurs dans F. Alors, f est continûment différentiable sur Ω si et seulement si f admet en tout point $a \in \Omega$, des dérivées partielles continues.*

Démonstration : Si f' existe sur Ω et est continue, il en est évidemment de même pour les f'_i. ■

Réciproquement, supposons que les f'_i existent et soient continues sur Ω. Pour simplifier, plaçons-nous dans le cas où $n = 2$, le cas général s'en déduit sans difficulté.
Soit $h \in E$. On écrit :

$$f(a+h) - f(a) = f(a_1 + h_1, a_2 + h_2) - f(a_1, a_2) =$$

$$f(a_1 + h_1, a_2 + h_2) - f(a_1 + h_1, a_2) + f(a_1 + h_1, a_2) - f(a_1, a_2)$$

Par le théorème des accroissements finis 1.5.6, on a :

$$f(a_1 + h_1, a_2 + h_2) - f(a_1 + h_1, a_2) = f'_2(a_1, a_2)h_2 + \varepsilon_1(h)\,\|h\|$$

avec $\|\varepsilon_1(h)\| \leq \sup_{\theta \in]0,1[} \|f'_2(a_1 + h_1, a_2 + \theta h_2) - f'_2(a_1, a_2)\|$.
D'où $\varepsilon_1(h) \to 0$ quand $h \to 0$.
On a aussi :

$$f(a_1 + h_1, a_2) - f(a_1, a_2) = f'_1(a_1, a_2)h_1 + \varepsilon_2(h_1)\,\|h_1\|$$

avec $\|\varepsilon_2(h_1)\| \to 0$ quand $h_1 \to 0$.
Ainsi f est différentiable en a et

$$f'(a) = \begin{pmatrix} f'_1(a) & f'_2(a) \end{pmatrix}$$

Comme f'_1 et f'_2 sont continues, il en est de même de f'. ■

Donnons deux conséquences immédiates du Théorème des Accroissements finis, qui seront utiles dans le chapitre sur les équations différentielles :

Définition 1.5.13. *Soient E et F deux espaces vectoriels normés et f une application différentiable dans un ouvert Ω de E, à valeurs dans F. On dira que f est lipschitzienne s'il existe une constante $M > 0$ telle que :*

$$\forall x, y \in E\ ,\ \|f(x) - f(y)\| \leq M\,\|x - y\|$$

Corollaire 1.5.14. *Soient E et F deux espaces vectoriels normés et f une application différentiable dans un ouvert Ω de E, à valeurs dans F. Alors, f est lipschitzienne si et seulement si sa dérivée f' est bornée sur E.*

Corollaire 1.5.15. *Si de plus f est définie sur un produit d'espaces $E = E_1 \times E_2 \times ... \times E_n$ et si elle est continûment différentiable, alors f est lipschitzienne si et seulement si ses dérivées partielles f'_i sont bornées sur E.*

1.6 Un exemple fondamental

Théorème 1.6.1. *Soit $I = [a,b]$ un intervalle de $\mathbb{R}$, E et F deux espaces de Banach, $\mathscr{C}(I;E)$(resp. $\mathscr{C}(I,F)$) l'espace des fonctions continues sur I, à valeurs dans E (resp. F) muni de la norme de la convergence uniforme : $\|X\|_\infty = \sup_{t\in T}\|X(t)\|_E$, (resp. $\|Y\|_\infty = \sup_{t\in T}\|Y(t)\|_F$).*

Soit L une application continue sur un ouvert Ω de $I\times E$, à valeurs dans F, possédant une dérivée partielle L'_x par rapport à la deuxième variable, continue sur Ω.

On considère l'application Φ qui à $X \in \mathscr{C}(I;E)$, dont le graphe est contenu dans Ω, associe $\Phi(X) \in \mathscr{C}(I;F)$ définie par :

$$\forall x \in I\ ,\ \Phi(X)(t) = L\big(t, X(t)\big)$$

Alors Φ est continûment différentiable et sa dérivée $\Phi'(X)$ est l'application linéaire continue qui à $\delta X \in \mathscr{C}(I;E)$ associe $\delta Y \in \mathscr{C}(I;F)$, définie par :

$$\forall x \in I\ ,\ \delta Y(t) = L'_x\big(t, X(t)\big)\delta X(t)$$

Démonstration : Vérifions tout d'abord que l'ensemble A des $X \in \mathscr{C}(I;E)$, dont le graphe est contenu dans Ω est ouvert : soit $X \in A$ et soit G_X son graphe. Comme X est continu et que I est compact, G_X est compact dans $I\times E$. Comme $G_X \subset \Omega$, chaque point $\big(t, X(t)\big) \in G_X$ a un voisinage V_t de la forme $I_t \times B\big(X(t), r_t\big)$ inclus dans Ω.

La famille $\{V_t, t\in I\}$ recouvre G_X donc par compacité, il existe un sous-recouvrement fini $\{V_{t_1}, V_{t_2}, ..., V_{t_n}\}$. Soit $r = \inf\{r_{t_k} \mid k = 1,2,...n\}$.
Alors, si $\|X' - X\|_\infty \le r$, le graphe de X' est inclus dans Ω. Par suite, A contient la boule de centre X et de rayon r, ce qui prouve bien que cet ensemble est ouvert.
D'autre part, l'application $\mathbb{L}$ telle que

$$\mathbb{L}(\partial X)(t) = L'_x\big(t, X(t)\big)\partial X(t)$$

est visiblement une application linéaire continue de $\mathscr{C}(I;E)$ dans $\mathscr{C}(I;F)$. On peut écrire :

$$\Phi\big(X + \partial X)(t) - \Phi(X)(t) - \mathbb{L}(\partial X)(t) =$$

$$L\big(t, X(t) + \partial X(t)\big) - L\big(t, X(t)\big) - L'_x\big(t, X(t)\big)\partial X(t)$$

D'où, pour tout $t \in [a,b]$:

$$\|\Phi(X + \partial X)(t) - \Phi(X)(t) - \mathbb{L}(\partial X)(t)\|$$

$$\|\partial X(t)\| \sup_{\theta\in[0,1]} \big\|L'_x\big(t, X(t) + \theta\partial X(t)\big) - L'_x\big(t, X(t)\big)\big\|$$

On a besoin d'un lemme sur les fonctions continues dans les espaces métriques :

lemme 1.6.2. *Soit f une application continue d'un espace métrique E dans un espace métrique F et K un compact de E.*
Alors, pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que

$$x \in K, x' \in E, d(x, x') \leq \eta \Rightarrow d\big(f(x), f(x')\big) \leq \varepsilon$$

Démonstration : On procède par l'absurde : si la conclusion du théorème était fausse, il existerait deux suites $(x_n)_{n\in\mathbb{N}}$ et $(x'_n)_{n\in\mathbb{N}}$ telles que

$$x_n \in K \text{ et } d(x_n, x'_n) \to 0, d\big(f(x_n), f(x'_n)\big) > \varepsilon$$

Comme par hypothèse, K est compact, quitte à extraire une sous-suite de la suite $(x_n)_{n\in\mathbb{N}}$, on peut supposer que cette suite converge. Soit ξ sa limite.
Alors, la suite $(x'_n)_{n\in\mathbb{N}}$ converge aussi vers ξ mais par contre, $d\big(f(x_n), f(x'_n)\big) > \varepsilon$, ce qui contredit la continuité de f. ■

On peut appliquer ce lemme pour démontrer le théorème 1.6.1 : soit $\varepsilon > 0$; d'après le lemme 1.6.2, appliqué à la fonction $(t, x) \to L'_x(t, x)$ avec $K = G_X$, il existe η tel que

$$\|\xi\| \leq \eta \Rightarrow \forall t \in I \ , \ \big\|L'_x\big(t, X(t) + \xi\big) - L'_x\big(t, X(t)\big)\big\| \leq \varepsilon$$

Pour tout ∂X tel que $\|\partial X\|_\infty \leq \varepsilon$, on aura alors

$$\forall t \in I \ , \ \|\Phi(X + \partial X)(t) - \Phi(X)(t) - \mathbb{L}(\partial X)(t)\| \leq \|\partial X\|_\infty \varepsilon$$

C'est à dire

$$\|\Phi(X + \partial X) - \Phi(X) - \mathbb{L}(\partial X)\|_\infty \leq \|\partial X\|_\infty \varepsilon$$

ce qui montre que Φ est différentiable au point X et que $\Phi'(X) = \mathbb{L}$.
Enfin, si $\|\partial X\|_\infty \leq \eta$, on a, en appliquant à nouveau le lemme, pour tout $t \in I$:

$$\big\|L'_x\big(t, X(t) + \partial X(t)\big) - L'_x\big(t, X(t)\big)\big\| \leq \varepsilon$$

et par suite,

$$\|\Phi'(X + \partial X) - \Phi'(X)\|_\infty \leq \varepsilon$$

ce qui montre la continuité de Φ' au point X. ■

2

THÉORÈME DES FONCTIONS IMPLICITES

2.1 Théorème d'existence des fonctions implicites

Définition 2.1.1. *Soient E, F et G trois espaces vectoriels normés et f une application de $E \times F$ dans G. On appelle fonction implicite définie par la relation $f(x,y) = c$, $c \in G$, toute application g d'une partie A de E dans une partie B de F telle que, restreinte à $A \times B$, la relation $f(x,y) = c$ soit équivalente à la relation $y = g(x)$.*

Rappelons le théorème du point fixe :

Théorème 2.1.2. *Soit E un espace métrique complet et Φ une application strictement contractante de E dans lui-même, c'est à dire qu'il existe $k \in]0,1[$ tel que*

$$\forall x_1, x_2 \in E \ , \ d\big(\Phi(x_1), \Phi(x_2)\big) \leq kd(x_1, x_2)$$

Alors Φ admet un point fixe unique, c'est à dire qu'il existe $x \in E$ tel que $\Phi(x) = x$.
De plus la suite $(x_n)_{n\in\mathbb{N}}$ telle que $x_{n+1} = \Phi(x_n)$ converge vers le point fixe.

On rappelle également une version paramétrée du thèorème du point fixe pour les contractions strictes, dont nous aurons besoin pour démontrer le théorème d'existence des fonctions implicites :

Théorème 2.1.3. *Soit E un espace métrique complet, T un espace topologique et Φ une application de $T \times E$ dans E telle que :*

1) $t \to \Phi(t,x)$ *est continue pour tout* $x \in E$.

2) *Il existe $k \in]0,1[$ tel que $\forall t \in T, \forall x_1, x_2 \in E \ , \ d\big(\Phi(t,x_1), \Phi(t,x_2)\big) \leq kd(x_1,x_2)$.*

Alors l'application $x \to \Phi(t,x)$ admet un point fixe unique $\Psi(t)$, c'est à dire tel que $\Psi(t) = \Phi(t, \Psi(t))$ et Ψ est continue de T dans E.

Définition 2.1.4. *On dira que les applications $x \to \Phi(t,x)$ vérifiant l'hypothèse 2) du théorème 2.1.2 sont uniformément contractantes de E dans lui-même.*

On peut maintenant énoncer le théorème des fonctions implicites :

Théorème 2.1.5. *(Théorème des fonctions implicites)*
Soient E un espace topologique, F et G deux espaces de banach et f une application d'un ouvert Ω de $E \times F$, à valeurs dans G.

Soient $a, b \in E \times F$ tels que $f(a,b) = 0$. On suppose de plus que f est continue sur Ω, que la dérivée partielle f'_y existe et est continue sur Ω et que $f'_y(a,b)$ est un isomorphisme de F sur G.

Alors, il existe un voisinage A de a dans E , un voisinage B de b dans F et une fonction continue g de A dans B, tels que la restriction à $A \times B$ de la relation $f(x,y) = 0$ soit équivalente à la relation : $y = g(x)$.

Démonstration : En composant f avec $\left(f'_y(a,b)\right)^{-1}$, on se ramène au cas où $F = G$ et $f'_y(a,b) = Id$. Par translation, on peut supposer de plus que $b = 0$.
Posons $\Phi(x,y) = y - f(x,y)$. On a alors :

$$\begin{cases} \Phi(a,0) = 0 \\ \Phi'_y(a,0) = 0 \end{cases}$$

Pour chaque $x \in E$, les valeurs de y satisfaisant la relation $f(x,y) = 0$ sont les points fixes de l'application $y \to \Phi(x,y)$.
Or on a :

$$\|\Phi(x,y_1) - \Phi(x,y_2)\| \leq k \|y_1 - y_2\|$$

avec

$$k \leq \sup_{y \in]y_1, y_2[} \left\|\Phi'_y(x,y)\right\|$$

Par continuité de Φ'_y, il existe un voisinage A_1 de a dans E et une boule fermée $\overline{B}$ de centre 0 et de rayon r dans F tels que :

$$\begin{cases} x \in A_1 \\ y \in \overline{B} \end{cases} \Rightarrow \left\|\Phi'_y(x,y)\right\| \leq \frac{1}{2}$$

D'où

$$\begin{cases} x \in A_1 \\ y_1, y_2 \in \overline{B} \end{cases} \Rightarrow \|\Phi(x,y_1) - \Phi(x,y_2\| \leq \frac{1}{2} \|y_1 - y_2\| \leq \frac{r}{2}$$

Par continuité de Φ, il existe un voisinage A_2 de a dans E tel que

$$x \in A_2 \Rightarrow \|\Phi(x,0)\| \leq \frac{r}{2}$$

Posons alors $A = A_1 \cap A_2$. On a donc :

$$\begin{cases} x \in A \\ y \in \overline{B} \end{cases} \Rightarrow \|\Phi(x,y)\| \leq \|\Phi(x,0)\| + \|[\Phi(x,y) - \Phi(x,0)]\| \leq \frac{r}{2} + \frac{r}{2} = r$$

Pour tout $x \in A$, les applications $y \to \Phi(x,y)$ sont donc des applications uniformément contractantes de l'espace métrique complet $\overline{B}$ dans lui-même. On peut donc appliquer le théorème 2.1.2 : pour chaque $x \in A$, il existe un unique point fixe $g(x)$ et la fonction g ainsi définie de A dans B est continue. ■

Remarque 2.1.6. *Dans le théorème 2.1.5, on peut supposer que les voisinages A et B sont ouverts.*

En effet, il suffit de remarquer qu'on peut toujours remplacer $\overline{B}$ par un voisinage $B_1 \subset \overline{B}$ sous réserve de remplacer A par $g^{-1}(B_1)$. Si B_1 est ouvert, $g^{-1}(B_1)$ le sera également. ■

2.2 Théorèmes de différentiabilité des fonctions implicites

Théorème 2.2.1. *(Différentiabilité des fonctions implicites)*
Soient E, F et G trois espaces de banach et f une application continue d'un ouvert Ω de $E \times F$, à valeurs dans G.
On suppose qu'il existe un ouvert A de E, un ouvert B de F et une fonction continue g de A dans B, tels que la restriction à $A \times B$ de la relation $f(x, y) = 0$ soit équivalente à la relation : $y = g(x)$.
On suppose de plus que f est différentiable en $(a, b) \in A \times B$ et que $f'_y(a, b)$ est un isomorphisme de F sur G.
Alors, g est différentiable en a et

$$g'(a) = -(f'_y(a,b))^{-1} \circ f'_x(a,b)$$

Démonstration : Soit $h \in E$. Posons $g(a+h) = b+k$. On a donc $f(a+h, b+k) = 0$, d'où

$$f'_x(a,b)h + f'_y(a,b)k + \epsilon(h,k)(\|h\| + \|k\|) = 0$$

avec $\epsilon(h, k) \to 0$ quand $h, k \to 0$.
On en déduit :

$$k = -\big(f'_y(a,b)\big)^{-1} f'_x(a,b)h - \big(f'_y(a,b)\big)^{-1} \epsilon(h,k)(\|h\| + \|k\|)$$

Comme $k \to 0$ lorsque $h \to 0$, il existe ϵ' tel que

$$\|h\| \leq \epsilon' \Rightarrow \left\| \big(f'_y(a,b)\big)^{-1} \right\| \|\epsilon(h,k)\| \leq \frac{1}{2}$$

D'où

$$\|k\| \leq \left\|(f'_y(a,b))^{-1}\right\| \|f'_x(a,b)\| \, \|h\| + \frac{1}{2}(\|h\| + \|k\|)$$

En d'autres termes, ceci revient à une majoration du type

$$\|k\| \leq \lambda \|h\|$$

On a alors :

$$k = -\big(f'_y(a,b)\big)^{-1} f'_x(a,b)h + \epsilon_1(h) \|h\|$$

avec $\|\epsilon_1(h)\| \to 0$ quand $h \to 0$, ce qui prouve le théorème 2.2.1. ■

2.3 Continuité des différentielles des fonctions implicites

Avant d'étudier la continuité des différentielles des fonctions implicites, nous avons besoin de résultats sur les algèbres de Banach :

Définition 2.3.1. *1) On appelle algèbre normée un espace vectoriel normé A, muni d'une loi de multiplication : $x, y \to xy$ définissant sur A une structure d'algèbre et telle que :*

$$\|xy\| \leq \|x\| \, \|y\|$$

2) Si l'algèbre admet un élément unité pour la loi multiplicative, on dira que c'est une algèbre unitaire.

3) Si l'espace vectoriel A est complet, on dira que l'algèbre associée est une algèbre de Banach.

Proposition 2.3.2. *Si A est une algèbre de Banach ayant un élément unité I et si $u \in A$ est tel que $\|u\| < 1$, alors $(I - u)$ est inversible et $(I - u)^{-1}$ tend vers I lorsque $u \to 0$.*

Démonstration : Si $u \in A$ est tel que $\|u\| < 1$, alors la série $\sum u^n$ est normalement convergente dans A car $\|u^n\| \leq \|u\|^n$. Elle est donc convergente puisque par hypothèse, A est en particulier un espace vectoriel normé complet. La somme s de cette série définit un élément de A qui vérifie :

$$(I - u)s = s(I - u) = I$$

$I - u$ est donc bien inversible dans A, d'inverse s.
De plus,

$$\left\|(I - u)^{-1} - I\right\| \leq \sum_{n=1}^{+\infty} \|u\|^n = (1 - \|u\|)^{-1} \|u\|$$

ce qui prouve que $(I - u)^{-1}$ tend vers I lorsque $u \to 0$. ■

Corollaire 2.3.3. *L'ensemble des éléments inversible d'une algèbre unitaire de Banach A est un ouvert sur lequel l'application $u \to u^{-1}$ est continue.*

Démonstration : Soit u_0 un élément inversible de A. Si $h \in A$ vérifie $\|h\| < \left\|u_0^{-1}\right\|^{-1}$, alors $\left\|hu_0^{-1}\right\| < 1$ et hu_0^{-1} est inversible dans A. Comme $u_0 + h = u_0(I + u_0^{-1}h)$, $u_0 + h$ est aussi inversible et on peut alors écrire :

$$(u_0 + h)^{-1} = (I + u_0^{-1}h)^{-1}u_0^{-1}$$

et si $h \to 0$, alors $(u_0 + h)^{-1}$ tend vers u_0^{-1}. ■

Proposition 2.3.4. *Si A est une algèbre de Banach unitaire, l'application $u \to u^{-1}$ est différentiable sur l'ensemble des éléments inversibles de A et sa dérivée en u_0 est l'application linéaire de A dans A :*

$$h \to -u_0^{-1}hu_0^{-1}$$

Démonstration : L'application $u \to u^{-1}$ est la fonction implicite définie par la relation $uv = I$, quand u et v appartiennent à l'ensemble des éléments inversibles de A.

Posons $f(u,v) = uv$ et appliquons le théorème de différentiabilité des fonctions implicites (théorème 2.2.1) :

$$\begin{cases} f'_u(u_0,v_0)h = hv_0 = hu_0^{-1} \\ f'_v(u_0,v_0)k = u_0 k \\ \left(f'_v(u_0,v_0)\right)^{-1}k = u_0^{-1}k \end{cases}$$

D'où

$$\left(f'_v(u_0,v_0)\right)^{-1} \circ f'_u(u_0,v_0)h = u_0^{-1}hu_0^{-1}$$

■

Remarque 2.3.5. *On appliquera souvent les résultats précédents dans le cas où A est l'algèbre$\mathscr{L}(E)$, où E est un espace de Banach ; la loi multiplicative dans ce cas est la loi de composition des applications linéaires, que l'on notera comme d'habitude $u, v \to u \circ v$.*

Dans le cas où $A = \mathscr{L}(E,F)$, on obtient :

Proposition 2.3.6. *Soient E et F deux espaces de Banach isomorphes et soit $Isom(E,F)$ et $Isom(F,E)$ l'ensembles des isomorphismes de E dans F et de F dans E. L'application $v \to v^{-1}$ est une application différentiable de $Isom(E,F)$ dans $Isom(F,E)$ dont la dérivée en v_0 est l'application linéaire de $Isom(E,F)$ dans $Isom(F,E)$*

$$h \to -v_0^{-1} \circ h \circ v_0^{-1}$$

Démonstration : Soient v_0 et $v \in Isom(E,F)$. Posons $u = v_0^{-1} \circ v$. Alors $u \in Isom(E)$. L'application $v \to v_0^{-1} \circ v$ est un isomorphisme de $Isom(E,F)$ dans $Isom(E)$. Par suite, grâce au corollaire 2.3.3, $Isom(E,F)$ est un ouvert de$\mathscr{L}(E,F)$ et de même, $Isom(F,E)$ est un ouvert de$\mathscr{L}(F,E)$.
L'application $v \to v^{-1}$, de $Isom(E,F)$ dans $Isom(F,E)$, est la fonction implicite définie par la relation $u \circ v = I$. On obtient sa dérivée comme dans la proposition 2.3.6. ■

On peut maintenant étudier la continuité des différentielles des fonctions implicites. On a deux versions possibles : dans la première, les ouverts A et B sont donnés et $f'_y(x,y)$ est un isomorphisme de F sur G pour $x, y \in A \times B$ et dans la seconde, on les construit à partir d'un point a, b où $f'_y(a,b)$ est un isomorphisme à l'aide du théorème des fonctions implicites 2.1.5.

Théorème 2.3.7. *(Continuité des différentielles de fonctions implicites 1)*
Soient E, F et G trois espaces de Banach et f une fonction continûment différentiable d'un ouvert Ω de $E \times F$ à valeurs dans G.

On suppose que la relation $f(x,y) = 0$ est équivalente sur le produit $A \times B$ de deux ouverts de E et F à la relation $y = g(x)$, où g est continue sur A et que pour tout $x, y \in A \times B$, $f'_y(x,y)$ est un isomorphisme de F sur G.
Alors g est continûment différentiable sur A.

Démonstration : D'après le théorème 2.2.1, on a, pour tout $x, y \in A \times B$:

$$g'(x) = -\left(f'_y(x,y)\right)^{-1} \circ f'_x(x,y)$$

La proposition 2.3.6 permet d'obtenir le théorème sans difficulté. ■

Théorème 2.3.8. *(Continuité des différentielles de fonctions implicites 2)*

Soient E, F et G trois espaces de Banach et f une fonction continûment différentiable d'un ouvert Ω de $E \times F$ à valeurs dans G.

On suppose que $f(a,b) = 0$ et $f'_y(a,b)$ est un isomorphisme de F sur G.

Alors il existe un voisinage ouvert A de a dans E, un voisinage ouvert B de b dans F tels que la restriction à $A \times B$ de la relation $f(x,y) = 0$ soit équivalente à une relation $y = g(x)$, où g est continûment différentiable sur A.

Démonstration : En appliquant le corollaire 2.3.3 à l'algèbre de Banach$\mathscr{L}(E,F)$, on voit que $Isom(E,F)$ est un ouvert de$\mathscr{L}(E,F)$ et par suite, comme par hypothèse, $f'_y(a,b)$ est un isomorphisme de F sur G, quitte à remplacer Ω par un ouvert plus petit, on peut supposer que $f'_y(x,y)$ est un isomorphisme de F sur G pour tout $x, y \in \Omega$.
D'après le théorème d'existence des fonctions implicites 2.1.5, il existe des ouverts A et B de E et F respectivement tels que $A \times B \in \Omega$ et tels que sur $A \times B$, la relation $f(x,y) = 0$ soit équivalente à la relation $y = g(x)$, où g est est continue sur A.
D'après le théorème de différentiation des fonctions implicites 2.2.1, g est différentiable sur A et

$$g'(x) = -\big(f'_y(x,y)\big)^{-1} \circ f'_x(x,y)$$

ce qui montre que g' est continue sur A. ■

Définition 2.3.9. *Soient E et F deux espaces de Banach. On appelle difféomorphisme, respectivement difféomorphisme de classe C^1, d'une partie ouverte A de E dans une partie ouverte B de F toute bijection de A sur B, différentiable de A dans B, respectivement continûment différentiable de A dans B,ainsi que son inverse.*

Théorème 2.3.10. *Soient E et F deux espaces de Banach et f une application de classe C^1 sur un ouvert de E, à valeurs dans F. Si $f'(a)$ est un isomorphime de E sur F, il existe un voisinage ouvert A de a dans E et un voisinage ouvert B de $b = f(a)$ dans F tels que la restriction de f à A soit un difféomorphisme de classe C^1 de A sur B.*

Démonstration : On applique le théorème 2.3.7 à la relation $y = f(x)$: il existe un voisinage ouvert A de a et un voisinage ouvert B de $b = f(a)$ tels que sur $A \times B$, la relation $y = f(x)$ soit équivalent à une relation $x = g(y)$ où g est continûment différentiable de B sur A
Cela signifie que la restriction de f à A est un difféomorphisme de classe C^1. ■

2.4 Notations différentielles

Dans la pratique, on utilisera des notations très suggestives, qui ont le très grand mérite de résumer les formules démontrées dans les paragraphes précédents.

Si f est une fonction différentiable de n variables $x_1, x_2, ..., x_n$, la différentielle de f sera notée df et on définit alors la relation suivante :

$$df(x_1, x_2, ..., x_n) = f'(x_1, x_2, ..., x_n) \begin{pmatrix} dx_1 \\ dx_2 \\ . \\ . \\ dx_n \end{pmatrix}$$

$$f'_1(x_1, x_2, ..., x_n)dx_1 + f'_2(x_1, x_2, ..., x_n)dx_2 + ... + f'_n(x_1, x_2, ..., x_n)dx_n$$

En particuliers, si $y - f(x) = 0$, on aura par convention :

$$dy = f'(x)dx$$

soit encore si $x \in \mathbb{R}$:

$$f'(x) = \frac{dy}{dx}$$

Cette convention de notation permet de retrouver des règles de différentiation :

Application 2.4.1. *Règle de différentiation des fonctions composées.*

Si $z = g(y)$ et $y = f(x)$, où f et g sont des fonctions différentiables, la dérivée de $g \circ f$ s'obtient en différentiant successivement les deux relations :

$$dz = g'(y)dy \ , \ dy = f'(x)dx$$

et en reportant la valeur de dy fournie par la deuxième formule dans la première formule :

$$dz = g'(y)f'(x)dx = g'\big(f(x)\big)f'(x)dx$$

et on retrouve ainsi la formule donnée au théorème 1.4.2.

Application 2.4.2. *Règle de différentiation des fonctions implicites.*

Soient E, F et G trois espaces de Banach et f une application de classe C^1 d'un ouvert Ω de $E \times F$, à valeurs dans G.
En différentiant la relation $f(x, y) = 0$, on trouve :

$$f'_x(x, y)dx + f'_y(x, y)dy = 0$$

Si pour $x = a$ et $y = b$, l'application linéaire $f'_y(a, b)$ est un isomorphisme de F sur G, alors la relation $f(x, y) = 0$ définit implicitement une relation $y = g(x)$ dans un voisinage de a, b et la différentielle de g se calcule à partir de la différentielle de la formule précédente :

$$dy = -\big(f'_y(x, y)\big)^{-1} f'_x(x, y)dx$$

On retrouve ainsi la formule démontrée au théorème 2.1.5.

Exemple 2.4.3. *Soient $\tilde{X}, \tilde{Y} \in \mathscr{L}(\mathbb{R}^n)$ tels que $\tilde{X}^2 = \tilde{Y}$ et que $\tilde{X}$ n'ait pas deux valeurs propres opposées. Alors, il existe des voisinages V de $\tilde{X}$ et W de $\tilde{Y}$ tels que on ait l'équivalence*

$$\left\{\begin{array}{l} Y = X^2 \\ X \in U \\ Y \in W \end{array}\right\} \iff X = \varphi(Y)$$

où φ est une fonction de classe C^1.

La fonction $X, Y \to X^2 - Y$ est de classe C^1. En différentiant la relation $X^2 - Y = 0$, on obtient :

$$X \circ dX + dX \circ X - dY = 0$$

On prend $X = \tilde{X}$ et on étudie l'opérateur :

$$dX \to \tilde{X} \circ dX + dX \circ \tilde{X}$$

C'est un opérateur linéaire de $\mathscr{L}(\mathbb{R}^n)$ dans lui-même. Pour qu'il soit un isomorphisme, il suffit qu'il soit injectif ou surjectif. Or si $\tilde{X}$ n'a pas de valeurs propres opposées, on peut montrer qu'il n'existe pas d'opérateur non nul $C \in \mathscr{L}(\mathbb{C}^n)$ tel que $\tilde{X}C = -C\tilde{X}$. A fortiori, il n'existe pas non plus de matrice réelle non nulle vérifiant cette égalité. Par suite, l'opérateur $dX \to \tilde{X} \circ dX + dX \circ \tilde{X}$ est injectif donc c'est un isomorphisme.
La conclusion est une conséquence du théorème des fonctions implicites.
La dérivée de φ est donnée par la formule :

$$X \circ dX + dX \circ X - dY = 0$$

dont la résolution explicite n'est pas classique.

3

DIFFÉRENTIABILITÉ À L'ORDRE SUPÉRIEUR

3.1 Dérivées secondes.

Définition 3.1.1. *1) Soient E et F deux espaces de Banach et f une application continûment différentiable d'un ouvert Ω de E, à valeurs dans F. Si la dérivée f' de f est différentiable en $a \in \Omega$, sa dérivée est appelée dérivée seconde de f en a et est notée $f''(a)$.*

2) Si pour tout $a \in \Omega$, $f''(a)$ existe, on dira que f est deux fois différentiable sur Ω.

3) Si f'' existe et est continue sur Ω, on dira que f est deux fois continûment différentiable sur Ω ou de classe C^2.

On a vu dans le chapitre 1 que pour tout $x \in \Omega$, $f'(x) \in \mathscr{L}(E, F)$. Donc f' est une application de $\Omega \subset E$ à valeurs dans $\mathscr{L}(E, F)$. Par suite, la dérivée seconde de f est une application de $\Omega \subset E$, à valeurs dans $\mathscr{L}(E, \mathscr{L}(E, F))$.
C'est à dire :

$$\forall x \in \Omega \ , \ \forall h \in E \ , \ f''(x)h \in \mathscr{L}(E, F)$$

et encore :

$$\forall x \in \Omega \ , \ \forall h \in E \ , \ \forall k \in E \ , \ (f''(x)h)k \in F$$

Pour comprendre la nature de ces espaces, on a besoin d'un lemme :

Lemme 3.1.2. *Il existe une isométrie linéaire surjective de $\mathscr{L}\big(E; \mathscr{L}(E, F)\big)$ dans $\mathscr{L}_2(E, F)$, espace des applications bilinéaires continues de $E \times E$ dans F.*

Démonstration : Soit $u \in \mathscr{L}\big(E; \mathscr{L}(E, F)\big)$. On pose

$$\forall h, k \in E \ , \ v(h, k) = (uh)k$$

Alors l'application v appartient à $\mathscr{L}_2(E, F)$ et l'on peut écrire :

$$\|v\| = \sup_{\|h\| \leq 1, \|k\| \leq 1} v(h, k) = \sup_{\|h\| \leq 1, \|k\| \leq 1} (uh)k =$$

$$\sup_{\|h\|\leq 1}\sup_{\|k\|\leq 1}(uh)k = \sup_{\|h\|\leq 1}\|uh\| = \|u\|$$

Réciproquement, si $v \in \mathscr{L}_2(E,F)$, l'application $uh : k \to (uh)k$ est linéaire et dépend linéairement de h. Donc $u \in \mathscr{L}(E, \mathscr{L}(E,F))$. Cette application est l'inverse de celle décrite ci-dessus. ■

Notations 3.1.3. *On identifiera toujours les espaces* $\mathscr{L}(E, \mathscr{L}(E,F))$ *et* $\mathscr{L}_2(E,F)$ *et on le notera* $\mathscr{L}_2(E,F)$*. La dérivée seconde d'une application de E dans F apparaît ainsi comme un application de E dans* $\mathscr{L}_2(E,F)$*.*

Rappel 3.1.4. *On dira qu'un élément* $u \in \mathscr{L}_2(E,F)$ *est symétrique si*

$$\forall h,k \in E \ , \ u(h,k) = u(k,h)$$

Le résultat suivant exprime la symétrie de la dérivée seconde :

Théorème 3.1.5. *(Symétrie de la dérivée seconde)*
Soient E et F deux espaces de Banach et f une application d'un ouvert Ω de E, à valeurs dans F. Si f est deux fois différentiable en $a \in \Omega$, $f''(a)$ est une application bilinéaire symétrique de E dans F.

Démonstration : Puisque Ω est ouvert, pour tout $h,k \in E$, il existe $r > 0$ tel que

$$\forall |\lambda|, |\mu| \leq r \Rightarrow a + \lambda h + \mu k \in \Omega$$

Posons, pour tout $t \in [-r, +r]$,

$$A = f\big(a + t(h+k)\big) - f(a+th) - f(a+tk) + f(a)$$

Alors

$$A = g(h) - g(0)$$

où

$$g(\xi) = f\big(a + t(\xi + k)\big) - f(a + t\xi)$$

En différentiant g, on trouve :

$$g'(\xi) = tf'\big(a + t(\xi+k)\big) - tf'(a+t\xi)$$

On a alors, par le théorème des Accroissements Finis :

$$\big\|A - t^2 f''(a)(h,k)\big\| = \big\|g(h) - g(0) - t^2 f''(a)(h,k)\big\| \leq$$

$$\|h\| \sup_{\xi\in]0,h[} \big\|g'(\xi) - t^2 f''(a)(.,k)\big\| =$$

$$\|h\| \sup_{\xi\in]0,h[} \big\|tf'\big(a + t(\xi+k)\big) - tf'(a+t\xi) - t^2 f''(a)(.,k)\big\|$$

Par définition de f'' on a :

$$f'\big(a + t(\xi+k)\big) = f'(a) + tf''(a)(.,\xi+k) + t\epsilon_1(t(\xi+k))\,\|\xi+k\|$$

$$f'(a+t\xi) = f'(a) + tf''(a)(.,\xi) + t\epsilon_2(t\xi)\,\|\xi\|$$

où $\epsilon_1(t(\xi+k))$ et $\epsilon_2(t\xi) \to 0$ quand $t \to 0$. D'où :

$$\left\|A - t^2 f''(a)(h,k)\right\| \leq$$

$$\|h\| \sup_{\xi\in]0,h[} \left\|t^2\epsilon_1(t(\xi+k))\right\| \|\xi+k\| + \left\|t^2\epsilon_2(t\xi)\right\| \|\xi\| \leq t^2\epsilon(t)$$

où $\epsilon(t) \to 0$ quand $t \to 0$.
On déduit de ces inégalités que

$$\frac{A}{t^2} \to f''(a)(h,k)$$

lorsque $t \to 0$. Comme A est symétrique par définition par rapport à h et k, il en est de même de $f''(a)(h,k)$. ∎

Remarque 3.1.6. *Ce résultat est une extension du principe de symétrie de Schwarz pour les fonctions de deux variable réelles.*

Exemple 3.1.7. *Soit $E = \mathbb{R}^n$ et $F = \mathbb{R}$. Si $f''(a)$ existe, c'est une forme bilinéaire symétrique. Les coefficients de la matrice de $f''(a)$ dans la base canonique $\{e_1, e_2, ..., e_n\}$ de $\mathbb{R}^n$ sont par définition $f''(a)(e_i, e_j)$. On notera cet élément $f''_{ij}(a)$ et par suite, si h et k sont deux vecteurs de E, de coordonnées $h = \begin{pmatrix} h_1 \\ h_2 \\ . \\ . \\ h_n \end{pmatrix}$ et $k = \begin{pmatrix} k_1 \\ k_2 \\ . \\ . \\ k_n \end{pmatrix}$, $f''(a)(h,k)$ s'écrit :*

$$f''(a)(h,k) = \sum_{i=1}^{n} f''_{ij}(a) h_i k_i$$

Cette représentation reste valable si E est un espace vectoriel de dimension n muni d'une base.

3.2 Dérivées sucessives

En itérant le procédé de différentiation, on a :

Définition 3.2.1. *1) Soient E et F deux espaces de Banach et f une application d'un ouvert Ω de E, à valeurs dans F. Si les dérivées $f', f'', ..., f^{(n)} = (f^{(n-1)})'$ de f existent en $a \in \Omega$, on dira que f est n fois différentiable en a.*

2) Si pour tout $a \in \Omega$, $f^{(n)}(a)$ existe, on dira que f est n fois différentiable sur Ω.

3) Si $f^{(n)}$ existe et est continue sur Ω, on dira que f est deux fois continûment différentiable sur Ω, ou bien de classe C^n.

Comme dans le paragraphe précédent, déterminons les espaces d'arrivée des dérivées successives d'une fonction :

Par commodité, nous noterons

$$\mathscr{L}_0(E,F) = F$$
$$\mathscr{L}_1(E,F) = \mathscr{L}(E,F)$$
$$\mathscr{L}_2(E,F) = \mathscr{L}\big(E;\mathscr{L}(E,F)\big)$$
$$\dots$$
$$\mathscr{L}_n(E,F) = \mathscr{L}\big(E;\mathscr{L}_{n-1}(E,F)\big)$$

Comme dans le cas $n = 2$ traité au paragraphe 3.1, l'espace $\mathscr{L}_n(E,F)$ apparaît naturellement comme l'ensemble des applications n-linéaires de E dans F. On identifiera toujours ces deux espaces.

On a vu dans les paragraphes précédents que pour tout $x \in \Omega$,

$$f'(x) \in \mathscr{L}_1(E,F)$$
$$f''(x) \in \mathscr{L}_2(E,F)$$

D'où en itérant,

$$f^{(n)}(x) \in \mathscr{L}_n(E,F)$$

Exemple 3.2.2. *Si $E = \mathbb{R}$, alors il est facile de voir que $\mathscr{L}_n(E,F)$ est isomorphe à F^n et*

$$f^{(n)}(a) = \lim_{h\to 0} \frac{f^{n-1}(a+h) - f^{n-1}(a)}{h}$$

Comme pour la dérivée seconde, on a un théorème de symétrie de la dérivée n-ième d'une application. Ce théorème se démontre simplement par récurrence à partir du théorème 3.1.5 mais nous admettrons ici sa démonstration pour ne pas alourdir ce texte :

Théorème 3.2.3. *(Symétrie des dérivées successives)*
Soient E et F deux espaces de Banach et f une application d'un ouvert Ω de E, à valeurs dans F. Si f est n fois différentiable en $a \in \Omega$, $f^{(n)}(a)$ est une application n linéaire symétrique de E dans F.

On rappelle qu'une application n linéaire est symétrique si sa valeur ne change pas lorsque que l'on échange deux variables.

Le résultat suivant est une extension du théorème 1.5.12. Il se démontre par récurrence à partir de ce théorème mais comme précédemment, nous n'en donnerons pas la démonstration pour ne pas alourdir.

Théorème 3.2.4. *Soit f une application d'un ouvert de $\mathbb{R}^n$, à valeurs dans un espace de Banach F et $m \geq 1$. Alors f est de classe C^m si et seulement si f a des dérivées partielles d'ordre inférieur ou égal à m continues sur Ω.*

Ce théorème s'étend sans problème lorsque E est un espace de dimension n, muni d'une base ou plus généralement lorsque E est un produit de n espaces de Banach, $E = E_1 \times E_2 \times \dots \times E_n$.

Les deux résultats suivants sur les différentiations successives des fonctions composées et les différentiations successives des fonctions implicites, se démontrent de même par récurrence à partir des théorèmes correspondants au cas $n = 1$ (1.4.2 et 2.2.1) ; nous ommettrons leur démonstration.

Théorème 3.2.5. *Soient E, F et G trois espaces de Banach, f une application d'un ouvert Ω de E, à valeurs dans F et g une application d'un ouvert Ω' de F contenant $f(\Omega)$, à valeurs dans G.*

1) Si f est m fois différentiable en a et g est m fois différentiable en $f(a)$, alors $g \circ f$ est m fois différentiable en a.

2) Si f et g sont m fois continûment différentiable, alors $g \circ f$ est m fois continûment différentiable.

Théorème 3.2.6. *Soient E, F et G trois espaces de Banach et f une fonction de classe C^m d'un ouvert Ω de $E \times F$ à valeurs dans G.*
On suppose que la relation $f(x, y) = 0$ est équivalente sur le produit $A \times B$ de deux ouverts de E et F à la relation $y = g(x)$, où g est continue sur A et que pour tout $x, y \in A \times B$, $f'_y(x, y)$ est un isomorphisme de F sur G.
Alors g est de classe C^m sur A.

En considérant l'application $u \to u^{-1}$ comme une fonction implicite associée à la relation $uv = I$ et en appliquant le théorème 3.2.6, on obtient alors :

Corollaire 3.2.7. *Soient E et F deux espaces de Banach isomorphes. L'application $u \to u^{-1}$ de $Isom(E, F)$ dans $Isom(F, E)$ est de classe C^m pour tout m, c'est à dire de classe C^∞.*

Définition 3.2.8. *Soient E et F deux espaces de Banach, A un ouvert de E, B un ouvert de F. On appelle difféomorphisme de classe C^m de A sur B, toute application bijective de A sur B, de classe C^m ainsi que son inverse.*

Une application directe du corollaire 3.2.7 donne alors :

Corollaire 3.2.9. *Soient E et F deux espaces de Banach, A un ouvert de E, B un ouvert de F. Si f est un difféomorphisme de A sur B et si f est de classe C^m, alors f est un difféomorphisme de classe C^m de A sur B.*

3.3 Formules de Taylor

Définition 3.3.1. *Soient E et F deux espaces normés.*

1) On appelle monôme de degré n, défini sur E, à valeurs dans F, toute application $x \to \Phi(x) = a(x, x, .., x)$ où a est une application n-linéaire symétrique de E dans F.

2) On appelle polynôme de degré n, défini sur E, à valeurs dans F, toute application $x \to P(x)$ avec

$$P(x) = \sum_{k=0}^{n} a_k(x, x, ..., x)$$

où $a_k \in \mathscr{L}_k(E, F)$ et est symétrique.

Théorème 3.3.2. *Soit $\Phi(x) = a(x, x, .., x)$ un monôme de degré n. On a :*

$$\Phi'(x) = na(x, .., x, x, .)$$

$$\Phi''(x) = n(n-1)a(x, .., x, ., .)$$

$$...$$

$$\Phi^{(n)}(x) = n!a(., ..., .)$$

et pour $N > n$,

$$\Phi^{(N)}(x) = 0$$

DÉMONSTRATION : Si $a \in \mathscr{L}_n(E, F)$ alors il est clair que l'application

$$y \to a(x_1, x_2, ..., x_{n-1}, y)$$

est dans $\mathscr{L}(E, F)$. Donc $a(x, ..., x, .)$ est un monôme de degré $n-1$, à valeurs dans $\mathscr{L}(E, F)$. De la même façon, l'application de deux variabes

$$y_1, y_2 \to a(x_1, x_2, ..., x_{n-2}, y_1, y_2)$$

est dans $\mathscr{L}_2(E, F)$. Donc $a(x, .., x, ., .)$ est un monôme de degré $n-2$, à valeurs dans $\mathscr{L}_2(E, F)$. Généralement, l'application de p variabes

$$y_1, ..., y_p \to a(x_1, ., ..., x_{n-p}, y_1, y_2, ..., y_p)$$

est dans $\mathscr{L}_p(E, F)$. Donc $a(x, .., x, ., .)$ est un monôme de degré $n-p$, à valeurs dans $\mathscr{L}_p(E, F)$. En raison de la linéarité de l'application $x_i \to a(x_1, x_2, ..., x_n)$ la i-ème dérivée partielle de a est

$$a'_i(x_1, x_2, ..., x_n)(h_i) = a(x_1, ..., x_{i-1}, h_i, x_{i+1}, ..., x_n)$$

et par suite

$$a'(x_1, x_2, ..., x_n)(h_1, h_2, .., h_n) =$$

$$\sum_{i=1}^{n} a(x_1, ..., x_{i-1}, h_i, x_{i+1}, ..., x_n)$$

Donc pour $\Phi(x) = a(x, x, ..., x)$, on obtient :

$$\Phi'(x)(h) = \sum_{i=1}^{n} a(x, ..., x, h, x, ..., x) =$$

$$na(x, x, ..., h)$$

c'est à dire

$$\Phi'(x) = na(x, ..x, x, .)$$

On obtient les formules suivantes en itérant ce procédé. ■

Théorème 3.3.3. *(Formule de taylor pour les polynômes)*
Soit P un polynôme de degré n défini dans E, à valeurs dans F. On a :

$$P(x) = P(0) + P'(0)(x) + \frac{1}{2}P''(0)(x,x) + ... + \frac{1}{k!}P^{(k)}(0)(x,x,...x) + ... + \frac{1}{n!}P^{(n)}(0)(x,x,...x)$$

Démonstration : Avec les notations de la définition 3.3.1, soit

$$P(x) = \sum_{k=0}^{n} a_k(x,x,...,x)$$

où $a_k \in \mathscr{L}_k(E,F)$ et est symétrique.
D'après le théorème 3.3.2, on a $P^{(k)}(0) = k!a_k$. ■

Corollaire 3.3.4. *Soient $P(x) = \sum_{k=0}^{n} a_k(x,x,...,x)$ et $Q(x) = \sum_{k=0}^{n} b_k(x,x,...,x)$ où $a_k, b_k \in \mathscr{L}_k(E,F)$ et sont symétriques. Si $P(x) = Q(x)$ pour tout $x \in E$, alors $a_k = b_k$ pour tout $k = 1,2,...,n$.*

Définition 3.3.5. *Deux fonctions définies sur un ouvert $\Omega \subset E$, à valeurs dans F sont dites n-tangentes en $a \in E$ si elles ont mêmes dérivées jusqu'à l'ordre n au point a.*

Corollaire 3.3.6. *1) Si f est une fonction définie sur un ouvert $\Omega \subset E$ contenant 0, à valeurs dans F, n fois différentiable en 0, il existe un polynôme n tangent à f en 0, c'est*

$$P(x) = f(0) + f'(0)(x) + \frac{1}{2}f''(0)(x,x) + ... + \frac{1}{k!}f^{(k)}(0)(x,x,...x) + ... + \frac{1}{n!}f^{(n)}(0)(x,x,...x)$$

2) Si f est une fonction définie sur un ouvert $\Omega \subset E$, à valeurs dans F, n fois différentiable en $a \in \Omega$, il existe un polynôme n-tangent à $h \to f(a+h)$ en a, c'est

$$P(h) = f(a) + f'(a)(h) + \frac{1}{2}f''(a)(h,h) + ... + \frac{1}{k!}f^{(k)}(a)(h,h,...,h) + ... + \frac{1}{n!}f^{(n)}(a)(h,h,...,h)$$

Théorème 3.3.7. *(Formule de Taylor-Lagrange)*
Soit f est une fonction définie sur un ouvert $\Omega \subset E$, à valeurs dans F, n fois différentiable en $a \in \Omega$ et soit $P(h)$ le polynôme n-tangent à $h \to f(a+h)$ en a. Alors il existe une fonction $\epsilon : E \to F$ telle que

$$f(a+h) = P(h) + \epsilon(h)\left\|h\right\|^n, \text{ avec } \epsilon(h) \to 0 \text{ quand } h \to 0$$

Démonstration : Pour simplifier, supposons $a = 0$ et procédons par récurrence :
Le théorème est vrai pour $n = 1$ par définition de la différentiabilité de f en 0.
Supposons que ce théorème est vrai jusqu'à l'ordre $n-1$ et soit f n fois différentiable en 0.
On remarque que le polynôme P, n-tangent à f en 0 vérifie :

$$f'(x) = P'(x) + \epsilon_1(x)\left\|x\right\|^{n-1}, \text{ avec } \epsilon_1(h) \to 0 \text{ quand } h \to 0$$

On peut appliquer le théorème des accroissements finis (1.5.5) sur $[0,x]$:

$$\|f(x) - P(x)\| \le \|x\| \sup_{\xi \in]0,x[} \|f'(x) - P'(x)\| \le \|x\|^n \sup_{\xi \in]0,x[} \|\epsilon_1(x)\|$$

D'où

$$f(x) = P(x) + \epsilon(x)\left\|x\right\|^n$$

avec

$$\|\epsilon(x)\| \le \sup_{\xi \in]0,x[} \|\epsilon_1(x)\|$$

qui tend vers 0 quand $x \to 0$. ■

Remarque 3.3.8. *L'existence d'un polynôme P et d'une fonction ϵ, tels que $f(a+h) = P(h) + \epsilon(h)\|h\|^n$, avec $\epsilon(h) \to 0$ quand $h \to 0$ n'entraîne pas que f est n fois différentiable en a.*

Théorème 3.3.9. *(Formule de Taylor-Young pour les fonctions à valeurs réelles)*
Soit f une fonction définie sur un ouvert $\Omega \subset E$, à valeurs réelles. Supposons que $[a, a+h] \subset \Omega$ et que f est n fois différentiable en tout point de $]a, a+h[$. Soit $P(h)$ le polynôme n-tangent à f en $a \in \Omega$. Alors il existe $c \in]a, a+h[$ tel que

$$f(a+h) = P(h) + \frac{1}{(n+1)!} f^{(n+1)}(c)(h, h, ..., h)$$

Démonstration : On suppose que ce théorème est connu si $E = \mathbb{R}$. Pour $t \in \mathbb{R}$, on pose

$$F(t) = f(a+th) \text{ et } p(t) = P(a+th)$$

Alors, pour tout $k \leq n$:

$$F^{(k)}(t) = f^{(k)}(a+th)(h, h, ..., h)$$
$$p^{(k)}(t) = P^{(k)}(a+th)(h, h, ..., h)$$

Par suite p est n-tangent à F pour $t = 0$. D'après le théorème de Taylor-Young pour les fonctions définies sur des ouverts de $\mathbb{R}$, à valeurs réelles :

$$F(1) = p(1) + \frac{1}{(n+1)!} F^{(n+1)}(\theta) \text{ , avec } \theta \in]0, 1[$$

c'est à dire la formule annoncée. ■

Corollaire 3.3.10. *Sous les hypothèses du théorème 3.3.9, on a :*

$$|f(a+h) - P(h)| \leq \frac{\|h\|^{n+1}}{(h+1)!} \sup_{x \in]a, a+h[} \left\| f^{(n+1)}(x) \right\|$$

Théorème 3.3.11. *(Formule de Taylor-Young" pour les fonctions à valeurs vectorielles)*
Soit f une fonction définie sur un ouvert $\Omega \subset E$, à valeurs dans F. Supposons que $[a, a+h] \subset \Omega$ et que f est n fois différentiable en tout point de $]a, a+h[$. Soit $P(h)$ le polynôme n-tangent à f en $a \in \Omega$. Alors il existe $c \in]a, a+h[$ tel que

$$\|f(a+h) - P(h)\| \leq \frac{\|h\|^{n+1}}{(n+1)!} \sup_{x \in]a, a+h[} \left\| f^{n+1}(x) \right\|$$

Démonstration : Soit $v \in F'$. On pose

$$f_v : x \to < f(x), v >$$

$$P_v : h \to < P(h), v >$$

Alors $P_v(h)$ est n-tangent à $f(a+h)$ en 0. D'où

$$\|f_v(a+h) - P_v(h)\| \leq \frac{\|h\|^{n+1}}{(n+1)!} \sup_{x \in]a, a+h[} \left\| f_v^{n+1}(x) \right\|$$

et par suite

$$\|f(a+h)-P(h)\| = \sup_{\|v\|\leq 1} |< f(a+h)-P(a+h), v >| =$$

$$\sup_{\|v\|\leq 1} \|f_v(a+h)-P_v(h)\| \leq \sup_{\|v\|\leq 1} \frac{\|h\|^{n+1}}{(n+1)!} \sup_{x\in]a,a+h[} \left\|f_v^{n+1}(x)\right\| =$$

$$\frac{\|h\|^{n+1}}{(n+1)!} \sup_{x\in]a,a+h[} \sup_{\|v\|\leq 1} \left\|f_v^{n+1}(x)\right\| =$$

$$\frac{\|h\|^{n+1}}{(n+1)!} \sup_{x\in]a,a+h[} \left\|f^{n+1}(x)\right\|$$

■

Remarque 3.3.12. *On peut dans ce cadre-là énoncer un corollaire analogue au corollaire 1.5.6.*

Théorème 3.3.13. *(Formule de Taylor avec reste intégral)*
Soit f une fonction de classe C^{n+1}, définie sur un ouvert $\Omega \subset E$, à valeurs dans F. Supposons que $[a, a+h] \in \Omega$ et soit $P(h)$ le polynôme n-tangent à $f(a+h)$ en 0. Alors :

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

Démonstration : Si $F = \mathbb{R}$, on suppose le théorème connu pour $E = \mathbb{R}$ et on l'applique à la fonction

$$F(t) = f(a+th)$$

Pour F quelconque, on pose pour tout $v \in F'$:

$$f_v(x) =< f(x), v > \text{ et } P_v(h) =< P(h, v >$$

Alors P_v est n-tangent à $f_v(a+h)$ en 0. On a donc en appliquant la première partie, $\forall v \in F'$:

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

ce qui entraîne la conclusion du théorème. ■

Exemple 3.3.14. *La formule de Taylor-Lagrange à l'ordre 2 pour la fonction $f(x,y) = \ln(x+e^y)$ au point $(1,0)$ est :*

$$\ln(1+h+e^k) = \ln 2 + \frac{1}{2}h + \frac{1}{2}k - \frac{1}{4}h^2 - \frac{1}{4}hk - \frac{1}{4}k^2 + (h^2+k^2)\varepsilon(h,k)$$

La fonction f est de classe C^2 sur$\mathbb{R}^+ \times \mathbb{R}$ et on a :

$$f'_x(x,y) = \frac{1}{x+e^y} \ , \ f'_y(x,y) = \frac{e^y}{x+e^y}$$

$$f''_{xx}(x,y) = -\frac{1}{(x+e^y)^2} \ , \ f''_{xy}(x,y) = -\frac{e^y}{(x+e^y)^2} \ , \ f''_{yy}(x,y) = -\frac{xe^y}{(x+e^y)^2}$$

ce qui donne bien la formule annoncée.

4

EQUATIONS DIFFÉRENTIELLES

4.1 Théorèmes d'existence et d'unicité

Définition 4.1.1.

1) Soient E et F deux espaces métriques et soit f une application d'un ouvert Ω de E, à valeurs dans F. On dit que f est lipschitzienne (ou vérifie la condition de Lipschitz) s'il existe une constante $k > 0$ telle que

$$\forall x_1, x_2 \in \Omega \ , \ d\big(f(x_1), f(x_2)\big) \leq k d(x_1, x_2)$$

2) On dira que f est localement lipschitzienne sur Ω si tout point $x_0 \in \Omega$ admet un voisinage sur lequel f est lipschitzienne.

Il est clair que toute fonction lipschitzienne sur un ouvert Ω est localement lipschitzienne. Par contre, l'inverse n'est pas vrai : dans le cas où f n'est que localement lipschitzienne, la constante de Lipschitz k peut dépendre du voisinage de chaque point et n'être pas bornée sur Ω tout entier.

Exemple 4.1.2. *Soient E et F deux espaces vectoriels normés. Toute application f continûment différentiable sur un ouvert Ω de E est localement lipschitzienne.*

En effet, pour tout $x_0 \in \Omega$, il existe une boule B de centre x_0, contenue dans Ω telle que f' soit bornée sur B. Pour tout $x_1, x_2 \in \Omega$, on aura :

$$\|f(x_1) - f(x_2)\| \leq \|x_1 - x_2\| \sup_{x \in B} \|f'(x)\|$$

De la même façon, dans une situation plus particulière que l'on considérera dans ce chapitre, on peut donner les définitions suivantes :

Définition 4.1.3.

1) Soient E un espace métrique et f une application d'un ouvert Ω de $\mathbb{R} \times E$, à valeurs dans E. On dit que f est lipschitzienne par rapport à la deuxième variable s'il existe une constante $k > 0$ telle que

$$\forall (t, x_1), (t, x_2) \in \Omega \ , \ d\big(f(t, x_1), f(t, x_2)\big) \leq k d(x_1, x_2)$$

2) On dira que f est localement lipschitzienne par rapport à la deuxième variable sur Ω si tout point $(t_0, x_0) \in \Omega$ admet un voisinage sur lequel f est lipschitzienne par rapport à la deuxième variable.

Exemple 4.1.4. *Soient E un espace vectoriel normé. Toute application f continûment différentiable sur un ouvert Ω de $\mathbb{R} \times E$ est localement lipschitzienne.*

En effet, pour tout $(t_0, x_0) \in \Omega$, il existe un voisinage V de (t_0, x_0), contenu dans Ω tel que f'_x soit bornée sur V. Pour tout $(t_1, x_1), (t_2, x_2) \in \Omega$, on aura :

$$\|f(t_1, x_1) - f(t_2, x_2)\| \leq \|x_1 - x_2\| \sup_{(t,x)\in V} \|f'_x(t, x)\|$$

Soit E un espace de Banach. On considère l'équation différentielle

$$(ED)\ X'(t) = f\big(t, X(t)\big)$$

où f est une fonction définie sur une partie $I \times B$ de $\mathbb{R} \times E$, à valeurs dans E.
On considérera aussi l'équation différentielle avec condition initiale en $(t_0, x_0) \in I \times B$:

$$(ED_0)\ X'(t) = f\big(t, X(t)\big)\ ,\ X(t_0) = x_0$$

Définition 4.1.5.

1) Une solution de (ED) est une fonction X de classe C^1 sur I, à valeurs dans B, qui vérifie l'équation :

$$\forall t \in I\ ,\ X'(t) = f\big(t, X(t)\big)$$

.

2) Une solution de (ED_0) est une fonction X de classe C^1 sur I, à valeurs dans B qui vérifie :

$$\forall t \in I\ ,\ X'(t) = f\big(t, X(t)\big) \text{ et } X(t_0) = x_0$$

Dans la pratique, I sera un intervalle de $\mathbb{R}$. Pour simplifier les notations, on supposera que $t_0 = 0$ et $I = [0, T]$.
En combinant avec les résultats qu'on obtiendrait par symétrie pour $t < 0$, on trouverait des résultats concernant les solutions sur $[-T, +T]$.

Théorème 4.1.6. *Soit E un espace de banach, $R > 0$, $B = B(0, R) = \{x \in E \mid \|x\| \leq R\}$ et f une fonction continue sur $[0, T] \times B$, à valeurs dans E.*

On suppose qu'il existe une fonction continue g définie sur $[0, T] \times [0, B]$, à valeurs positives telle que

$$\left\{ \begin{array}{l} \|x\| \leq \rho \leq R \\ t \in [0, T] \end{array} \right\} \Rightarrow \|f(t, x\| \leq g(t, \rho)$$

et que l'équation différentielle réelle

$$(ED_{\mathbb{R}})\ h'(t) = f\big(t, h(t)\big)$$

admet une solution telle que $\forall t \in [0,T]$, $h(t) \in [0,R]$.
On suppose de plus que f *est lipschitzienne par rapport à la deuxième variable sur le produit* $[0,T] \times B\big(0,h(T)\big)$, *c'est à dire qu'il existe* $k > 0$ *telle que*

$$\forall x_1, x_2 \in E \text{ tels que } \|x_1\|, \|x_2\| \le h(T), \|f(x_1) - f(x_2)\| \le \|x_1 - x_2\|$$

Alors l'équation différentielle

$$(ED)\ X'(t) = f\big(t, X(t)\big)$$

admet une solution définie sur $[0,T]$ *pour toute condition initiale* $X(0) = x_0$ *vérifiant* $\|x_0\| \le h(0)$.

Démonstration : Posons $\mathscr{F} = \{X \in \mathscr{C}([0,T];E) \mid \|X(t)\| \le h(t)\}$ et soit Φ la fonction, définie pour $X \in \mathscr{F}$ par :

$$\forall t \in [0,T]\ ,\ \Phi\big(X\big)(t) = x_0 + \int_0^t f\big(\theta, X(\theta)\big)\, d\theta$$

Vérifions que Φ applique $\mathscr{F}$ dans lui-même :

$$\big\|\Phi\big(X\big)(t) - x_0\big\| \le \int_0^t \big\|f\big(\theta, X(\theta)\big)\big\|\, d\theta \le$$

$$\int_0^t g\big(\theta, X(\theta)\big)\, d\theta = h(t) - h(0)$$

D'où

$$\big\|\Phi\big(X\big)(t)\big\| \le \|x_0\| + h(t) - h(0) \le h(t)$$

c'est à dire que $\Phi\big(X\big) \in \mathscr{F}$
En appliquant l'hypothèse f lipschitzienne, on a :

$$\big\|\Phi\big(X_1\big)(t) - \Phi\big(X_2\big)(t)\big\| = \left\|\int_0^t [f\big(\theta, X_1(\theta)\big) - f\big(\theta, X_2(\theta)\big)\right\|\, d\theta \le$$

$$\int_0^t \big\|f\big(\theta, X_1(\theta)\big) - f\big(\theta, X_2(\theta)\big)\big\|\, d\theta \le$$

$$\int_0^t k\, \|X_1(\theta) - X_2(\theta)\|\, d\theta \le kt\, \|X_1 - X_2\|_\infty$$

D'où

$$\big\|\Phi\big(X_1\big) - \Phi\big(X_2\big)\big\|_\infty \le kT\, \|X_1 - X_2\|_\infty$$

Φ serait une contraction de $\mathscr{F}$ dans lui-même si $kT < 1$, ce qui en général n'est pas réalisé. Pour pouvoir utiliser le théorème du point fixe pour les contractions (Théorème 2.1.2), nous devons donc poursuivre le raisonnement :

$$\big\|\Phi^2\big(X_1\big)(t) - \Phi^2\big(X_2\big)(t)\big\| = \left\|\int_0^t [f\big(\theta, \Phi(X_1)(\theta)\big) - f\big(\theta, \Phi(X_2)(\theta)\big)\right\|\, d\theta \le$$

$$\int_0^t \left\| f\big(\theta, \Phi(X_1)(\theta)\big) - f\big(\theta, \Phi(X_2)(\theta)\big) \right\| d\theta \leq$$

$$\int_0^t k \left\| \Phi\big(X_1\big)(\theta) - \Phi\big(X_2\big)(\theta) \right\| d\theta \leq$$

$$k^2 \int_0^t t \left\| X_1 - X_2 \right\|_\infty = \frac{k^2 t^2}{2} \left\| X_1 - X_2 \right\|_\infty$$

D'où

$$\left\| \Phi^2\big(X_1\big) - \Phi^2\big(X_2\big) \right\|_\infty \leq \frac{k^2 T^2}{2} \left\| X_1 - X_2 \right\|_\infty$$

Par une récurrence immédiate, on trouve aussi, pour tout $n \in \mathbb{N}$:

$$\left\| \Phi^n\big(X_1\big) - \Phi^n\big(X_2\big) \right\|_\infty \leq \frac{k^n T^n}{n!} \left\| X_1 - X_2 \right\|_\infty$$

On en déduit que pour n suffisemment grand, Φ^n est une contraction. Comme $\mathscr{F}$ est complet, Φ^n et donc aussi Φ a un point fixe unique d'après le théorème 2.1.2. Ce point fixe est l'unique solution dans $\mathscr{F}$ de l'équation différentielle (ED_0). ■

Remarque 4.1.7.

i) On peut vérifier que la solution X de (ED) telle que $X(0) = x_0$, dépend continûment de x_0.

ii) L'unique solution de (ED) vérifiant $X(0) = x_0$ est la limite de $\Phi^n(X_0)$ quand $n \to +\infty$, pour tout $X_0 \in \mathscr{F}$, la fonction Φ étant définie par :

$$\Phi\big(X\big)(t) = x_0 + \int_0^t f\big(\theta, X(\theta)\big)\, d\theta$$

Le théorème précédent s'applique pour obtenir le résultat qu'on utilisera en pratique :

Théorème 4.1.8. *(Cauchy-Lipschitz)*
Soit f une application continue d'un ouvert Ω de $\mathbb{R} \times E$, à valeurs dans E, localement lipschitzienne.
Alors, pour tout $(t_0, x_0) \in \Omega$, il existe $\alpha > 0$ tel que l'équation

$$(ED)\ X'(t) = f\big(t, X(t)\big)$$

ait une solution unique définie sur $[t_0 - \alpha, t_0 + \alpha]$ et vérifiant $X(t_0) = x_0$.
De plus deux solutions X_1 et X_2, définies sur un même intervalle I contenant t_0 et telles que $X_1(t_0) = X_2(t_0)$ sont égales.

Remarque 4.1.9. *Dans la pratique, ce théorème sera fréquemment appliqué pour des fonctions f lipschitziennes.*

DÉMONSTRATION : Existence : Comme f est continue, il existe $T_1 > 0$ et $R > 0$ tel que f soit bornée sur le produit $[0, T_1] \times B(x_0, R)$. On peut ausi supposer que f est lipschitzienne par rapport à x sur ce produit.

Posons

$$M = \sup\{\|f(t,x)\| \ \mid t \in [0,T_1], \|x - x_0\| \leq R\}$$

On va appliquer le théorème 4.1.6 avec $g(t,\rho) = M$ pour $\rho \leq R$. On prend $h(0) = 0$ et l'on a $h(t) = Mt$ donc $h(t) \leq R$ pour $t \leq \frac{R}{M}$. Posons $T = \inf\{T_1, \frac{R}{M}\}$. Le théorème 4.1.6 s'applique sur $[0,T]$ avec :

$$\mathscr{F} = \{X \in \mathscr{C}\big([0,T];E\big) \mid \ \|X(t) - x_0\| \leq Mt\}$$

Il existe une solution $X \in \mathscr{F}$.

Unicité : première étape. Soient $X1$ et X_2 deux solutions définies sur $[0,T]$, telles que $X_1(0) = X_2(0) = x_0$. Soient R_0 et T_1 tels que sur le produit $[0,T_1] \times B(x_0,R_0)$, f soit localement lipschitzienne par rapport à x et bornée. Soit $T = \inf\{T_1,T_2\}$. On peut appliquer le théorème 4.1.6 avec, pour $\rho \leq R_0$:

$$g(t,\rho) = M = \sup\{\|f(t,x)\| \ \mid t \in [0,T], \|x - x_0\| \leq R_0\}$$

On a

$$\|X_1(t) - x_0\| \leq \int_0^t \big\|f\big(t, X(\theta)\big)\big\| \, d\theta$$

d'où $X_1 \in \mathscr{F} = \{X \in \mathscr{C}\big([0,T];E\big) \mid \|X(t)\| \leq Mt\}$. De même, $X_2 \in \mathscr{F}$. D'après le théorème 4.1.6, $\mathscr{F}$ ne contient qu'une seule solution et par suite, on a $X_1 = X_2$ sur $[0,T]$.

Unicité : deuxième étape. Soient $X1$ et X_2 deux solutions définies sur un même intervalle I et telles que $X_1(t_0) = X_2(t_0)$. D'après l'étape précédente, l'ensemble

$$A = \{t \in I \mid X_1(t) = X_2(t)\}$$

est un ouvert de I. En raison de la continuité de X_1 et X_2, A est fermé. Comme I est connexe, A étant un ouvert fermé de I, il est soit égal à I soit égal à ϕ.
Comme $t_0 \in A$, $A = I$. ■

Définition 4.1.10. *On appelle solution maximale d'une équation différentielle, toute solution définie sur un intervalle de $\mathbb{R}$ et qui ne peut être prolongée à un intervalle plus grand.*

Théorème 4.1.11. *(Existence de solutions maximales)*

Sous les hypothèses du théorème de Cauchy-Lipschitz 4.1.8, pour $(t_0,x_0) \in \Omega$, il existe une solution maximale X telle que $X(t_0) = x_0$.

Démonstration : On applique le lemme de Zorn à l'ensemble de couples (J,X), formés d'un intervalle J et d'une solution telle que $X(t_0) = x_0$, définie sur J, avec la relation d'ordre :

$$(J_1,X_1) \leq (J_2,X_2) \iff J_1 \subset J_2 \text{ et } X_2 \text{ prolonge } X_1$$

Cette ensemble est inductif, il admet donc un élément maximal (J,X). La solution X est clairement une solution maximale de l'équation différentielle. ■

Exemple 4.1.12. *L'équation différentielle linéaire*

$$(ED) X'(t) = \frac{1}{3}\left[X(t)^2 \sin(t) + X(t)^3 \cos(t)\right] \ , \ X(0) \in [-1, +1]$$

a une solution unique vérifiant $|X(t)| \leq 2$ *pour* $t \in [0, \frac{1}{2}$.

En effet, on pose $f(t,x) = \frac{1}{3}\left[x^2 \sin(t) + x^3 \cos(t)\right]$. F est évidemment de classe C^1 sur $\mathbb{R}^2$, donc localement lipschitzienne par rapport à x.

Or

$$|x| \leq \rho \leq 2 \Rightarrow |f(t,x)| \leq \rho^2$$

L'équation différentielle

$$h'(t) = h^2(t) \ , \ h(0) = 1$$

a pour solution $h(t) = \frac{1}{1-t}$.

De plus

$$\left|\frac{1}{1-T}\right| \leq 2 \iff T \leq \frac{1}{2}$$

Donc d'après le théorème d'existence de solution des équations différentielles, il existe une solution unique de (ED) sur $[0, \dfrac{1}{2}]$ vérifiant $X(0) \in [-1, +1]$.

4.2 Equations différentielles linéaires sans deuxième membre

Définition 4.2.1. *Soit E un espace de banach. On appelle équation différentielle linéaire une équation différentielle de la forme :*

$$X'(t) = A(t)X(t) \text{ avec } \forall t \in I \ , \ A(t) \in \mathscr{L}(E)$$

Théorème 4.2.2. *(Solution d'une équation différentielle linéaire)*

Si l'application $t \to A(t)$ est continue de I dans$\mathscr{L}(E)$, l'équation différentielle linéaire

$$(ED) \ X'(t) = A(t)X(t)$$

admet une et une seule solution, définie sur I et vérifiant la condition initiale : $X(t_0) = x_0$.

Démonstration : On pose $f(t,x) = A(t)x$

Supposons d'abord que $I = [0,T]$, T fini et $t_0 = 0$. En posant

$$k = \sup\{\|A(t)\| \ | \ t \in [0,T]\}$$

on obtient :

$$\forall t \in [0,T] \ , \ \|f(t,x_1) - f(t,x_2)\| \leq k \|x_1 - x_2\|$$

On peut donc appliquer le théorème d'existence des solutions d'équations différentielles 4.1.6 avec $R = +\infty$, $g(t,\rho) = k\rho$ et $h(t) = h(0)e^{kt}$.

Si $\|x_0\| \leq h(0)$, il existe une solution et une seule dans $C([0,T]; E)$ qui vérifie de plus

$$\forall t \in [0,T] \ , \ \|X(t)\| \leq h(0)e^{kt}$$

En revenant au cas général, avec les notations de l'énoncé, cette démonstration montre que l'équation différentielle (ED) a une solution unique vérifiant $X(t_0) = x_0$ sur tout intervalle finis J inclus dans I et contenant x_0.

Ceci entraîne que toute solution maximale est définie sur I tout entier. ■

Remarque 4.2.3. *L'application qui, à une solution X d'une équation différentielle linéaire (ED), définie sur un intervalle I, associe $X(t_0)$ est linéaire et bijective.*

Cette remarque permet d'énoncer le corollaire suivant :

Corollaire 4.2.4. *1) L'ensemble des solutions d'une équation différentielle linéaire à valeurs dans un espace vectoriel E est un espace vectoriel algébriquement isomorphe à E.*

2) Si I est un intervalle fermé borné, l'application qui à x_0 associe la solution X telle que $X(t_0) = x_0$ vérifie :

$$\|X(t)\| \leq \|x_0\| \, e^{k|t-t_0|} \text{ avec } k = \sup_{t \in I} \|A(t)\|$$

et par suite cette application est continue.

Définition 4.2.5. *(Résolvante)*
On appelle résolvante de l'équation différentielle linéaire, définie sur un intervalle I de $\mathbb{R}$, à valeurs dans un espace vectoriel E

$$(ED) \; X'(t) = A(t)X(t)$$

l'application $R : I \times I \to \mathscr{L}(E)$ telle que

$$\forall t_1, t_2 \in I \; , \; X(t_2) = R(t_1, t_2)X(t_1)$$

On peut énoncer les propriétés fondamentale de la résolvante :

Théorème 4.2.6. *(Propriétés de la résolvante)*

i) $R(t_3, t_2) \circ R(t_2, t_1) = R(t_3, t_1)$

ii) $R(t_1, t_1) = Id$, *où I_E est l'identité de E*

iii) $R(t_2, t_1) = \big(R(t_1, t_2)\big)^{-1}$.

Démonstration : *i)* Pour toute solution X d l'équation différentielle (ED), on a :

$$X(t_3) = R(t_3, t_2)X(t_2) = R(t_3, t_2)R(t_2, t_1)X(t_1)$$

et aussi

$$X(t_3) = R(t_3, t_1)X(t_1)$$

d'où la première relation puisque $X(t_1)$ est arbitraire. ■

Les relations *ii)* et *iii)* decoulent directement de la définition de la résolvante R.

Théorème 4.2.7. *(Caractérisation de la résolvante)*
On considère l'équation différentielle linéaire, définie sur un intervalle I de $\mathbb{R}$, à valeurs dans E :

$$(ED) \; X'(t) = A(t)X(t)$$

et on fixe $\theta \in I$.
La résovante R de (ED) est l'unique solution de l'équation différentielle dans $\mathscr{L}(E)$:

$$R'_t(t, \theta) = A(t) \circ R(t, \theta) \; , \text{ avec } R(\theta, \theta) = I_E$$

DÉMONSTRATION : Considérons à priori l'équation différentielle :

$$R'_t(t,\theta) = A(t) \circ R(t,\theta)$$

C'est une équation différentielle linéaire à valeurs dans$\mathscr{L}(E)$. En effet, l'opérateur

$$\Phi(t) : U \to A(t) \circ U$$

est un endomorphisme de$\mathscr{L}(E)$.

De plus, par un calcul facile, on a :

$$\|\Phi(t)\| = \|A(t)\| \text{ et aussi } \|\Phi(t_1) - \Phi(t_2)\| = \|A(t_1) - A(t_2)\|$$

Par suite, par le théorème de Cauchy-Lipschitz 4.1.8, l'équation différentielle considérée admet une solution unique vérifiant $R(\theta,\theta) = I_E$

Pour $X(\theta) \in E$, posons $X(t) = R(t,\theta)X(\theta)$, alors :

$$X'(t) = R'_t(t,\theta)X(\theta) = A(t) \circ R(t,\theta)X(\theta) = A(t)X(t)$$

La solution $R(t,\theta)$ de l'équation diférentielle considérée au départ est donc bien la valeur en (t,θ) de la résolvante de l'équation

$$X'(t) = A(t)X(t)$$

■

Remarque 4.2.8. *Si l'on sait que $R(t,\theta)$ est dérivable par rapport à t, on peut obtenir le résultat précédent plus simplement.*

On dérive par rapport à t la relation

$$X(t) = R(t,\theta)X(\theta)$$

et on obtient :

$$X'(t) = R'_t(t,\theta)X(\theta)$$

Comme par ailleurs,

$$X'(t) = A(t)X(t) = A(t) \circ R(t,\theta)X(\theta)$$

et que $X(\theta)$ est arbitraire, on obtient bien

$$R'_t(t,\theta) = A(t) \circ R(t,\theta)$$

. ■

Remarque 4.2.9. *En utilisant la relation $R(t,\theta) = \big(R(\theta,t)\big)^{-1}$, on voit que R est dérivable par rapport à θ, de dérivée :*

$$R_2'(t,\theta) = \frac{\partial}{\partial\theta}\big(R(\theta,t)\big)^{-1} = \big(R_1'(\theta,t)\big)^{-1} = \big(R(\theta,t)\big)^{-1} \circ A(\theta)^{-1}$$

Remarque 4.2.10. *L'équation différentielle vérifiée par la résolvante est d'une part plus compliquée que l'équation initiale vérifiée par la fonction car elle est à valeurs dans $\mathscr{L}(E,E)$ et non pas dans E mais d'autre part, cela peut aussi se révéler être un avantage car on peut exploiter la structure d'algèbre de Banach unitaire de $\mathscr{L}(E,E)$.*

Cas particulier. *Soit (ED) $X'(t) = A(t)X(t)$. Si pour tous $t,u \in I$, $A(t)A(u) = A(u)A(t)$, alors*

$$R(t,t_0) = \exp\Big(\int_{t_0}^{t} A(s)\,ds\Big)$$

Démonstration : Posons $M(t) = \exp\Big(\int_{t_0}^{t} A(s)\,ds\Big) \in \mathscr{L}(E,E)$. L'hypothèse de commutation entraîne que $\int_{t_1}^{t_2} A(s)\,ds$ et $\int_{t_3}^{t_4} A(s)\,ds$ commutent pour tous $t_1,t_2,t_3,t_4 \in I$. Donc :

$$M(t+h) = \exp\Big(\int_{t_0}^{t} A(s)\,ds + \int_{t}^{t+h} A(s)\,ds\Big) = \exp\Big(\int_{t}^{t+h} A(s)\,ds\Big) M(t)$$

Or, on a : $\int_{t}^{t+h} A(s)\,ds = hA(t) + h\epsilon(t)$ et en utilisant le développement en série de l'exponentielle :

$$M(t+h) = \big(Id + hA(t) + h\epsilon(t)\big)M(t) = M(t) + hA(t)M(t) + h\epsilon_1(t)$$

ce qui montre que $M'(t) = A(t)M(t)$. M vérifie l'équation caractéristique de la résolvante dans $\mathscr{L}(E,E)$ et donc est la résolvante de (ED).

$\blacksquare$

Exemples. *i)* *Si A et B sont des matrices qui commutent et $A(t) = f(t)A + g(t)B$ où f et g sont des fonctions scalaires, alors*

$$R(t,t_0) = \exp\Big(\int_{t_0}^{t} f(s)\,dsA + \int_{t_0}^{t} g(s)\,dsB\Big) = \exp\Big(\int_{t_0}^{t} f(s)\,dsA\Big)\exp\Big(\int_{t_0}^{t} g(s)\,dsB\Big)$$

ii) *Le système*

$$\begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix} = \begin{pmatrix} 1/t & t \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$$

défini sur $\mathbb{R}_*^+$ ne vérifie pas l'hypothèse de commutation et on peut montrer que

$$R(t,t_0) \neq \exp\Big(\int_{t_0}^{t} A(s)\,ds\Big)$$

4.3 Equations différentielles linéaires avec deuxième membre

Définition 4.3.1. *On appelle équation différentielle linéaire, définie sur un intervalle I de $\mathbb{R}$, à valeurs dans E, avec deuxième membre une équation différentielle de la forme*

$$X'(t) = A(t)X(t) + B(t) \text{avec } \forall t \in I \ , \ A(t) \in \mathscr{L}(E) \text{ et } B \in C(I; E)$$

Théorème 4.3.2. *(Solution d'une équation différentielle linéaire avec deuxième membre) Si l'application $t \to A(t)$ est continue de I dans $\mathscr{L}(E)$, l'équation différentielle linéaire avec deuxième membre :*

$$(ED) \ X'(t) = A(t)X(t) + B(t)$$

admet une et une seule solution, définie sur I et vérifiant la condition initiale : $X(t_0) = x_0$. Cette solution est donnée par la formule

$$X(t) = R(t, t_0)x_0 + \int_{t_0}^{t} R(t, \theta)B(\theta)\, d\theta$$

Démonstration : La démonstration de l'existence et l'unicité de la solution est la même que celle des équations différentielles linéaires sans deuxième membre (4.2.2)
Pour obtenir la forme de la solution, on utilise la méthode dite de variations des constantes, qui déja connue dans le cas où $E = \mathbb{R}$:
On pose

$$X(t) = R(t, t_0)Y(t)$$

D'où

$$X'(t) = R'_t(t, t_0)Y(t) + R(t, t_0)Y'(t)$$

C'est à dire

$$X'(t) = A(t)R(t, t_0)Y(t) + R(t, t_0)Y'(t)$$

En reportant dans l'équation différentielle, il vient :

$$R(t, t_0)Y'(t) = B(t)$$

C'est à dire

$$Y'(t) = R(t_0, t)B(t)$$

La condition initiale est $Y(t_0) = x_0$, d'où

$$Y(t) = x_0 + \int_{t_0}^{t} R(t_0, \theta)B(\theta)\, d\theta$$

et en revenant à X :

$$X(t) = R(t, t_0)x_0 + R(t, t_0)\int_{t_0}^{t} R(t_0, \theta)B(\theta)\, d\theta =$$

$$R(t, t_0)x_0 + \int_{t_0}^{t} R(t, t_0)R(t_0, \theta)B(\theta)\, d\theta =$$

$$R(t, t_0)x_0 + \int_{t_0}^{t} R(t, \theta)B(\theta)\, d\theta$$

■

Corollaire 4.3.3. *L'ensemble des solutions d'une équation différentielle linéaire avec deuxième membre sur un espace vectoriel normé E est un espace affine dont l'espace vectoriel associé est isomorphe à E.*

4.4 Equations différentielles linéaires à coefficients constants

Rappelons d'abord quelques propriétés de l'exponentielle d'une application linéaire :

Définition 4.4.1. *Soit E un espace de banach et $A \in \mathscr{L}(E)$. L'exponentielle de A est définie par*

$$e^A = \sum_{n=0}^{+\infty} \frac{A^n}{n!}$$

où A^n est la n-ième puissance de composition de A par lui-même.

Remarque 4.4.2. *La série ci-dessus est normalement convergente donc convergente. En effet, on peut écrire*

$$\left\| \frac{A^n}{n!} \right\| \leq \frac{1}{n!} \left\| A \right\|^n$$

et la série numérique $\sum \frac{1}{n!} \left\| A \right\|^n$ est convergente.

Théorème 4.4.3. *(Propriétés de l'exponentielle)*
Si A et $B \in \mathscr{L}(E)$ commutent, on a

$$e^{A+B} = e^A \circ e^B$$

Démonstration : C'est la même que dans le cas scalaire : comme les séries définissant e^A et e^B sont normalement convergentes, on a le droit de sommer comme on veut, d'où :

$$e^A \circ e^B = \sum_{m,n=0}^{+\infty} \frac{A^m B^n}{n!m!} = \sum_{p=0}^{+\infty} \sum_{m=0}^{p} \frac{A^m B^{p-m}}{m!(p-m)!} =$$

$$\sum_{p=0}^{+\infty} \frac{1}{p!} \sum_{m=0}^{p} C_m^p A^m B^{p-m} = \sum_{p=0}^{+\infty} \frac{1}{p!} (A+B)^p = e^{A+B}$$

■

Théorème 4.4.4. *(Résolvante d'une équation différentielle linéaire à coefficients constants)*
La résolvante de l'équation différentielle linéaire à coefficients constants

$$(ED)\ X'(t) = AX(t)\ ,\ \text{avec } A \in \mathscr{L}(E)$$

est

$$R(t_1, t_2) = e^{(t_1 - t_2)A}$$

DÉMONSTRATION : Comme la série définissant e^{tA} est normalement convergente, on a le droit de dériver terme à terme. D'où

$$\frac{d}{dt}\left(e^{tA}\right) = \sum_{n=1}^{+\infty} \frac{t^{n-1}A^n}{n!}$$

Donc aussi

$$\frac{d}{dt}\left(e^{tA}\right) = Ae^{tA}$$

Si l'on pose $R(t,\theta) = e^{(t-\theta)A}$, on aura :

$$R'_t(t,\theta) = Ae^{(t-\theta)A} = AR(t,\theta) \text{ et } R(\theta,\theta) = I_E$$

ce qui montre que R ainsi défini est bien la résolvante de l'équation différentielle linéaire à coefficients constants considérée. ∎

Corollaire 4.4.5. *La solution de l'équation différentielle linéaire à coefficients constants avec deuxième membre*

$$(ED)\ X'(t) = AX(t) + B(t)$$

telle que $X(t_0) = x_0$ *est :*

$$X(t) = e^{(t-t_0)A}x_0 + \int_{t_0}^{t} e^{(t-\theta)A}B(\theta)\,d\theta$$

Dans le cas particulier où E est un espace vectoriel de dimension finie, on peut exprimer ces résultats à l'aide des valeurs propres de A :

Supposons que la dimension de E est n. Soient $\{\lambda_i \mid i = 1,2,...k\}$ les valeurs propres de A, $\{r_i \mid i = 1,2,...k\}$ leur ordre de multiplicité dans le polynôme caractéristique de A et $\{\pi_i \mid i = 1,2,...k\}$ les projecteurs spectraux associés.

Alors, il existe des endomorphismes nilpotents $\{N_i \mid i = 1,2,...k\}$ d'indice k_i tels que :

$$A = \sum_{i=1}^{k}(\lambda_i + N_i)\pi_i$$

$$e^{tA} = \sum_{i=1}^{k} e^{\lambda_i t}e^{tN_i}\pi_i = \sum_{i=1}^{k} e^{\lambda_i t}(I_E + tN_i + ... + \frac{t^{k_i-1}}{(k_i-1)!}N_i^{k_i-1})\pi_i$$

4.5 Equations différentielles linéaires d'ordre p à coefficients constants

On considère l'équation différentielle

$$(ED_p)\ a_p y^{(p)} + ... + a_1 y' + a_0 y = b(t)$$

où y est la fonction inconnue définie sur $I \in \mathbb{R}$, à valeurs dans $\mathbb{R}$, b est une fonction définie sur I à valeurs dans $\mathbb{R}$ et les coefficients a_i , $i = 1, ...p$ sont des constantes appartenant à $\mathbb{R}$, avec $a_p \neq 0$.
Posons

$$A = \begin{pmatrix} 0 & 1 & 0 & ... & 0 \\ 0 & 0 & 0 & ... & 0 \\ . & & & & \\ . & & & & \\ . & & & & \\ 0 & 0 & 0 & ... & 1 \\ c_0 & c_1 & c_2 & ... & c_{p-1} \end{pmatrix} \ , \ c_j = -\frac{a_j}{a_p} \ , \ B(t) = \begin{pmatrix} 0 \\ 0 \\ . \\ . \\ . \\ 0 \\ b(t) \end{pmatrix}$$

alors (ED_p) est équivalente au système :

$$(S_p)\ Y'(t) = AY(t) + B(t)$$

où l'on a posé

$$Y(t) = \begin{pmatrix} y(t) \\ y'(t) \\ y''(t) \\ . \\ . \\ . \\ y^{p-1}(t) \end{pmatrix}$$

On déduit des résultats généraux :

Théorème 4.5.1. *L'ensemble des solutions de l'équation différentielle (ED_p) est un espace affine dont l'espace vectoriel associé est de dimension p.*

4.6 Wronskien d'un système de solutions

On considère l'équation différentielle

$$(ED)\ Y'(t) = A(t)Y(t)$$

où les fonctions sont à valeurs dans l'espace $\mathbb{R}^m$.

Définition 4.6.1. *Le Wronskien d'un système de m solutions $(Y_1, ..., Y_m)$ de (ED) est*

$$W(t) = \det(Y_1(t), ..., Y_m(t))$$

Théorème 4.6.2. *Si $V_j = Y_j(t_0)$ pour $j = 1, 2, ...m$, alors $Y_j(t) = R(t, t_0)V_j$ et on a*

$$W(t) = \exp\big(\int_{t_0}^{t} tr(A(s)\,ds\big)\det(V_1, ..., V_m)$$

5

OPTIMISATION

5.1 Préliminaires

Dans ce chapitre, on va étudier des problèmes du type suivant : étant données une fonction à valeurs réelles, définie sur une partie Ω d'un espace vectoriel normé E et une partie $\Delta \subset \Omega$, on recherche les points de Δ en lesquels la trace de f sur Δ est minimale.

Généralement, Δ sera défini par des égalités appelées *liaisons* et des inégalités appelées *contraintes*.

Définition 5.1.1. *Soit $\tilde{x} \in \Delta$ et $f : \Omega \subset E \to \mathbb{R}$.*

1) On dit que $f(x)$ est minimale en $\tilde{x}$ sur Δ si

$$\forall x \in \Delta \ , \ f(x) \geq f(\tilde{x})$$

2) On dit que $f(x)$ est strictement minimale en $\tilde{x}$ sur Δ si

$$\forall x \in \Delta \ , \ x \neq \tilde{x} \ , \ f(x) > f(\tilde{x})$$

3) On dit que $f(x)$ est localement minimale en $\tilde{x}$ sur Δ si

$$\exists V \text{ voisinage de } \tilde{x} \text{ tel que } \forall x \in V \cap \Delta \ , \ f(x) \geq f(\tilde{x})$$

4) On dit que $f(x)$ est localement strictement minimale en $\tilde{x}$ sur Δ si

$$\exists V \text{ voisinage de } \tilde{x} \text{ tel que } \forall x \in V \cap \Delta \ x \neq \tilde{x} \ , \ f(x) > f(\tilde{x})$$

Il est clair que si f est (strictement) minimale en $\tilde{x}$, elle est (strictement) localement minimale en ce point.

Cas particulier 5.1.2. *On suppose que $E = \mathbb{R}$ et que f est dérivable sur un ouvert Ω de $\mathbb{R}$.*

1) Si $\tilde{x}$ est à l'intérieur de Δ et si f est localement minimale en $\tilde{x}$ sur Δ, alors $f'(\tilde{x}) = 0$.

2) Soit $a \leq b$ et $\Delta = [a, b]$. Si f est localement minimale en a pour $x \in [a, b]$, alors $f'(a) \geq 0$ et si f est localement minimale en b pour $x \in [a, b]$, alors $f'(b) \leq 0$.

Démonstration : *1)* Pour tout h dans un voisinage de 0 et tel que $\tilde{x} + h \in \Delta$, on a :

$$f(\tilde{x} + h) - f(\tilde{x}) \geq 0$$

Donc par définition de la dérivée :

$$\lim_{h>0\ ,\ h\to 0} \frac{f(\tilde{x} + h) - f(\tilde{x})}{h} \geq 0$$

$$\lim_{h<0\ ,\ h\to 0} \frac{f(\tilde{x} + h) - f(\tilde{x})}{h} \leq 0$$

D'où

$$f'(\tilde{x}) = \lim_{h\to 0} \frac{f(\tilde{x} + h) - f(\tilde{x})}{h} = 0$$

2) se démontre de même. ■

Cas particulier 5.1.3. *On suppose encore que $E = \mathbb{R}$ et que f est dérivable sur un ouvert Ω de $\mathbb{R}$. Si de plus f est convexe, alors :*

1) f est localement minimale sur Ω en $\tilde{x} \in \Omega$ si et seulement si elle est minimale en $\tilde{x}$ sur Ω.

2) Supposons que f est deux fois dérivable en $\tilde{x}$, alors si $f'(\tilde{x}) = 0$ et si $f''(\tilde{x}) > 0$, f admet un minimun en $\tilde{x}$ sur Ω.

Démonstration : On rappelle qu'une fonction $f : \Omega \to \mathbb{R}$ est *convexe* si

$$\forall x, y \in \Omega\ ,\ \forall t \in [0, 1]\ ,\ f\big(tx + (1 - t)y\big) \leq tf(x) + (1 - t)f(y)$$

1) Supposons que f soit localement minimale en x et ne soit pas minimale en ce point. Alors il existe $y \in \Omega$ tel que $f(y) < f(x)$.
Mais par convexite,

$$\forall t \in [0, 1]\ ,\ f\big(tx + (1 - t)y\big) \leq tf(x) + (1 - t)f(y) < f(x)$$

Quand $t \to 1$, $tx + (1 - t)y \to x$ et ceci contredit la minimalité locale de f en x.
2) Par la formule de Taylor, on peut écrire :

$$f(\tilde{x} + h) - f(\tilde{x}) = \frac{h^2}{2} f''(\tilde{x}) + h^2\epsilon(h)\ ,\ \text{avec } \epsilon(h) \to 0\ ,\ \text{quand } h \to 0$$

Quand h est suffisamment petit, $f''(\tilde{x}) + \epsilon(h) \geq 0$ et par suite, $f(\tilde{x} + h) - f(\tilde{x}) \geq 0$. Donc f admet un minimum local en $\tilde{x}$ et d'après *1)*, ce minimum est global. ■

5.2 Optimisation sur un ouvert

Théorème 5.2.1. *(Optimisation sur un ouvert)*

Soit f une fonction différentiable sur un ouvert Ω d'un espace vectoriel normé E, à valeurs réelles. Si f est localement minimale en $\tilde{x}$ pour $x \in \Omega$, alors $f'(\tilde{x}) = 0$.

DÉMONSTRATION : Soit $h \in E$. Posons $F(t) = f(\tilde{x} + th)$. Alors, F est définie dans un voisinage de 0 et est localement minimale en 0. D'où $F'(0) = 0$, c'est à dire $f'(\tilde{x})h = 0$. ■

Remarque 5.2.2. *Ce thèorème reste valable si Ω est d'intérieur non vide et si $\tilde{x}$ est dans l'intérieur de Ω.*

Cas particulier 5.2.3. *On suppose que f est différentiable sur un ouvert Ω de E.*

Si f est convexe, f est localement minimale sur Ω en $\tilde{x} \in \Omega$ si et seulement si elle est minimale en $\tilde{x}$ sur Ω.

DÉMONSTRATION : On utilise la démonstration du cas scalaire (5.1.3) en l'appliquant à la fonction $F(t) = f(\tilde{x} + th)$. ■

5.3 Optimisation sur une variété affine

Définition 5.3.1. *1) On appelle variété affine d'un espace vectoriel normé E un ensemble de la forme $H = a + V$ où $a \in E$ et V est un sous-espace vectoriel de E.*

2) V est appelé direction de H.

On remarquera que

$$V = \{x_1 - x_2 \mid x_1, x_2 \in H\}$$

et que H est fermé si et seulement si V l'est.

Théorème 5.3.2. *(Condition nécessaire du premier ordre sur une variété affine) Soient E un espace vectoriel normé, H une variété affine fermée de E et V la direction de H. Soit f une fonction différentiable d'un ouvert Ω de E, à valeurs dans $\mathbb{R}$. Soit $\tilde{x} \in \Omega \cap H$.*

Pour que f soit minimale en $\tilde{x}$ sur $\Omega \cap H$, il faut que

$$\forall h \in V \ , \ f'(\tilde{x})h = 0$$

DÉMONSTRATION : Posons

$$\forall h \in V \ , \ F(t) = f(\tilde{x} + th)$$

Alors F est localement minimal en 0, d'où $F'(0) = 0$.
Or $F'(0) = f'(\tilde{x})h$ d'où le résultat. ■

Remarque 5.3.3. *Soit F un espace vectoriel normé. Supposons que $H = \{x \in E \mid Ax = b\}$, où $A \in \mathscr{L}(E, F)$ et $b \in F$. Alors $V = \{x \in E \mid Ax = 0\}$ et la condition nécessaire du premier ordre 5.3.2 s'écrit :*

$$Ah = 0 \Rightarrow f'(\tilde{x})h = 0$$

Pour exprimer cette condition, on a besoin d'un théorème de Banach, qui prolonge le théorème de Banach énoncé dans la partie 1.2.

Théorème 5.3.4. *(Théorème de factorisation)*
Soient E, F et G trois espaces de banach, $A \in \mathscr{L}(E, F)$ et $B \in \mathscr{L}(E, G)$. On suppose que

$$Ax = 0 \Rightarrow Bx = 0$$

Si de plus A est surjective ou si F est de dimension finie, alors il existe $C \in \mathscr{L}(F, G)$ tel que

$$B = CA$$

Théorème 5.3.5. *(Règle du multiplicateur)*
Soient E et F deux espaces de banach, $A \in \mathscr{L}(E, F)$ et $b \in F$. Soit f une fonction réelle dérivable définie sur un ouvert Ω de E, à valeurs réelles. Soit

$$H = \{x \in \Omega \mid Ax = b\} \text{ et } \tilde{x} \in H$$

Si A est surjective ou si F est de dimension finie, alors si f est localement minimale en $\tilde{x}$ sur H, il existe $\lambda \in F'$ tel que

$$f'(\tilde{x}) = \lambda A$$

Démonstration : On a en effet :

$$Ah = 0 \Rightarrow f'(\tilde{x})h = 0$$

■

Définition 5.3.6. *L'élément $\lambda \in F'$ introduit par le théorème 5.3.5 est appelé multiplicateur de Lagrange.*

Exemple 5.3.7. *Si $F = \mathbb{R}^p$, on a*

$$Ax = \begin{pmatrix} A^1x \\ A^2x \\ . \\ . \\ A^px \end{pmatrix}$$

où pour tout $i = 1, 2, ..., p$, $A^i \in E'$ et

$$b = \begin{pmatrix} b^1 \\ b^2 \\ . \\ . \\ b^p \end{pmatrix}$$

H est l'ensemble des $x \in E$ tels que

$$\forall i = 1, 2, ..., p \ , \ A^i x = b^i$$

On aura

$$\lambda = \begin{pmatrix} \lambda_1 & \lambda_2 & . & . & \lambda_p \end{pmatrix}$$

et la règle du multiplicateur s'écrit :

$$f'(\tilde{x}) = \sum_{i=1}^{p} \lambda_i A^i$$

Chaque λ_i est appelé multiplicateur de Lagrange relatif à chacune des liaisons $A^i x = b_i$.

5.4 Optimisation avec liaisons.

Soient E et F deux espaces de banach.

Définition 5.4.1. *1) On appelle surface d'un espace vectoriel normé E un ensemble de la forme $S = \{x \in E \ , \ g(x) = 0\}$ où g est une fonction de classe C^1 d'un ouvert Ω de E dans F .*

2) $g'(\tilde{x})$ est appelée direction de S au point $\tilde{x}$.

On remarquera que $S = g^{-1}(0)$ est fermée dans E.

Théorème 5.4.2. *(Condition nécessaire du premier ordre sur une surface) Soient E un espace vectoriel normé, S une surface de E. Soit f une fonction différentiable d'un ouvert Ω de E, à valeurs dans $\mathbb{R}$. Soit $\tilde{x} \in \Omega \cap S$. On suppose de plus :*

1) La dérivée $g'(\tilde{x})$ de g en $\tilde{x}$ est surjective

2) Le sous espace vectoriel $K = Kerg'(\tilde{x})$ admet un supplémentaire fermé G

Alors, pour que f soit minimale en $\tilde{x}$ sur S, il faut que

$$\forall u \in Kerg'(\tilde{x}) \ , \ f'(\tilde{x})u = 0$$

Démonstration : : Grâce à l'hypothèse *2)*, tout vecteur $x \in E$ s'écrit de façon unique comme $x = (x_1, x_2)$ où $x_1 \in K, x_2 \in G$. Les fonctions f et g apparaissent donc comme des fonctions de deux variables.

La dérivée de g en $\tilde{x}$ s'écrit :

$$g'(\tilde{x}) = g'(\tilde{x}_1, \tilde{x}_2) = \begin{pmatrix} g'_1(\tilde{x}_1, \tilde{x}_2) & g'_1(\tilde{x}_1, \tilde{x}_2) \end{pmatrix}$$

Comme $K = Kerg'(\tilde{x})$, $g'_1(\tilde{x}_1, \tilde{x}_2) = 0$ dans$\mathscr{L}(K, F)$.

L'application linéaire de G dans F, $g'_1(\tilde{x}_1, \tilde{x}_2)$ est surjective par hypothèse, injective par construction et continue car g est supposée de classe C^1. Par le théorème de l'application ouverte (conséquence admise du théorème de Banach), c'est un isomorphisme de G dans F.

Par le théorème des fonctions implicites, il existe des voisinages V de $\tilde{x}_1$ dans K, W de $\tilde{x}_2$ dans G et une fonction h de classe C^1 de V sur W tels que sur $V \times W$, la relation $g(x_1, x_2) = 0$ soit équivalente à $x_2 = h(x_1)$.
De plus, on a, pour $x_1 \in V$:

$$h'(x_1) = g_2'\big(x_1, h(x_1)\big)^{-1} \circ g_1'(\big(x_1, h(x_1)\big)$$

En particulier, $h'(\tilde{x}_1) = 0$.
Si f est localement minimale en $\tilde{x} = (\tilde{x}_1, \tilde{x}_2)$ sur S, alors la fonction : $\phi : x_1 \to \phi(x_1) = f\big(x_1, h(x_1)\big)$ est minimale en $\tilde{x}_1$ sur V et d'après le théorème 5.2.1, la dérivée de ϕ au point $\tilde{x}_1$ s'annule sur K.
Or, en appliquant les résultats du chapitre 1 sur la composition d'applications différentiables, la dérivée de ϕ en $\tilde{x}_1$ vaut :

$$f_1'\big(\tilde{x}_1, h(\tilde{x}_1)\big) + f_2'\big(\tilde{x}_1, h(\tilde{x}_1)\big)h'(\tilde{x}_1)$$

$$= f_1'\big(\tilde{x}_1, h(\tilde{x}_1)\big)$$

En identifiant le couple $(u, 0)$ de $K \times G$ au vecteur $u \in K \subset E$, dire que la dérivée de ϕ en $\tilde{x}_1$ s'annule sur K équivaut à la condition

$$\forall u \in Kerg'(\tilde{x}) \ , \ f'(\tilde{x})u = 0$$

Remarque 5.4.3. *Si E et F sont de dimension finie, la condition 2) du théorème 5.4.2 est automatiquement vérifiée.*

Grâce au théorème de factorisation énoncé dans le paragraphe précédent, on obtient une façon d'exprimer ce résultat à l'aide d'un multiplicateur de Lagrange :

Théorème 5.4.4. *(Règle du multiplicateur avec liaison)*
Soient E et F deux espaces de banach, S une surface de E définie par une fonction g de classe C^1 définie sur un ouvert Ω de E. Soit f une fonction différentiable, définie sur l'ouvert Ω de E, à valeurs réelles. Soit $\tilde{x} \in \Omega \cap S$. On suppose de plus que :

1) *La dérivée $g'(\tilde{x})$ de g en $\tilde{x}$ est surjective*
2) *Le sous espace vectoriel $K = Kerg'(\tilde{x})$ admet un supplémentaire fermé G*

alors si f est localement minimale en $\tilde{x}$ sur S, il existe $\lambda \in F'$ tel que

$$f'(\tilde{x}) = \lambda g'(\tilde{x})$$

Exemple. *Calculez le volume maximal d'une boîte rectangulaire de surface latérale dépliée égale à a^2.*

Remarque. *Dans le cas où on chercher à optimiser une fonction f avec une contrainte, c'est à dire sur un ensemble de la forme : $\{x \in E \ , \ g(x) \leq 0\}$ où g est une fonction continue, on peut considérer deux cas :*

1) *On optimise sur l'ouvert $\{x \in E \ , \ g(x) < 0\}$ en appliquant le théorème 5.2.1*
2) *On optimise sur la surface $\{x \in E \ , \ g(x) = 0\}$ en appliquant le théorème 5.4.2*

et on compare.

Extension 5.4.5. *Les notations étant celles du théorème 5.4.2, supposons que* $F = F_1 \times F_2$ *et que* $g(x) = \begin{pmatrix} g_1(x) \\ g_2(x) \end{pmatrix}$ *où* g_1 *est affine* ($g_1(x) = Ax + b$) *et* F_2 *est de dimension finie. Si la restriction de* $g_2'(\tilde{x})$ *à* $KerA$ *est surjective et si* f *est localement minimale en* $\tilde{x}$ *sur* $S = \{x \in \Omega \ , \ g(x) = 0\}$, *il existe* $\lambda \in (F_2)'$ *tel que*

$$\forall u \in KerA \ , \ f'(\tilde{x})u - \lambda g_2'(\tilde{x})u = 0$$

5.5 Calcul des variations

Le problème consiste à rechercher les fonctions qui minimisent certaines intégrales. On a besoin d'un lemme préliminaire :

Lemme 5.5.1. *L'espace des fonctions de classe* C^1 *définies sur un intervalle* $[a, b]$, *à valeurs dans un espace de Banach* E, *noté* $\mathscr{C}^1([a, b]; E)$ *est un espace de banach pour la norme :* $\|X\|_\infty + \|X'\|_\infty$. *L'ensemble des fonctions* $X \in \mathscr{C}^1([a, b]; E)$ *telles que* $X(a) = \alpha$, $X(b) = \beta$ *est un sous espace affine fermé de* $\mathscr{C}^1([a, b]; E)$.

Théorème 5.5.2. *(Conditions nécessaires du premier ordre en calcul des variations)*
Soit E *un espace vectoriel normé de dimension finie,* $[a, b]$ *un intervalle de* $\mathbb{R}$, L *une application continue de* $[a, b] \times E^2$ *dans* $\mathbb{R}$, *admettant des dérivées partielles* L'_x *et* L'_u *par rapport à la deuxième et la troisième variable, continues sur* $[a, b] \times E^2$.
Pour que

$$\int_a^b L\big(t, X(t), X'(t)\big)\, dt$$

soit localement minimale en $\tilde{X}$ *pour l'ensemble des fonction* X *de classe* C^1 *et telles que* $X(a) = \alpha$ *et* $X(b) = \beta$, *il faut que l'on ait :*

$$\frac{\partial}{\partial t} L'_u\big(t, \tilde{X}(t), \tilde{X}'(t)\big) = L'_x\big(t, \tilde{X}(t), \tilde{X}'(t)\big)$$

Définition 5.5.3. *L'équation*

$$\frac{\partial}{\partial t} L'_u\big(t, \tilde{X}(t), \tilde{X}'(t)\big) = L'_x\big(t, \tilde{X}(t), \tilde{X}'(t)\big)$$

s'appelle l'équation d'Euler.

Démonstration : On va appliquer le théorème 5.3.2 :
L'application

$$X, U \rightarrow I(X, U) = \int_a^b L\big(t, X(t), U(t)\big)\, dt$$

est continûment différentiable et sa dérivée en $\tilde{X}, \tilde{U}$ est donnée par

$$I'(\tilde{X}, \tilde{U})(\partial X, \partial U) = \int_a^b \left[L'_x\big(t, X(t), U(t)\big)\partial X(t) + L'_u\big(t, X(t), U(t)\big)\partial U(t)\right] dt$$

Si $\tilde{X}$ minimise $I(X, X')$ pour $X \in \mathscr{C}^1([a,b]; E)$, $X(a) = \alpha$, $X(b) = \beta$, alors $\tilde{X}, \tilde{X}'$ minimise $I(X, U)$ pour

$$\begin{cases} X \in \mathscr{C}^1([a,b]; E) \ , \ X(a) = \alpha \ , \ X(b) = \beta \\ U \in \mathscr{C}([a,b]; E) \ , \ U = X' \end{cases}$$

La condition nécessaire du premier ordre s'écrit pour

$$\partial X \in \mathscr{C}^1([a,b]; E), \partial U \in \mathscr{C}([a,b]; E)$$

$$\left\{ \begin{array}{l} \partial X(a) = 0 \\ \partial X(b) = 0 \\ \partial U = \partial X' \end{array} \right\} \Rightarrow I'(\tilde{X}, \tilde{X}')(\partial X, \partial U) = 0$$

Compte tenu de la relation $\partial X(a) = 0$, ∂X s'exprime ainsi en fonction de ∂U :

$$\partial X(t) = \int_a^t \partial U(s)\, ds$$

La condition $\partial X(b) = 0$ s'écrit

$$\int_a^b \partial U(s)\, ds = 0$$

Définissons :

$$G = \{\partial U \in \mathscr{C}([a,b]; E) \mid \int_a^b \partial U(s)\, ds = 0\}$$

Nous avons à écrire que

$$I'(\tilde{X}, \tilde{X}')(\partial X, \partial U) = 0$$

pour tout $\partial U \in G$, ∂X étant donné par $\partial X(t) = \displaystyle\int_a^t \partial U(s)\, ds$.

Nous allons faire une intégration par parties sur le premier terme de $I'(\tilde{X}, \tilde{X}')(\partial X, \partial U)$ pour l'exprimer à l'aide de ∂U : Posons

$$F(t) = \int_a^t L'_x\big(s, \tilde{X}(s), \tilde{X}'(s)\big)\, ds$$

Alors $F \in \mathscr{C}^1([a,b]; F)$ et

$$\int_a^b L'_x\big(t, \tilde{X}(t), \tilde{X}'(t)\big)\partial X(t)\, dt = \int_a^b F'(t)\partial X(t)\, dt =$$

$$F(b)\partial X(b) - F(a)\partial X(a) - \int_a^b F(t)(\partial X)'(t)\, dt$$

D'où pour $\partial U \in G$ et $\partial X(t) = \displaystyle\int_a^t \partial U(s)\,ds$,

$$I'(\tilde{X}, \tilde{X}')(\partial X, \partial U) = \int_a^b \big[- F(t) + L'_u\big(t, \tilde{X}(t), \tilde{X}'(t)\big)\big] \partial U(t)\,dt$$

Cette expression doit être nulle pour tout $\partial U \in G$. D'après le théorème de factorisation, il existe $\lambda \in E'$ tel que

$$\int_a^b \big[- F(t) + L'_u\big(t, \tilde{X}(t), \tilde{X}'(t)\big)\big] \partial U(t)\,dt = \lambda \int_a^b \partial U(t)\,dt$$

pour tout $\partial U \in \mathscr{C}([a,b];E)$.
D'où

$$L'_u\big(t, \tilde{X}(t), \tilde{X}'(t)\big) - F(t) = \lambda$$

ce qui par dérivation par rapport à t donne le résultat annoncé. ■

Remarques 5.5.4. *1) Si on voulait exprimer $I'(\tilde{X}, \tilde{X}')(\partial X, \partial U)$ en fonction de ∂X en utilisant la relation $\partial U = \partial X'$, cela conduirait à faire une intégration par parties sur le deuxième terme de $I'(\tilde{X}, \tilde{X}')(\partial X, \partial U)$. Or on ne sait pas si $L'_u\big(t, \tilde{X}(t), \tilde{X}'(t)\big)$ est dérivable par rappport à t. On peut vérifier que si c'est le cas, cette autre méthode donne le même résultat.*

2) La méthode utilisée consiste essentiellement à minimiser par rapport à U sur l'ensemble

$$\{U \in \mathscr{C}([a,b];E) \mid \int_a^b U(t)\,dt = X(b) - X(a)\}$$

La constante λ est le multiplicateur associé à cette liaison.

3) On pourrait aussi garder les deux variables X et U et appliquer la règle des multiplicateurs avec les liaisons :

$$X(a) = \alpha \ , \ X(b) = \beta \ , \ X' = U$$

Exemple 5.5.5. *i) En remplaçant les liaisons $X(a) = \alpha$, $X(b) = \beta$ par $AX(a) = \alpha$, $BX(b) = \beta$, où $A \in \mathscr{L}(E, \mathbb{R}^p)$, $B \in \mathscr{L}(E; \mathbb{R}^q)$, la condition nécessaire du premier ordre devient la condition d'Euler et les conditions transversales :*

$$L'_u\big(a, \tilde{X}(a), \tilde{X}'(a)\big) = \lambda A \ , \ L'_u\big(b, \tilde{X}(b), \tilde{X}'(b)\big) = \mu B$$

ii) Si l'on conserve la liaison $X(a) = \alpha$ mais que l'on supprime la liaison $X(b) = \beta$, la condition du premier ordre est l'équation d'Euler plus la condition

$$L'_u\big(b, \tilde{X}(b), \tilde{X}'(b)\big) = 0$$

Exemple. *Soit $E = (\mathbb{R}^n, \|.\|_2)$. Trouver la ligne de longueur minimale dans $\mathbb{R} \times E$ joignant $\{t = a, AX(a) = \alpha\}$ et $\{t = b, BX(b) = \beta\}$, où A et B sont des applications linéaires données de E dans $\mathbb{R}^p$.*

Exemple. *Minimiser $I(r) = \int_0^{2\pi} (r^2 + r'^2)^{1/2}\, d\theta$ sur l'ensemble des fonction $r(\theta)$ de classe C^1 telles que $r(0) = r(2\pi)$ et $J(r) = \displaystyle\int_0^{2\pi} \frac{r^2}{2}\, d\theta = S$, S étant fixé. Ceci revient à déterminer les courbes fermées en coordonnées polaires de longueur minimale limitant une aire donnée.*

Démonstration : : On applique l'extension 5.4.5 et le calcul des variations.

5.6 Conditions du deuxième ordre en dimension finie

Théorème 5.6.1. *(Conditions du deuxième ordre sur un ouvert)*

1) Soit E un espace vectoriel de dimension finie, f une fonction de classe C^2, définie sur un ouvert Ω de E. Soit $\tilde{x} \in \Omega$. Si l'on a

$$f'(\tilde{x}) = 0 \text{ et } \forall h \neq 0 \ , \ f''(\tilde{x})(h, h) > 0$$

alors f est localement minimale en $\tilde{x}$ sur Ω.

2) Réciproquement, si f est localement minimale en $\tilde{x}$, on a :

$$f'(\tilde{x}) = 0 \text{ et } \forall h \in E \ , \ f''(\tilde{x})(h, h) \geq 0$$

Démonstration : *1)* On a :

$$f(\tilde{x} + h) = f(\tilde{x}) + \frac{1}{2} f''(\tilde{x})(h, h) + \varepsilon(h)\, \|h\|^2$$

avec $\varepsilon(h) \to 0$ quand $h \to 0$.
Par hypothèse, $f''(\tilde{x})$ est une forme quadratique définie positive sur l'espace E qui est de dimension finie. Il existe donc des constantes $a, b > 0$ telles que

$$a\, \|h\|^2 \leq f''(\tilde{x})(h, h) \leq b\, \|h\|^2$$

Par suite

$$f(\tilde{x} + h) \geq f(\tilde{x}) + \left[\frac{a}{2} + \varepsilon(h)\right] \|h\|^2$$

Donc pour $\|h\|$ assez petit, on aura :

$$f(\tilde{x} + h) > f(\tilde{x})$$

et $\tilde{x}$ est bien un minimum local pour f sur Ω.
2) Il suffit de considérer la fonction

$$F(t) = f(\tilde{x} + th)$$

et d'écrire que si $\tilde{x}$ est un minimum local de f sur Ω, alors $F'(0) = 0$ et $F''(0) \geq 0$. ■

Remarque 5.6.2. *Si f est convexe et deux fois différentiable en $\tilde{x}$, alors si $f'(\tilde{x}) = 0$ et si $f''(\tilde{x})$ est une forme quadratique définie positive, f admet un minimun en $\tilde{x}$ sur Ω.*

On remarque que comme dans le cas scalaire, si f est localement minimale en $\tilde{x}$, elle est minimale en ce point.

Théorème 5.6.3. *(Conditions du deuxième ordre avec liaisons)*
Soit E un espace vectoriel de dimension finie, f une fonction de classe C^2, définie sur un ouvert Ω de E et g une application de classe C^2 de E dans $\mathbb{R}^p$. Soit $\tilde{x} \in \Omega$ tel que $g(\tilde{x}) = 0$. Si l'on a :

i) $g'(\tilde{x})$ *surjective*

ii) il existe $\lambda \in (\mathbb{R}^p)'$ *tel que* $f'(\tilde{x}) = \lambda g'(\tilde{x})$

et en posant $L = f - \lambda \circ g$ *:*

iii) $\forall h \neq 0 \ , \ h \in Ker g'(\tilde{x}) \ , \ L''(\tilde{x})(h,h) > 0$

alors f est localement minimale en $\tilde{x}$ sur $S = \{x \in \Omega \ , \ g(x) = 0\}$.

Démonstration : : admise

Exercice 5.6.4. *f et g étant comme dans le théorème 5.6.3 et vérifiant l'hypothèses i) de ce théorème, montrer que la propriété $\forall h \in Ker g'(\tilde{x}) \ , \ L''(\tilde{x})(h,h) \geq 0$ est une condition nécessaire pour que f soit minimale sur S.*

Exemple. *Déterminer les extréma de $f(x,y,z) = 2(x+y+z) - xyz$ sur la sphère définie par : $x^2 + y^2 + z^2 \leq 9$*

Index des définitions

Table des Matières