



Calcul différentiel et Equations différentielles

M. Le Gonidec

Table des matières

I	Motivations et rappels	5
I.1	Dérivabilité : le cas des fonctions réelles de la variable réelle	5
I.1.1	Définition	5
I.1.2	Nombre dérivé et tangente à la courbe	6
I.1.3	Quelques règles de calcul des dérivées	6
I.1.4	Théorème et inégalité des accroissements finis	6
I.1.5	Dérivée et fonction réciproque	7
I.1.6	extrema locaux	7
I.1.7	Limites de suite de fonctions dérivables	7
I.2	A propos d'une généralisation	7
I.3	Rappels sur les espaces vectoriels normés...	9
I.3.1	Définitions	9
I.3.2	Les outils...	11
I.4	Applications linéaires et multilinéaires	14
I.4.1	Définitions	14
I.4.2	Continuité et norme d'une application multilinéaire	16
II	Applications différentiables	17
II.1	Différentiabilité	17
II.1.1	Fonctions tangentes	17
II.1.2	Différentielle d'une application	18
II.1.3	Premiers exemples simples	20
II.2	Règles de calcul	20
II.2.1	Linéarité	21
II.2.2	Différentielles d'applications particulières	21
II.2.3	Différentielle d'une application composée	22
II.2.4	Différentielle d'une application à valeurs dans un espace vectoriel produit	23
II.3	Différentielle d'une application définie sur un espace vectoriel produit & dérivées partielles	25
II.3.1	Applications partielles et différentielles partielles	25
II.3.2	Lien entre différentielles partielles et différentielle d'une application	27
II.3.3	Cas particulier des applications de $\mathbb{R}^n$ dans $\mathbb{R}^p$	28
II.3.4	Dérivées directionnelles	31

II.3.5 Liens entre différentiabilité et Gâteaux-différentiabilité	32
II.4 Remarque sur la $\mathbb{R}$ -différentiabilité et la $\mathbb{C}$ -différentiabilité	32
III Inégalité de la moyenne	35
III.1 Enoncé(s) du théorème	35
III.1.1 Rappel : le cas des fonctions réelles de la variable réelle	35
III.1.2 Fonctions vectorielles de la variable réelle	36
III.1.3 Fonctions d'un $\mathbb{K}$ -evn dans un autre $\mathbb{K}$ -evn	38
III.2 Application : Retour sur les liens entre différentiabilité et dérivées partielles	39
III.3 Application : Convergence de suites de fonctions différentiables	42
IV Différentielles d'ordres supérieurs et extrema locaux	45
IV.1 différentielle seconde	45
IV.2 différentielles d'ordre supérieur	48
IV.3 formule de Taylor-Young	49
IV.4 Points critiques et extrema locaux	49
V Inversion locale et théorème des fonctions implicites	51
V.1 Problématique et isomorphismes topologiques...	51
V.2 Homéomorphismes et difféomorphismes de classe $\mathcal{C}^1$	53
V.3 Le théorème d'inversion locale	56
V.3.1 Enoncé du théorème	56
V.3.2 Le théorème du point fixe et la preuve du théorème d'inversion locale	57
V.4 Le théorème des fonctions implicites	59
VI Equations différentielles	61
VI.1 Généralités	61
VI.2 Le problème de Cauchy	64

Chapitre I

Motivations et rappels

L'idée fondamentale du calcul diff, c'est l'extension des résultats usuels concernant la dérivabilité des fonctions réelles de la variable réelle à des fonctions d'un espace vectoriel normé, vers un autre espace vectoriel normé.

Dans des problèmes physiques, on a rarement des fonctions de $\mathbb{R}$ dans $\mathbb{R}$, mais plutôt de $\mathbb{R}^n$ dans $\mathbb{R}^p$ ou d'un espace de suites (espace vectoriel de dimension infini) vers un autre espace vectoriel, fini ou infini... Dans ce cadre, il est très intéressant d'avoir un analogue de la notion de dérivabilité et de fonction dérivée. L'objet de ce cours est donc de présenter cet analogue : la différentiabilité et les applications différentielles.

Mais d'abord, quelques rappels pour fixer les choses...

I.1 Dérivabilité : le cas des fonctions réelles de la variable réelle

I.1.1 Définition

Définition I.1. Soit $f : \mathbb{R} \rightarrow \mathbb{R}$ une fonction. On dit que f est **dérivable en un point** $a \in \mathbb{R}$ si et seulement si la limite $\lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a}$ existe et est finie.

Lorsque cette limite finie existe, on la note $f'(a)$. Elle est appelé **nombre dérivé de f en a** .

Soit Ω un sous-ensemble de $\mathbb{R}$. On dit que f est **dérivable sur Ω** si f est dérivable en tout point de Ω .

Si $f : \mathbb{R} \rightarrow \mathbb{R}$ est dérivable et que sa fonction dérivée est continue, on dit que f est **de classe $\mathcal{C}^1$ sur $\mathbb{R}$** , on encore $f \in \mathcal{C}^1(\mathbb{R}, \mathbb{R})$.

Remarque I.2. On dit que f est **dérivable à gauche** (resp. **à droite**) **en un point** $a \in \mathbb{R}$ si et seulement si la limite $\lim_{x \rightarrow a, x < a} \frac{f(x) - f(a)}{x - a}$ (resp. $\lim_{x \rightarrow a, x > a} \frac{f(x) - f(a)}{x - a}$) existe et est finie.

Une première conséquence (directe) de la définition :

Proposition I.3. Soit $f : \mathbb{R} \rightarrow \mathbb{R}$ une fonction dérivable en un point $a \in \mathbb{R}$. La fonction f est continue en a .

I.1.2 Nombre dérivé et tangente à la courbe

On suppose qu'une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ est dérivable en $a \in \mathbb{R}$. La définition du nombre dérivé assure aussi l'égalité suivante :

$$f(x) = f(a) + f'(a)(x - a) + o(x - a).$$

Ainsi, la droite affine $T_f(a)$ d'équation $y = f(a) + f'(a)(x - a)$ est la meilleure approche linéaire de f au voisinage de a . Cette droite est la **tangente à f en a** .

Graphiquement, cela s'interprète comme suit : La courbe représentative de f vient s'écraser au voisinage de a sur la droite $T_f(a)$.

I.1.3 Quelques règles de calcul des dérivées

Soient f et g deux fonctions de $\mathbb{R}$ dans $\mathbb{R}$ et $\lambda \in \mathbb{R}$.

Linéarité : $(f + g)'(a) = f'(a) + g'(a)$ et $(\lambda f)'(a) = \lambda f'(a)$.

Produit d'applications : $(f \times g)'(a) = f'(a)g(a) + f(a)g'(a)$.

Puissance d'une application : $\forall k \in \mathbb{Z}^*, (f^k)'(a) = k f'(a)(f(a))^{k-1}$.

Composition d'applications (avec f dérivable en $g(a)$) : $(f \circ g)'(a) = f'(g(a))g'(a)$.

I.1.4 Théorème et inégalité des accroissements finis

I.1.4.1 Théorème des accroissements finis

Théorème I.4. Soit $f : [a, b] \rightarrow \mathbb{R}$ une application continue et dérivable sur $]a, b[$.

$$\exists c \in]a, b[, f(b) - f(a) = (b - a)f'(c)$$

Graphiquement, le théorème des accroissements finis indique que, pour toute sécante d'une courbe différentiable, il existe une tangente parallèle à la sécante. Pour donner une image, on peut illustrer ainsi le théorème : "Si un véhicule parcourt une distance à la vitesse moyenne de 60km/h, alors son compteur a indiqué au moins une fois la vitesse précise de 60 km/h."

Ce théorème est une conséquence du Théorème de Rolle : Soit $f : [a, b] \rightarrow \mathbb{R}$ une application continue et dérivable sur $]a, b[$ telle que $f(a) = f(b)$. Alors il existe (au moins) un réel $c \in]a, b[$ tel que $f'(c) = 0$. Il suffit d'appliquer ce résultat à la fonction $g : x \mapsto f(x) - f(a) - \frac{f(b)-f(a)}{b-a}(x - a)$.

I.1.4.2 Inégalité de la moyenne

Théorème I.5. Soient $f : [a, b] \rightarrow \mathbb{R}$ et $g : [a, b] \rightarrow \mathbb{R}$ deux applications continues et dérivable sur $]a, b[$. Si pour tout $t \in]a, b[$, on a $f'(t) \leq g'(t)$, alors on a :

$$\forall (x, y) \in [a, b]^2, x < y \implies f(y) - f(x) \leq g(y) - g(x).$$

En particulier, s'il existe un nombre $M > 0$ tel que, pour tout $t \in]a, b[$, on a $f'(t) \leq M$, alors :

$$\forall (x, y) \in [a, b]^2, x < y \implies f(y) - f(x) \leq M(y - x).$$

Il suffit pour obtenir ce résultat d'appliquer le théorème des accroissements finis à $h = g - f$. Le cas particulier s'obtient pour $g(t) = Mt$.

Pour donner une image, on peut illustrer ainsi le théorème : “Si un premier véhicule roule toujours moins vite qu'un deuxième véhicule, à temps égal, le premier aura tout parcouru moins que le second.”

I.1.5 Dérivée et fonction réciproque

Théorème I.6. Soit $f : \mathbb{R} \rightarrow \mathbb{R}$ une application continuellement dérivable sur un voisinage ouvert de $a \in \mathbb{R}$.

Si $f'(a) \neq 0$, alors f est inversible au voisinage de a : elle admet une fonction réciproque f^{-1} , définie au voisinage de $f(a)$. De plus, f^{-1} est dérivable sur un voisinage de $f(a)$ et on a : $(f^{-1})'(f(a)) = \frac{1}{f'(a)}$.

I.1.6 extrema locaux

Soit $f : \mathbb{R} \rightarrow \mathbb{R}$ une fonction. la fonction f admet un minimum (resp. maximum) local en $a \in \mathbb{R}$ s'il existe un voisinage ouvert de a sur lequel on a : $f(a) \leq f(x)$ (resp. $f(a) \geq f(x)$).

Lorsque f est dérivable et que la dérivée est continue, cela implique qu'on a une condition nécessaire pour que f admette un extremum local en $a \in \mathbb{R}$: $f'(a) = 0$ (Cette condition s'obtient facilement avec la définition du nombre dérivé).

Cette condition n'est pas suffisante : les “points selles” ont une dérivée nulle mais ne sont pas des extrema. Ex : $x \mapsto x^3$ a une dérivée nulle en 0, mais 0 n'est ni un minimum ni un maximum local.

Pour avoir une caractérisation des extrema locaux, il faut que la dérivée change de signe au point a , c'est à dire que f' soit monotone au voisinage de a . Cela se traduit, lorsque la fonction f est 2 fois dérivable par le fait que $f''(x)$ est de signe constant au voisinage de a .

I.1.7 Limites de suite de fonctions dérivables

Théorème I.7. Soit $(f_n)_{n \geq 0}$ une suite de fonctions à valeurs réelles, dérivables sur un intervalle I de $\mathbb{R}$. On suppose que :

- Il existe un $c \in I$ tel que la suite numérique $(f_n(c))_{n \geq 0}$ converge.
- La suite $(f'_n)_{n \geq 0}$ converge uniformément sur I .

Alors :

- Pour tout $x \in I$, la suite $(f_n(x))_{n \geq 0}$ converge.
- La fonction $x \mapsto \lim_{n \rightarrow \infty} f_n(x)$ est dérivable sur I et sa dérivée est donnée par la fonction $x \mapsto \lim_{n \rightarrow \infty} f'_n(x)$.

I.2 A propos d'une généralisation

Pour généraliser ces résultats à un cadre plus large voici quelques remarques sur les outils dont on va avoir besoin et (quelques unes des) difficultés que l'on va avoir à franchir.

Pour les fonctions réelles de la variables réelle, la définition du nombre dérivé à l'aide du quotient limite $\lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a}$ n'est pas très adaptée à une généralisation : dans des espaces plus gros que $\mathbb{R}$, quotienter par $x - a$ n'a pas de sens...

On va donc privilégier la notion de tangente. Au voisinage d'un point a , on peut approcher la fonction f par sa tangente :

$$f(x) = f(a) + f'(a)(x - a) + o(x - a).$$

La généralisation s'axe donc plutôt sur l'idée que l'on veut approcher une fonction par une fonction simple (qui serait l'équivalent de équation d'une droite) lorsqu'on se rapproche d'un point :

Une fonctions f sera donc "*différentiable en a* " s'il existe une "*fonction simple*" S telle que

$$f(x) = f(a) + S(x - a) + o(x - a)$$

lorsqu'on est "*proche de a* ". Cela implique déjà plein de choses :

1. Que signifie "être proche d'un point" ?

On doit pouvoir parler de distance entre 2 points et 2 images. On doit donc de travailler dans des espaces muni d'une distance (espaces métriques) ce qui permet également d'avoir accès à des notions tels que les ouverts, les fermés, les voisinages de points, etc...

2. Quelles fonctions simples choisir ?

Sur la représentation graphique d'une fonction dérivable f , au voisinage de a , la courbe vient s'écraser sur la droite tangente à f en a .

Si on donne de la profondeur au dessin (comme si on rajouter une autre variable...) la "nappe" obtenue vient s'écraser sur un plan...

Tout comme une fonction de $\mathbb{R}$ dans $\mathbb{R}$ dérivable en a est approchée par une droite passant par a au voisinage de a , une fonction de $\mathbb{R}^2$ dans $\mathbb{R}$ est différentiable en (a, b) devrait être approchée par l'équation d'un plan passant par $f(a, b)$ au voisinage de (a, b) .

l'équation d'une droite passant par $f(a)$ est toujours une application affine du type $f(a) + \lambda(x - a)$, la partie " $\lambda(x - a)$ " étant linéaire en l'accroissement $(x - a)$. De la même manière, l'équation d'un plan passant par $f(a, b)$ est toujours une application affine du type $f(a, b) + \alpha(x - a) + \beta(y - b)$, la partie " $\alpha(x - a) + \beta(y - b)$ " étant linéaire en l'accroissement $(x - a, y - b)$.

L'idée, c'est donc de dire qu'une fonction est différentiable en un point si elle est "presque" affine au voisinage de ce point, avec la même structure : $f(a) + S(x - a)$ où S est une application linéaire en l'accroissement $(x - a)$...

Pour chercher nos applications simples S , que l'on appellera **différentielles**, on va donc se tourner vers des espaces où on peut parler d'applications linéaires, c'est-à-dire les **espaces vectoriels**...

Dans les espaces vectoriels, les distances sympathiques sont en général issues d'une norme, que l'on utilise plus fréquemment dans les espaces vectoriels. Ainsi, Le cadre le plus acceptable est donc celui des **espaces vectoriels normés**, munis de la topologie associées à cette norme.

En ayant donc gommé les traces des objets propres à $\mathbb{R}$, on peut dire qu'une fonction f , d'un espace vectoriel normé vers un autre est dérivable en a si elle s'écrit $f(x) = f(a) + S(x - a) + o(\|x - a\|)$, où S est une application linéaire, que l'on appelle différentielle de f .

Oui, mais... une dernière petite chose :

Lorsqu'on est dans le cadre réel, on peut avoir l'existence d'une dérivée à gauche et d'une dérivée à droite d'un même point, sans que pour autant la fonction soit dérivable en ce point. Il y a donc une notion d'homogénéité de comportement : Pour une fonction f admettant une dérivée à droite $f'_d(a)$ et une dérivée à gauche $f'_g(a)$ mais non dérivable en a , on peut très bien écrire $f(x) = f(a) + S(x - a) + o(x - a)$ où S est une application linéaire par morceaux ($S(h) = f'_d(a)h$ ou $S(h) = f'_g(a)h$ selon si h est positif ou négatif) : selon d'où l'on vient, l'application ne se comporte pas de la même manière.

Ici, la fonction S n'est pas linéaire bien sûr, mais cela préfigure notamment un problème qui apparaît dans les espaces vectoriels normés de dimension infinie. On peut avoir des applications linéaires qui se comportent très bien selon certaines directions de l'espace, mais pas bien du tout selon d'autres, avec éventuellement avec des problèmes de discontinuité... On choisira donc de chercher nos différentielles parmi les **applications linéaires continues** pour éviter ces problèmes.

C'est aussi notamment pour cela qu'il n'est pas rare qu'on suppose également que nos evn soient complets, ce qui facilite aussi plein de petites choses ennuyeuses que l'on détaillera un peu plus tard, mais ce n'est pas obligatoire.

Cela dit, on aura souvent à faire à des **espaces vectoriels normés complets** (toute suite de Cauchy y converge), autrement appelés **espaces de Banach** (nom donné par Maurice Frechet en l'honneur de Stefan Banach, mathématicien polonais de la première moitié du XX^e siècle, un des fondateurs de l'analyse fonctionnelle qui a étudié ces espaces dans sa thèse). Au passage, les espaces vectoriels $\mathbb{R}^n$ et $\mathbb{C}^n$ sont des espaces de Banach et ce sont ceux qu'on utilisera le plus souvent...

Après ces quelques remarques, il est donc naturel de proposer la généralisation suivante :

On considère $f : E \rightarrow F$, où E et F sont des espaces vectoriels normés. On dira que f est **différentiable en** $a \in E$ s'il existe une application linéaire continue $L_a : E \rightarrow F$ telle que :

$$f(x) = f(a) + L_a(x - a) + o(\|x - a\|).$$

C'est le point de départ du calcul différentiel...

I.3 Rappels sur les espaces vectoriels normés...

I.3.1 Définitions

Définition I.8. Soit $\mathbb{K}$ un corps, typiquement $\mathbb{R}$ ou $\mathbb{C}$.

Un espace E est un **$\mathbb{K}$ -espace vectoriel** s'il est muni de :

- une loi d'addition interne "+" de E^2 dans E , appelée addition ou somme vectorielle,
- une loi de composition externe à gauche "." de $K \times E$ dans E , appelée multiplication par un scalaire,

Ces deux lois sont telles que :

1. $(E, +)$ est un groupe abélien,
2. La loi de composition externe à gauche vérifie :

— La loi “ $\cdot$ ” est distributive à gauche par rapport à la loi “ $+$ ” de E :

$$\lambda \cdot (u + v) = \lambda \cdot u + \lambda \cdot v,$$

— La loi “ $\cdot$ ” est distributive à droite par rapport à l’addition du corps K :

$$(\lambda_1 + \lambda_2) \cdot u = \lambda_1 \cdot u + \lambda_2 \cdot u,$$

— La loi “ $\cdot$ ” est associative à droite par rapport à la multiplication dans K :

$$(\lambda_1 \lambda_2) \cdot u = \lambda_1 \cdot (\lambda_2 \cdot u),$$

— L’élément neutre multiplicatif du corps $\mathbb{K}$, noté 1, est neutre à gauche pour la loi “ $\cdot$ ” :

$$1 \cdot u = u.$$

Définition I.9. Un $\mathbb{K}$ -espace vectoriel E est dit **normé** lorsqu’il est muni d’une norme, c’est-à-dire d’une application $N : E \rightarrow \mathbb{R}^+$ satisfaisant les hypothèses suivantes :

1. séparation : $\forall x \in E, N(x) = 0 \implies x = 0_E$,
2. homogénéité : $\forall (\lambda, x) \in \mathbb{K} \times E, N(\lambda \cdot x) = |\lambda|N(x)$,
3. sous-additivité : $\forall (x, y) \in E^2, N(x + y) \leq N(x) + N(y)$.

Remarque I.10. Le dernier point (sous-additivité) est en fait l’inégalité triangulaire, dont une conséquence immédiate est qu’on a : $\forall (x, y) \in E^2, N(x) - N(y) \leq N(x + y) \leq N(x) + N(y)$.

Usuellement on note la norme $\|\cdot\|$, avec éventuellement un indice.

Quelques exemples d’espaces vectoriels normés :

1. Les ensembles $\mathbb{R}^n$ et $\mathbb{C}^n$ sont des espaces vectoriels (de dimension finie) qui peuvent être munis de plusieurs normes, les plus classiques étant les normes $p : \|\cdot\|_p$ pour $p \in \mathbb{N}^*$, définies pour $x = (x_1, x_2, \dots, x_n)$ par :

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}},$$

ou encore, ils peuvent être munis de la norme infinie : $\|\cdot\|_\infty$, définie par $\|x\|_\infty = \sup_{i=1}^n |x_i|$. Notez que l’emploi de la notation $\|\cdot\|_\infty$ n’est pas anodin car $\lim_{p \rightarrow \infty} \|x\|_p = \|x\|_\infty$ quelque soit $x \in \mathbb{C}^n$.

2. L’ensemble $S_{00}(\mathbb{R})$ des suites réelles nulles à partir d’un certain rang est un espace vectoriel (de dimension infinie) que l’on peut munir de normes similaires à celles décrites ci-dessus : les normes $\|\cdot\|_p$ pour $p \in \mathbb{N}^*$, sont définies pour $x = (x_n)_{n \geq 0}$ par :

$$\|x\|_p = \left(\sum_{n \geq 0} |x_n|^p \right)^{\frac{1}{p}},$$

ou encore, ils peuvent être munis de la norme infinie $\|\cdot\|_\infty$, définie par $\|x\|_\infty = \sup_{n \geq 0} |x_n|$.

3. Un autre espace vectoriel normé de dimension infinie est l'ensemble $\mathcal{C}^0(I, \mathbb{C})$ des fonctions continues sur un intervalle $I \subset \mathbb{R}$ et à valeurs dans $\mathbb{C}$, peut être muni de la norme $\|\cdot\|_\infty$, définie par : $\|f\|_\infty = \sup_{x \in I} |f(x)|$.

Cette norme est dite de “de la convergence uniforme” car une suite de fonctions $(f_n)_{n \geq 0}$ de $\mathcal{C}^0(I, \mathbb{C})$ converge uniformément vers f si et seulement si $\|f_n - f\|_\infty$ tends vers 0.

I.3.2 Les outils...

I.3.2.1 Voisinages et parties connexes d'un evn

Soient a un élément d'un evn $(E, +, \cdot, \|\cdot\|)$, et $r \in \mathbb{R}^+$. on définit les ensembles suivants :

— **Boule ouverte centrée en a et de rayon r** :

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

— **Boule fermée centrée en a et de rayon r** :

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

Ce sont les ensembles du type $B(a, r)$ qui génèrent la topologie sur E .

On utilisera également la **sphère de centre a et de rayon r** :

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

Attention : la forme des boules ouvertes et fermées dépendent entièrement de la norme choisie. La topologie peut donc être modifiée. D'autre part, lorsque l'on travaille avec plusieurs normes indicées (ex : les normes $\|\cdot\|_p$ sur $\mathbb{R}^n$), on utilisera les mêmes indices pour les boules définie par les différentes normes. Par exemple :

- si $\mathbb{R}^2$ est muni de la norme 1, $B_1(0, 1)$ est le carré du plan délimité par les droites d'équations $y = \pm(1 - x)$ et $y = \pm(1 + x)$,
- si $\mathbb{R}^2$ est muni de la norme 2, $B_2(0, 1)$ est le disque centré en 0 et de rayon 1.
- si $\mathbb{R}^2$ est muni de la norme infinie, $B_\infty(0, 1)$ est le carré du plan délimité par les droites d'équations $y = \pm 1$ et $x = \pm 1$,

Remarque I.11. Dans le cadre des espaces vectoriels (et uniquement dans ce cadre!), on a les propriétés sympathiques suivantes :

- $B(a, r)$ est un ensemble ouvert,
- $\overline{B}(a, r)$ est un ensemble fermé,
- et si $r > 0$, alors $\overline{B(a, r)} = \overline{B}(a, r)$.

Attention : cela n'est pas valable sous d'autres hypothèses. Ex : Dans $\mathbb{Z}$ muni de la distance usuelle : $B(0, 2) = \{-1, 0, 1\} = \overline{B}(0, 1) \neq \overline{B}(0, 1) = \{0\}$. L'ensemble $B(0, 2)$ est, par ailleurs, à la fois ouvert et fermé...

Définition I.12. Soit E un evn et $a \in E$. On appelle **voisinage d'un point a** un sous-ensemble qui contient un ouvert contenant ce point (cette définition est en fait vraie dans les espaces topologiques quelconques).

On note $\mathcal{V}(a)$ l'ensemble des voisinages de a .

Exemple : Une boule ouverte ou fermée centrée en un point et de rayon strictement positif est un voisinage de ce point.

Propriétés I.13. Soit a un point d'un evn E .

1. $E \in \mathcal{V}(a)$ (donc $\mathcal{V}(a) \neq \emptyset$),
2. $A, B \in \mathcal{V}(a) \implies A \cap B \in \mathcal{V}(a)$
3. $A \in \mathcal{V}(a)$ et $A \subset B \subset E \implies B \in \mathcal{V}(a)$
4. $\emptyset \notin \mathcal{V}(a)$.

Ces propriétés indiquent en fait que l'ensemble des voisinages de a forment un filtre sur E pour l'inclusion.

Définition I.14. Soit E un evn. un sous-ensemble A de E est appelé **partie convexe** de E si :

$$\forall (x, y) \in A^2, \forall t \in [0, 1], tx + (1 - t)y \in A$$

Attention : Ne pas confondre conVexité et conNexité...

- La conVexité formalise la propriété suivante : Si 2 points sont dans A , alors tout le segment qui les relie est également dans A .
- La conNexité formalise le concept “d'objet fait d'un seul morceau” : Si 2 points sont dans A , alors il existe un chemin continu qui permet de les relier en restant toujours dans A . Ex : une étoile est conNexe, mais pas conVexe.

On peut aussi distinguer la connexité simple (“absence de trou”) et la connexité par arcs en utilisant le groupe d'homotopie mais c'est une autre histoire...

Propriété I.15. Si $\|\cdot\|$ est une norme sur un evn E , alors, pour tout $a \in E$ et tout $r > 0$, $B(a, r)$ est convexe.

I.3.2.2 Suites de Cauchy et espaces de Banach

Définitions I.16. Soit E un evn.

Une suite $(x_n)_{n \geq 0}$ d'éléments de E est dite **suite de Cauchy** si :

$$\forall \epsilon > 0, \exists n_\epsilon \forall i, j > n_\epsilon, \|x_i - x_j\| \leq \epsilon$$

Un espace métrique dans lequel tout suite de Cauchy converge est dit **complet**.

Un espace vectoriel normé complet est appelé **Espace de Banach**.

Remarque : Plus généralement, on peut définir les suites de Cauchy dans un espace métrique quelconque (E, d) , en remplaçant la norme $\|x_i - x_j\|$ par $d(x_i, x_j)$.

Attention : Un evn peut être complet lorsqu'il est muni d'une certaine norme mais ne plus l'être pour une autre norme...

Exemples I.17. Les $\mathbb{R}$ -ev $\mathbb{R}^n$ et $\mathbb{C}^n$, munis d'une des normes $\|\cdot\|_p$ ou de la norme $\|\cdot\|_\infty$ sont des espaces de Banach. Il sont de plus de dimension finie.

Du côté des espaces vectoriels de dimension infinie, on trouve par exemple $\mathcal{C}^0([0, 1], \mathbb{C})$ muni de sa norme infinie (présentée au paragraphe I.3.1 ou l'ensemble des suites réelles bornées muni de sa norme infinie parmi les espaces de Banach.

Remarque I.18. Si E et F sont 2 espaces de Banach, alors le produit $E \times F$ muni d'une norme p usuelle $\|(x, y)\|_{p, E \times F} = (\|x\|_E^p + \|y\|_F^p)^{\frac{1}{p}}$ ou de la norme infinie $\|(x, y)\|_{\infty, E \times F} = \sup \{\|x\|_E, \|y\|_F\}$ est aussi un espace de Banach.

Remarque I.19. D'une manière générale, lorsqu'on fait du calcul diff, tout ce qui est vrai dans $\mathbb{R}^n$ et qui ne fait pas intervenir explicitement une base) va être vrai dans un espace de Banach, même s'il y a quelques exceptions...

I.3.2.3 Continuité des applications d'un evn dans un autre et équivalence de normes

Définition I.20. Soit E et F deux evn et Ω un ouvert de E , contenant un point a .

Une fonction $f : E \rightarrow F$ est continue en a si :

$$\forall \epsilon > 0, \exists \eta_\epsilon > 0, \forall x \in B_E(a, \eta_\epsilon), f(x) \in B_F(f(a), \epsilon).$$

La continuité d'une application dépend donc intimement des normes choisies sur E et F via les boules ouvertes qui entrent en jeu. . Cela est ennuyeux car à priori, il n'y a aucune raison de choisir l'une ou l'autre norme. Pour chaque norme choisie, les boules ouvertes et fermées sont modifiées, et la topologie qu'elle génère aussi... Ainsi, une fonction peut être continue pour un couple de normes mais pas pour un autre..

Dans ce cas-là, y en a-t-il des normes à privilégier ? pas nécessairement...

En fait, la notion de continuité est identique lorsque les normes sont remplacées par d'autres normes qui leur sont équivalentes :

Définition I.21. Soit E un espace vectoriel sur lequel sont définies deux normes $\|\cdot\|_a$ et $\|\cdot\|_b$. On dit que ces deux normes sont des **normes équivalentes** si :

$$\exists (m, M) \in \mathbb{R}_+^*, \forall x \in E, m\|x\|_b \leq \|x\|_a \leq M\|x\|_b.$$

Théorème I.22.

Si E est un $\mathbb{R}$ -evn de dimension finie, alors toutes les normes sur E sont équivalentes.

Par exemple, sur $\mathbb{R}^2$, on a : $\|x\|_\infty \leq \|x\|_2 \leq \sqrt{2}\|x\|_\infty$ donc la norme 2 et la norme infinie sont équivalentes.

Notez que ce Théorème est ABSOLUMENT FAUX en dimension infinie (lorsqu'on considère par exemple des ensemble de suites ou des ensembles de fonctions)...

Conséquence importante : Les evn $\mathbb{R}^n$ et $\mathbb{C}^n$ font partie des evn de dimension finie... Sur ces espaces, les normes sont donc toutes équivalentes. Ainsi :

Lorsque les applications sont définies et à valeurs dans $\mathbb{R}^n$ ou $\mathbb{C}^n$, la notion de continuité est indépendante de la norme choisie.

Remarque : Si un evn E quelconque muni d'une norme $\|\cdot\|$ est complet, alors il est complet lorsqu'on le muni de toute norme équivalente à $\|\cdot\|$. Ainsi, on a la caractérisation suivante : Un evn de dimension finie est un espace de Banach si et seulement s'il est complet pour une seule norme.

I.4 Applications linéaires et multilinéaires

I.4.1 Définitions

Soit $\mathbb{K} = \mathbb{R}$ ou $\mathbb{C}$.

Définition I.23. Soient E et F deux $\mathbb{K}$ -ev.

Une application $L : E \rightarrow F$ est une **application linéaire** si elle satisfait les deux propriétés suivantes :

- $\forall (x, y) \in E^2, L(x + y) = L(x) + L(y)$,
- $\forall x \in E, \forall \lambda \in \mathbb{K}, L(\lambda x) = \lambda L(x)$.

Une application linéaire est en fait un morphisme d'espaces vectoriels.

On note $L(E, F)$ l'ensemble des **applications linéaires de E dans F** .

On rencontrera dans la suite du cours des applications linéaires bijectives de E dans F , appelées aussi **isomorphismes**, dont l'ensemble sera noté $\mathcal{I}som(E, F)$.

Remarque : L'ensemble des applications linéaires de E dans F est également un $\mathbb{K}$ -espace vectoriel.

Définition I.24. Soit un entier $n \geq 1$ et soient $E_1, E_2, \dots, E_n$ et F des $\mathbb{K}$ -ev.

Une application $L : E_1 \times E_2 \times \dots \times E_n \rightarrow F$ est une **application n -linéaire** si elle est linéaire en chacune de ses variables.

C'est-à-dire, quelque soient $i \in \{1, 2, \dots, n\}$ et $a = (a_1, a_2, \dots, a_n) \in E_1 \times E_2 \times \dots \times E_n$, l'application :

$$\begin{aligned} L_j^a : E_j &\rightarrow F \\ x &\mapsto L(a_1, a_2, \dots, a_{j-1}, a_j + x, a_{j+1}, \dots, a_n) \end{aligned}$$

est une application linéaire.

Une application n -linéaire pour un certain entier n est dite aussi multilinéaire.

Remarque :

Lorsque $n = 1$, cette définition coïncide avec la définition d'une application linéaire.

Lorsque $n = 2$, on parle souvent de **bilinéarité** au lieu de 2-linéarité. Voici quelques exemples d'applications bilinéaires :

- l'application $L_1 : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ définie par $L(x, y) = xy$,
- l'application $L_2 : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par $L((h_1, h_2), (k_1, k_2)) = \alpha h_1 k_1 + \beta h_2 k_2$ où α et β sont des réels quelconques,
- l'application $L_2 : S_{00}(\mathbb{R}) \times S_{00}(\mathbb{R}) \rightarrow \mathbb{R}$ définie par $L((u_n)_{n \geq 0}, (v_n)_{n \geq 0}) = \sum_{n \geq 0} \alpha_n u_n v_n$, pour une certaine suite réelle $(\alpha_n)_{n \geq 0}$.

I.4.1.1 Continuité et norme d'un application linéaire

Proposition I.25. Soient E et F deux $\mathbb{K}$ -evn et $L : E \rightarrow F$ une application linéaire.

Les conditions suivantes sont équivalentes :

1. L est continue en tout point de E ,
2. L est continue en 0_E ,
3. Il existe un réel $M \geq 0$ tel que, pour tout $x \in E$, $\|L(x)\|_F \leq M\|x\|_E$,

4. L est bornée sur $B_E(0_E, 1)$,
5. L est bornée sur $S_E(0_E, 1)$,
6. L est bornée sur une partie bornée de E .

On note $\mathcal{L}(E; F)$ l'ensemble des applications linéaires continues de E dans F .

Remarque : Une application linéaire continue est en fait lipschitzienne et elle est également uniformément continue.

Définition I.26. Soient E et F deux $\mathbb{K}$ -evn et $L : E \rightarrow F$ une application linéaire. On appelle **norme de L** la quantité $\|L\| = \sup_{x \in E} \frac{\|L(x)\|_F}{\|x\|_E}$.

On obtient très facilement les égalités suivantes :

$$\|L\| = \sup_{x \in E} \frac{\|L(x)\|_F}{\|x\|_E} = \sup_{\|x\|_E=1} \|L(x)\|_F = \sup_{\|x\|_E \leq 1} \|L(x)\|_F = \sup_{\|x\|_E \leq r} \frac{\|L(x)\|_F}{\|x\|_E} = \sup_{\|x\|_E=r} \frac{\|L(x)\|_F}{r},$$

où r est un quelconque réel strictement positif.

Remarque : L'emploi du terme norme est justifié car $\|\cdot\|$ est bien une norme sur $\mathcal{L}(E; F)$.

Proposition I.27. Soient E et F deux $\mathbb{K}$ -evn et $L : E \rightarrow F$ une application linéaire. L est continue si et seulement si $\|L\|$ est finie.

Proposition I.28. cas des ev de dimension finie.

Soient E et F deux $\mathbb{K}$ -evn.

Si E est un ev de dimension finie, alors toute application linéaire $L : E \rightarrow F$ est continue.

Preuve : Comme toutes les normes sur E sont équivalentes (grâce à la dimension finie), il suffit de montrer la continuité de L pour une seule des normes sur E .

On choisit une base $(e_1, e_2, \dots, e_n)$ de E . Pour un élément $x = x_1 e_1 + x_2 e_2 + \dots + x_n e_n$ de E , on munit E de sa norme 1 relative à cette base : $\|x\|_1 = \sum_{i=1}^n |x_i|$. On a :

$$\|L(x)\|_F = \|x_1 L(e_1) + x_2 L(e_2) + \dots + x_n L(e_n)\|_F \leq \|x_1 L(e_1)\|_F + \|x_2 L(e_2)\|_F + \dots + \|x_n L(e_n)\|_F.$$

On en déduit donc que $\|L(x)\|_F \leq |x_1| \|L(e_1)\|_F + |x_2| \|L(e_2)\|_F + \dots + |x_n| \|L(e_n)\|_F$ et ainsi :

$$\|L(x)\|_F \leq \sup_{i=1}^n \|L(e_i)\|_F \sum_{i=1}^n |x_i| = \sup_{i=1}^n \|L(e_i)\|_F \|x\|_1,$$

donc la norme de L est finie et L est donc continue.

Exercice I.29. Soit E l'ensemble des suites réelles bornées, muni de la norme infinie $\|(u_n)_{n \geq 0}\| = \sup_{n \geq 0} |u_n|$. On considère les applications $S : E \rightarrow E$ et $\Delta : E \rightarrow E$ définies de la manière suivante : $S((u_n)_{n \geq 0}) = (u_{n+1})_{n \geq 0}$ et $\Delta((u_n)_{n \geq 0}) = (u_{n+1} - u_n)_{n \geq 0}$. Montrer que S et Δ sont des applications linéaires et donnez leur normes.

Exercice I.30. Soit $E = \mathbb{R}[X]$ l'ensemble des polynômes à coefficients réels muni de la norme $\|P\| = \sup_{x \in [0,1]} |P(x)|$. On considère l'application :

$$\begin{aligned} \Phi : E &\rightarrow \mathbb{R} \\ P &\mapsto P(3). \end{aligned}$$

1. Montrer que Φ est une application linéaire continue.
2. Soit $P_n \in E$ défini par $P(X) = \frac{1}{2^n} X^n$ pour tout $n \geq 1$.
 - (a) Montrer que $\lim_{n \rightarrow \infty} \|P_n\| = 0$ et que $\lim_{n \rightarrow \infty} |\Phi(P_n)| = +\infty$.
 - (b) En déduire que Φ n'est pas continue.

I.4.2 Continuité et norme d'une application multilinéaire

Proposition I.31. Soit un entier $n \geq 1$ et soient $E_1, E_2, \dots, E_n$ et F des $\mathbb{K}$ -ev.

Soit $L : E_1 \times E_2 \times \dots \times E_n \rightarrow F$ une application n -linéaire.

Les conditions suivantes sont équivalentes :

1. L est continue en tout point de $E_1 \times E_2 \times \dots \times E_n$,
2. L est continue en $0_{E_1 \times E_2 \times \dots \times E_n}$,
3. Il existe un réel $M \geq 0$ tel que, pour tout $x = (x_1, x_2, \dots, x_n) \in E_1 \times E_2 \times \dots \times E_n$,
 $\|L(x)\|_F \leq M \|x_1\|_{E_1} \|x_2\|_{E_2} \dots \|x_n\|_{E_n}$,

On note $\mathcal{L}(E_1, E_2, \dots, E_n; F)$ l'ensemble des applications n -linéaires continues de $E_1 \times E_2 \times \dots \times E_n$ dans F .

Définition I.32. Soit un entier $n \geq 1$, soient $E_1, E_2, \dots, E_n$ et F des $\mathbb{K}$ -ev et soit $L : E_1 \times E_2 \times \dots \times E_n \rightarrow F$ une application n -linéaire.

On appelle **norme de L** la quantité

$$\|L\| = \sup_{x \in E_1 \times E_2 \times \dots \times E_n} \frac{\|L(x)\|_F}{\|x_1\|_{E_1} \|x_2\|_{E_2} \dots \|x_n\|_{E_n}}.$$

En notant $\|x\| = \prod_{i=1}^n \|x_i\|_{E_i}$ la norme produit, on obtient très facilement les égalités suivantes :

$$\|L\| = \sup_{x \in E_1 \times E_2 \times \dots \times E_n} \frac{\|L(x)\|_F}{\|x\|} = \sup_{\|x\|=1} \|L(x)\|_F = \sup_{\|x\| \leq 1} \|L(x)\|_F = \sup_{\|x\| \leq r} \frac{\|L(x)\|_F}{\|x\|} = \sup_{\|x\|=r} \frac{\|L(x)\|_F}{r},$$

où r est un quelconque réel strictement positif.

Remarque : L'emploi du terme norme est justifié car $\|\cdot\|$ est bien une norme sur $\mathcal{L}(E_1, E_2, \dots, E_n; F)$.

Proposition I.33. Soient E et F deux $\mathbb{K}$ -evn et $L : E_1 \times E_2 \times \dots \times E_n \rightarrow F$ une application n -linéaire. L est continue si et seulement si $\|L\|$ est finie.

Proposition I.34. cas des ev de dimension finie.

Soit un entier $n \geq 1$ et soient $E_1, E_2, \dots, E_n$ et F des $\mathbb{K}$ -ev. Si tous les E_i sont des evn de dimension finie, alors :

Toute application linéaire $L : E_1 \times E_2 \times \dots \times E_n \rightarrow F$ est continue.

Chapitre II

Applications différentiables

Dans tout ce chapitre et les suivants $\mathbb{K}$ désigne le corps $\mathbb{R}$ ou $\mathbb{C}$.

II.1 Différentiabilité

II.1.1 Fonctions tangentes

Définition II.1. Soient E et F deux $\mathbb{K}$ -evn, Ω un ouvert de E , $a \in \Omega$ et $n \geq 1$ un entier.

On dit que $f : E \rightarrow F$ est **tangente à 0_F à l'ordre n en a** si $f(x) = o(\|x - a\|_E^n)$, c'est-à-dire que :

$$\forall \epsilon > 0, \exists \eta > 0, \|x - a\|_E < \eta \Rightarrow \|f(x)\|_F \leq \epsilon \|x - a\|_E^n.$$

Lorsque $n = 1$, on dit simplement que f est tangente à 0_F en a .

Deux fonctions $f : E \rightarrow F$ et $g : E \rightarrow F$ sont **tangentes à l'ordre n en a** si $(f - g)$ est tangente à 0 à l'ordre n en a .

Propriétés II.2.

1. *Expression à l'aide limite* : $f : E \rightarrow F$ est tangente à 0_F à l'ordre n en a peut aussi s'écrire :

$$\lim_{x \rightarrow a} \frac{1}{\|x - a\|_E^n} f(x) = 0_F.$$

2. *Addition et multiplication par un scalaire* : Si $f : E \rightarrow F$ et $g : E \rightarrow F$ sont tangente à 0 à l'ordre n en a et que $\lambda \in \mathbb{K}$, alors les fonctions $f + g$ et λf sont également tangente à 0 à l'ordre n en a .
3. *Relation d'équivalence* : La relation entre les fonctions de Ω dans F "être tangente à l'ordre n en a " est une relation d'équivalence.
4. *Changement de normes sur E et F* : Si on remplace sur E et F les normes par des normes qui leur sont équivalents, on ne modifie pas la notion de tangence :
Deux fonctions tangentes à l'ordre n en a pour les anciennes normes le seront toujours pour les nouvelles.

5. *Comparaison des ordres* : Si $f : E \rightarrow F$ et $g : E \rightarrow F$ sont tangentes à l'ordre n en a , alors elle le sont encore à l'ordre m pour tout $m \leq n$.
6. *Tangence et continuité* : Si $f : E \rightarrow F$ est tangente à 0_F à l'ordre n en a , alors f est continue en a et $f(a) = 0_F$.
Par conséquent, si $f : E \rightarrow F$ et $g : E \rightarrow F$ sont tangentes à l'ordre n en a et que f est continue en a , alors g l'est aussi.
7. *Cas des applications linéaires* : Si $L : \Omega \rightarrow F$ est une application linéaire tangente à 0_F à l'ordre n en 0_E , alors L est nulle.

Exercice II.3. Montrer les affirmations des points 4. et 7.

II.1.2 Différentielle d'une application

Définition II.4. Soient E et F deux $\mathbb{K}$ -evn, Ω un ouvert de E . et

On dit qu'une application $f : \Omega \rightarrow F$ est **différentiable en** $a \in \Omega$ s'il existe une application linéaire continue $L_a : E \rightarrow F$ telles que :

$$f(x) = f(a) + L_a(x - a) + o(\|x - a\|_E). \quad (\text{II.1})$$

Autrement dit, f et l'application affine $x \mapsto f(a) + L_a(x - a)$ sont tangentes en a .

Dans ce cas, l'application linéaire continue L_a est appelée **différentielle de f en a** , parfois appelée **différentielle au sens de Fréchet**.

On trouve différentes notations dans la littérature pour la différentielle de f en a . En voici quelques unes :

$$Df(a), Df_a, \partial f(a), \partial f_a, df(a), df_a, f'(a), \dots$$

Remarque II.5. En fait, si on réécrit l'équation II.1 comme suit :

$$f(a + h) - f(a) = L_a(h) + o(\|h\|_E),$$

On peut interpréter cette égalité ainsi : au voisinage de a , l'accroissement de la fonction se comporte comme un terme linéaire par rapport à l'accroissement des variables, modulo un terme correctif négligeable.

A ce propos, puisque la notion même de différentiabilité en un point utilise un voisinage de celui-ci, tous les résultats de différentiabilité considèreront des fonctions définies sur des ouverts, pour s'assurer que la fonction considérée est bien définie sur un voisinage de tous les points de son ensemble de définition.

Remarque II.6. Comparaison avec le cas des fonctions réelles de la variables réelle Si $f : \mathbb{R} \rightarrow \mathbb{R}$, alors f est différentiable en $a \in \mathbb{R}$ si et seulement si elle est dérivable au sens usuel.

De ce fait, dans certains livres (notamment la bible de Cartan), on parle aussi de dérivabilité et de dérivée (au lieu d'employer le terme différentiabilité et différentielles) même dans des espaces plus gros que $\mathbb{R}$.

Ce qui peut perturber un peu, c'est que dans le cadre plus général des fonctions de E dans F , ce n'est pas tout à fait la notion du "nombre dérivé" que l'on exporte...

Si $f : \Omega \rightarrow \mathbb{R}$ est dérivable en a avec $\Omega \subset \mathbb{R}$, le nombre dérivé $f'(a)$ apparaît dans la différentielle de f en a de la façon suivante :

$$Df(a) : \begin{array}{ccc} \mathbb{R} & \rightarrow & \mathbb{R} \\ h & \mapsto & f'(a) \cdot h, \end{array}$$

En fait, $Df(a)$ est une application linéaire de $\mathcal{L}(\mathbb{R}, \mathbb{R})$. L'utilisation massive du nombre dérivé (plutôt que de la différentielle) découle du fait que $\mathcal{L}(\mathbb{R}, \mathbb{R})$ est isomorphe à $\mathbb{R}$: se donner d'une application linéaire continue de $\mathbb{R}$ dans $\mathbb{R}$: $Df(a)$, c'est en fait donner sa pente : $f'(a)$.

A titre indicatif (en attendant la suite et les histoires de matrices jacobiniennes pour les fonctions de $\mathbb{R}^n$ dans $\mathbb{R}^p$ au paragraphe II.3.3...), ce qui étend la notion de nombre dérivée pour des fonctions de $\mathbb{R}$ dans $\mathbb{R}^n$, c'est la notion de gradient de f en a (voir Paragraphe II.2.4).

Définition II.7. Soient E et F deux $\mathbb{K}$ -evn, Ω un ouvert de E .

On dit qu'une application $f : \Omega \rightarrow F$ est **différentiable sur** Ω (ou seulement **différentiable** si aucune confusion n'est possible) si elle l'est en tout point de Ω .

Lorsque l'application

$$Df : \begin{array}{ccc} \Omega & \rightarrow & \mathcal{L}(E; F) \\ x & \mapsto & Df(x) \end{array}$$

est continue, on dit que f est **continuellement différentiable sur** Ω ou **de classe** $\mathcal{C}^1$. L'ensemble des applications continuellement différentiables sur Ω est noté $\mathcal{C}^1(\Omega, F)$ (ou seulement $\mathcal{C}^1(\Omega)$ si aucune confusion n'est possible).

Remarque II.8. Dans le cas des fonctions de $\mathbb{R}$ dans $\mathbb{R}$ la continuité de Df est équivalente à la continuité de l'application $x \mapsto f'(x)$ puisque $f'(x)$ détermine entièrement $Df(x)$.

Propriétés II.9.

1. *Différentiabilité et continuité* : Si une application $f : \Omega \rightarrow F$ est différentiable en $a \in \Omega$, alors elle est continue en a .
2. *Unicité de la différentielle* : Si $f : \Omega \rightarrow F$ est différentiable en $a \in \Omega$, sa différentielle en a est unique.
3. *Changement de normes* : La différentiabilité d'une application $f : \Omega \rightarrow F$ en un point $a \in \Omega$ n'est pas affectée par un changement des normes sur E ou F par des normes équivalentes.

Justifications :

1. L'application f est équivalente à une application linéaire continue en a , (puisque elle lui est tangente en a) elle est donc elle également continue en a .
2. En supposant qu'il existe deux applications linéaires continues $L_1 : E \rightarrow F$ et $L_2 : E \rightarrow F$ telles que $f(x) = f(a) + L_1(x - a) + o(\|x - a\|_E)$ et $f(x) = f(a) + L_2(x - a) + o(\|x - a\|_E)$, on a alors : $L_1(x - a) - L_2(x - a) = o(\|x - a\|_E)$. Ainsi l'application $L_1 - L_2$ est tangente à zéro en a . D'après la propriété II.2.7, $L_1 - L_2$ est donc nulle.
3. C'est une conséquence directe de la propriété II.2.4.

II.1.3 Premiers exemples simples

Exercice II.10. On munit $\mathbb{R}^2$ de la norme euclidienne $\|\cdot\|_2$. On considère la fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par $f(x, y) = x^2 + xy + y^2$.

1. Soient $(x, y) \in \mathbb{R}^2$ et $(h, k) \in \mathbb{R}^2$. Calculer $f(x + h, y + k) - f(x, y)$.
2. On définit, pour tout couple $(x, y) \in \mathbb{R}^2$ fixé, l'application

$$\begin{aligned} L_{(x,y)} : \quad \mathbb{R}^2 &\rightarrow \mathbb{R} \\ (h, k) &\mapsto \langle (2x + y, 2y + x), (h, k) \rangle. \end{aligned}$$

où $\langle \cdot, \cdot \rangle$ représente le produit scalaire usuel sur $\mathbb{R}^2$.

Montrer que $L_{(x,y)}$ est une application linéaire continue et calculer sa norme.

3. Montrer que $L_{(x,y)} = Df(x, y)$ pour tout $(x, y) \in \mathbb{R}^2$.

Exercice II.11. Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par $f(x, y) = \frac{x}{y^2}$. Justifier l'existence et calculer la différentielle de f au point $(1, 1)$.

Indication : On pourra faire un développement limité de $f(1+h, 1+k)$ lorsque $\|(h, k)\| = \sup(|h|, |k|)$ tend vers 0.

Remarque : La méthode développée dans cet exercice est classique pour trouver la différentielle sans utiliser les dérivées partielles.

Exercice II.12. (Fonctions vectorielles de la variable réelle)

Soient n applications $f_i : \mathbb{R} \rightarrow \mathbb{R}$ dérivables en un point $a \in \mathbb{R}$ et $f : \mathbb{R} \rightarrow \mathbb{R}^n$ l'application définie par $f(x) = (f_1(x), f_2(x), \dots, f_n(x))$.

1. Montrer que $f(x) - f(a) = (f'_1(a), f'_2(a), \dots, f'_n(a)) \cdot (x - a) + o(x - a)$.
2. En déduire que f est différentiable en a et donner sa différentielle $Df(a)$ au point a . Attention : $Df(a) \in \mathcal{L}(\mathbb{R}, \mathbb{R}^n)$

Pour les fonctions de $\mathbb{R}^n$ dans $\mathbb{R}^m$, on aimerait bien généraliser ce procédé pour calculer les différentielles... D'où la question : et si c'était aussi simple que pour les fonctions de $\mathbb{R}$ dans $\mathbb{R}^n$? On regarde ce qui se passe dans les différentes composantes lorsqu'on dérive par rapport à chacune des variables et on additionne le tout ?

Pas tout à fait... mais quand même un peu, enfin presque... L'existence de dérivées par rapport aux différentes coordonnées n'est pas tout à fait la même chose que la différentiabilité... on verra cela à la section II.3.4.

II.2 Règles de calcul

Les règles de calculs présentées dans ce paragraphe permettent en pratique de calculer les différentielles de fonctions "traditionnelles".

II.2.1 Linéarité

Proposition II.13. Soient E et F deux $\mathbb{K}$ -evn, Ω un ouvert de E et un point $a \in \Omega$.

Si on note $\mathcal{D}(\Omega; F; a)$ l'ensemble des applications $f : \Omega \rightarrow F$ différentiables en a et $\mathcal{D}(\Omega; F)$ l'ensemble des applications $f : \Omega \rightarrow F$ différentiables sur Ω , alors :

1. $\mathcal{D}(\Omega; F; a)$ et $\mathcal{D}(\Omega; F)$ sont des espaces vectoriels,
2. Les applications :

$$\begin{aligned} D \cdot (a) : \mathcal{D}(\Omega; F; a) &\rightarrow \mathcal{L}(E; F) \\ f &\mapsto Df(a) \end{aligned}$$

et

$$\begin{aligned} D : \mathcal{D}(\Omega; F) &\rightarrow \mathcal{L}(E; F) \\ f &\mapsto Df \end{aligned}$$

sont linéaires.

Autrement dit :

Si $f : \Omega \rightarrow F$ et $g : \Omega \rightarrow F$ sont deux applications différentiables en $a \in \Omega$ et si $\lambda \in \mathbb{K}$ alors $f + g$ et λf sont différentiables en $a \in \Omega$ et on a :

$$D(f + g)(a) = D(f)(a) + D(g)(a), \quad \text{et} \quad D(\lambda f)(a) = \lambda D(f)(a).$$

Démonstration en exercice.

II.2.2 Différentielles d'applications particulières

Propriétés II.14. Soient E , E_1 , E_2 et F des $\mathbb{K}$ -evn.

1. **Applications constantes :**

Toute application constante est différentiable et sa différentielle est nulle.

2. **Applications linéaires continues :**

Toute application linéaire continue $L \in \mathcal{L}(E; F)$ est différentiable sur E et sa différentielle en tout point est L elle-même.

3. **Applications bilinéaires continues :**

Une application bilinéaire continue $B \in \mathcal{L}(E_1, E_2; F)$ est différentiable sur $E_1 \times E_2$. Sa différentielle au point $(a_1, a_2) \in E_1 \times E_2$ est l'application linéaire $DB(a_1, a_2) \in \mathcal{L}(E_1 \times E_2; F)$ définie par :

$$DB(a_1, a_2) : (h_1, h_2) \mapsto B(h_1, a_2) + B(a_1, h_2).$$

Justifications :

2. Puisque f est linéaire, on a $L(x) - L(a) = L(x - a) + 0$.

3. D'abord, il y a un non-dit dans l'énoncé : Pour pouvoir parler de continuité, il faut avoir fixé une norme sur $E_1 \times E_2$. On peut fixer par exemple $\|(x_1, x_2)\|_{E_1 \times E_2} = \|x_1\|_{E_1} + \|x_2\|_{E_2}$ (ou encore $\|(x_1, x_2)\|_{E_1 \times E_2} = \sup(\|x_1\|_{E_1}, \|x_2\|_{E_2})$ puisqu'elles sont équivalentes).

Puisque l'application $L : (h_1, h_2) \mapsto B(a_1, h_2) + B(h_1, a_2)$ est clairement linéaire sur $E_1 \times E_2$. Montrons qu'elle est continue. Puisque B est bilinéaire continue, on a :

$$\|B(a_1, h_2) + B(h_1, a_2)\|_F \leq \|B\| \|a_1\|_{E_1} \|h_2\|_{E_2} + \|B\| \|h_1\|_{E_1} \|a_2\|_{E_2}.$$

Ainsi, on a :

$$\begin{aligned}\|B(a_1, h_2) + B(h_1, a_2)\|_F &\leq \|B\| (\|a_1\|_{E_1} \|h_2\|_{E_2} + \|h_1\|_{E_1} \|a_2\|_{E_2}), \\ \|B(a_1, h_2) + B(h_1, a_2)\|_F &\leq \|B\| \sup(\|a_1\|_{E_1}, \|a_2\|_{E_2}) (\|h_1\|_{E_1} + \|h_2\|_{E_2}),\end{aligned}$$

et donc :

$$\|B(a_1, h_2) + B(h_1, a_2)\|_F \leq \|B\| \sup(\|a_1\|_{E_1}, \|a_2\|_{E_2}) \|(h_1, h_2)\|_{E_1 \times E_2}.$$

L'application $L : (h_1, h_2) \mapsto B(a_1, h_2) + B(h_1, a_2)$ est donc linéaire et continue. Il reste donc à montrer que $B(a_1 + h_1, a_2 + h_2) - B(a_1, a_2) = L(h_1, h_2) + o(\|(h_1, h_2)\|_{E_1 \times E_2})$. Or, par bilinéarité, on a :

$$B(a_1 + h_1, a_2 + h_2) = B(a_1, a_2) + B(h_1, a_2) + B(a_1, h_2) + B(h_1, h_2),$$

donc

$$B(a_1 + h_1, a_2 + h_2) - B(a_1, a_2) = L(h_1, h_2) + B(h_1, h_2).$$

Or B étant bilinéaire et continue, on a

$$\|B(h_1, h_2)\|_F \leq \|B\| \|h_1\|_{E_1} \|h_2\|_{E_2} \leq \|B\| \|(h_1, h_2)\|_{E_1 \times E_2}^2 = o(\|(h_1, h_2)\|_{E_1 \times E_2}),$$

ce qui conclut la preuve.

Notez que cette dernière propriété se généralise aux applications multilinéaires continue (la preuve étant sensiblement la même) :

Propriété II.15. Applications multilinéaires continues

Soient $E_1, E_2, \dots, E_n$ et F des $\mathbb{K}$ -evn.

Une application bilinéaire continue $B \in \mathcal{L}(E_1, E_2, \dots, E_n; F)$ est différentiable sur $E_1 \times E_2 \times \dots \times E_n$. Sa différentielle au point $a = (a_1, a_2, \dots, a_n) \in E_1 \times E_2 \times \dots \times E_n$ est l'application linéaire $DB(a) \in \mathcal{L}(E_1 \times E_2 \times \dots \times E_n; F)$ définie par :

$$DB(a) : (h_1, h_2, \dots, h_n) \mapsto B(h_1, a_2, \dots, a_n) + B(a_1, h_2, \dots, a_n) + \dots + B(a_1, a_2, \dots, a_{n-1}, h_n).$$

II.2.3 Différentielle d'une application composée

Proposition II.16. Soient E, F et G trois $\mathbb{K}$ -evn, Ω un ouvert de E , $\mathcal{V}$ un ouvert de F .

Soient $f : \Omega \rightarrow F$ et $g : \mathcal{V} \rightarrow G$ deux applications et $a \in \Omega$ tel que $f(a) \in \mathcal{V}$.

Si f est différentiable en a et que g l'est aussi en $f(a)$, alors l'application $g \circ f$ est différentiable en a et sa différentielle en ce point est donnée par :

$$D(g \circ f)(a) = Dg(f(a)) \circ Df(a).$$

Remarque : Cette formule généralise bien la formule de la dérivée d'une fonction composée de $\mathbb{R}$ dans $\mathbb{R}$: $(g \circ f)'(x) = f'(x)g'(f(x))\dots$ ou plutôt

$$(g \circ f)'(x) = g'(f(x)) \times f'(x).$$

En effet, la composition de 2 application linéaires de $\mathbb{R}$ dans $\mathbb{R}$ $L_1 : x \mapsto \lambda_1 x$ et $L_2 : x \mapsto \lambda_2 x$ est une opération commutative et elle se traduit par la multiplication des coefficients : $L_1 \circ L_2(x) = L_2 \circ L_1(x) = \lambda_1 \lambda_2 x$. Ce ne sera plus le cas en dimension supérieure !

Preuve : On fixe $b = f(a)$ et, pour tout $x \in \Omega$ et $y \in \mathcal{V}$, on pose

$$\Delta_a^f(x) = f(x) - b - Df(a)(x - a) \quad \text{et} \quad \Delta_b^g(y) = g(y) - g(b) - Dg(b)(y - b),$$

de sorte que $\Delta_a^f(x) = o(\|x - a\|_E)$ et $\Delta_b^g(y) = o(\|y - b\|_F)$. On obtient facilement l'égalité suivante :

$$g(f(x)) - g(b) - Dg(b) \circ Df(a)(x - a) = \Delta_b^g(f(x)) + Dg(b) \left(\Delta_a^f(x) \right).$$

Or $Dg(b)$ est une application linéaire continue, donc

$$\left\| Dg(b) \left(\Delta_a^f(x) \right) \right\| \leq \|Dg(b)\| \left\| \Delta_a^f(x) \right\|.$$

ce qui implique que $Dg(b) \left(\Delta_a^f(x) \right) = o(\|x - a\|)$ puisque $\Delta_a^f(x) = o(\|x - a\|)$.

Il reste donc à montrer que $\Delta_b^g(f(x)) = o(\|x - a\|)$ pour pouvoir conclure la preuve.

On sait que $\Delta_b^g(f(x)) = o(\|f(x) - b\|)$, or $f(x) - b = Df(a)(x - a) + \Delta_a^f(x)$. Ainsi,

$$\|f(x) - b\| \leq \|Df(a)\| \|x - a\| + \left\| \Delta_a^f(x) \right\|.$$

On a donc $f(x) - b = O(\|x - a\|)$ Puisque $\Delta_a^f(x) = o(\|x - a\|)$.

Puisqu'on a, à la fois $\Delta_b^g(f(x)) = o(\|f(x) - b\|)$ et $f(x) - b = O(\|x - a\|)$, on a donc $\Delta_b^g(f(x)) = o(\|x - a\|)$, ce qui termine la preuve.

On a également une version globale de cette formule :

Proposition II.17. Soient E, F et G trois $\mathbb{K}$ -evn, Ω un ouvert de E , $\mathcal{V}$ un ouvert de F .

Soient $f : \Omega \rightarrow F$ et $g : \mathcal{V} \rightarrow G$ deux applications.

Si f est différentiable (resp. $\mathcal{C}^1$) sur Ω et que g l'est aussi sur $\mathcal{V}$, alors l'application $g \circ f$ est différentiable (resp. $\mathcal{C}^1$) sur son domaine de définition : $\Omega \cap f^{-1}(\mathcal{V})$.

II.2.4 Différentielle d'une application à valeurs dans un espace vectoriel produit

Les résultats présentés ici sont sous une forme générale mais dans la pratique, les espaces produits que l'on rencontrera dans les exos seront rarement autre chose que $\mathbb{R}^n$ ou $\mathbb{C}^n$

Définitions II.18. Soient E et F deux $\mathbb{K}$ -evn et Ω un ouvert de E . On suppose (ici et dans la suite du paragraphe) que $F = F_1 \times F_2$ est un $\mathbb{K}$ -evn produit, que l'on muni d'une norme équivalente à la norme sup : $\|(y_1, y_2)\|_\infty = \sup(\|y_1\|_{F_1}, \|y_2\|_{F_2})$.

Les **projections de l'espace produit sur ses facteurs** $p_1 : F \rightarrow F_1$ et $p_2 : F \rightarrow F_2$ sont définies par $p_1((x, y)) = x$ et $p_2((x, y)) = y$.

Les **injections des facteurs dans l'espace produit** $u_1 : F_1 \rightarrow F$ et $u_2 : F_2 \rightarrow F$ sont définies par $u_1(x) = (x, 0)$ et $u_2(y) = (0, y)$.

On appelle **composantes d'une application** $f : E \rightarrow F_1 \times F_2$, les applications $f_i : E \rightarrow F_i$ définies par $f_i = p_i \circ f$.

On a alors : $f = u_1 \circ f_1 + u_2 \circ f_2$.

Remarques II.19. On a, avec les mêmes notations que dans les définitions précédentes :

1. Les applications p_i et u_i sont linéaires et continues.
2. $p_i \circ u_i = 1_{F_i}$ et $u_1 \circ p_1 + u_2 \circ p_2 = 1_F$.
3. L'espace $\mathcal{L}(E; F)$ est isomorphe à $\mathcal{L}(E; F_1) \times \mathcal{L}(E; F_2)$ (l'isomorphisme a en plus le bon gout de préserver la norme) :
Une application linéaire continue de E dans F se décompose en une application linéaire continue de E dans F_1 et une application linéaire continue de E dans F_2 .
4. Les définitions et les propriétés données précédemment s'étendent très facilement au produit de plus de 2 evn.

Proposition II.20. Soient E, F_1 et F_2 trois $\mathbb{K}$ -evn et Ω un ouvert de E .

Une application $f : E \rightarrow F_1 \times F_2$ est différentiable en un point $a \in E$ si et seulement si ces composantes ($f_i = p_i \circ f$) le sont. Dans ce cas, on a les égalités suivantes :

$$Df(a) = (Df_1(a), Df_2(a)) = u_1 \circ Df_1(a) + u_2 \circ Df_2(a),$$

$$\forall i \in \{1, 2\}, Df_i(a) = p_i \circ Df(a).$$

De même f est différentiable (resp C^1) sur Ω si ses composantes le sont.

preuve : On munit $F_1 \times F_2$ de la norme infinie du produit ($\|(y_1, y_2)\|_\infty = \sup(\|y_1\|_{F_1}, \|y_2\|_{F_2})$).

Si on suppose f différentiable. Puisque les p_i sont des applicaitons linéaires continues sont différentiables, alors les applications composées $p_i \circ f$ le sont aussi. Dans ce cas, on a :

$$Df_i(a) = D(p_i \circ f)(a) = Dp_i(f(a)) \circ Df(a),$$

or $Dp_i(x) = p_i$ quelque soit $x \in F_1 \times F_2$ puisque p_i est linaire et continue, donc :

$$Df_i(a) = p_i \circ Df(a).$$

Si maintenant on suppose que les composantes f_i sont toutes deux différentiables en $a \in \Omega$. On pose

$$\forall x \in E, L_a(x) = (Df_1(a)(x), Df_2(a)(x)).$$

On a :

$$\|f(x) - f(a) - L_a(x)\|_\infty = \sup_{i=1,2} \|f_i(x) - f_i(a) - Df_i(a)(x - a)\|_{E_i} = o(\|x - a\|).$$

Ainsi, f est différentiable et sa différentielle en a est bien donnée par $Df(a) = (Df_1(a), Df_2(a))$.

Ci-dessous, une application usuelle de la Proposition II.20 (Pourquoi au fait ?) :

Proposition II.21. Dérivée d'une application produit Soient E un $\mathbb{K}$ -evn, Ω un ouvert de E et $a \in \Omega$ et soient $f : E \rightarrow \mathbb{K}$ et $g : E \rightarrow \mathbb{K}$ deux fonctions différentiables en a .

L'application $\phi = fg$ est différentiable au point a et on a :

$$\forall h \in E, D\phi(a)(h) = f(a)Dg(a)(h) + g(a)Df(a)(h).$$

Ce résultat est en fait le cas particulier de la proposition suivante :

Proposition II.22. Dérivée d'une application bilinéaire continue Soient E et F deux $\mathbb{K}$ -evn, Ω un ouvert de E et $a \in \Omega$.

Soient $f : E \rightarrow F$ et $g : E \rightarrow F$ deux fonctions différentiables en a .

L'application $\phi = B(f, g)$ où $B \in \mathcal{L}(E, E; F)$: ϕ est différentiable en a et on a :

$$\forall h \in E, D\phi(a)(h) = B(f(a), Dg(a)(h)) + B(g(a), Df(a)(h)).$$

Démonstration laissée en exercice.

Exercice II.23. En utilisant les deux dernières propositions et la formule donnant la différentielle d'une application composée, retrouver les formules usuelles (sous conditions d'existence à expliciter) qui donnent les expressions de $D(f^n)(a)$ pour $n \geq 1$ et $D\left(\frac{1}{f}\right)(a)$ en fonction de f , Df et a .

Cas particulier des fonctions vectorielles de la variable réelle :

Définition II.24. Soient $F_1, F_2, \dots, F_n$ des $\mathbb{R}$ -evn et une application $f : \mathbb{R} \rightarrow F_1 \times F_2 \times \dots \times F_n$ est différentiable en un point $a \in \mathbb{R}$. On note f_i la composante de f sur F_i .

Le vecteur $\overrightarrow{\text{Grad}f}(a) = (f'_1(a), f'_2(a), \dots, f'_n(a))$ est appelé **gradient de f en a** .

On retrouve dans la littérature d'autres notation pour le gradient : $\text{Grad}f(a)$, $\overrightarrow{\nabla}f(a)$ ou $\nabla f(a)$

...

Proposition II.25. Soient $F_1, F_2, \dots, F_n$ des $\mathbb{R}$ -evn et une application $f : \mathbb{R} \rightarrow F_1 \times F_2 \times \dots \times F_n$ est différentiable en un point $a \in \mathbb{R}$. On note f_i la composante de f sur F_i .

La différentielle $Df(a)$ est l'application définie par :

$$\forall h = (h_1, h_2, \dots, h_n) \in \mathbb{R}^n, Df(a)(h) = \langle \overrightarrow{\text{Grad}f}(a), h \rangle$$

où $X \cdot Y$ représente le produit scalaire de X et Y dans $\mathbb{R}^n$.

Remarque II.26. Le gradient de f en a est l'analogue du nombre dérivé $f'(a)$ pour les fonctions vectorielles de la variable réelle. Ce n'est pas la différentielle de f en a mais il la caractérise totalement.

II.3 Différentielle d'une application définie sur un espace vectoriel produit & dérivées partielles

II.3.1 Applications partielles et différentielles partielles

Définitions II.27. Applications et différentielles partielles

Soient E_1, E_2 et F trois $\mathbb{K}$ -evn, Ω un ouvert de $E_1 \times E_2$ et $a = (a_1, a_2)$ un point de Ω .

Soit $f : \Omega \rightarrow F$ une application. On note f comme une fonction de 2 variables : $f : (x_1, x_2) \mapsto f(x_1, x_2)$.

On appelle **première application partielle associé à f au point a** l'application $x \mapsto f(x, a_2)$, définie sur $\Omega_1 = \{x \in E_1 \mid (x, a_2) \in \Omega\}$.

De même, on définit la **deuxième application partielle associé à f au point a** par l'application $x \mapsto f(a_1, x)$, définie sur $\Omega_2 = \{x \in E_2 \mid (a_1, x) \in \Omega\}$.

Si la i^{e} application partielle associée à f en a est différentiable en $a_i \in E_i$, sa différentielle en ce point est appelée **différentielle partielle de f par rapport à sa i^{e} variable au point a** . On la note, selon l'air du temps ou le contexte :

$$D_i f(a), D_{x_i} f(a), \partial_{x_i} f(a), \partial_i f(a), d_{x_i} f(a), \partial_i f(a), d_i f(a), f'_i(a) \dots$$

On parle de **dérivées partielles de f par rapport à sa i^{e} variable au point a** lorsque les $\mathbb{K}$ -evn E_i sont tous égaux à $\mathbb{K}$. Dans ce cas, les applications partielles sont des application de $\mathbb{K}$ dans F et calculer la différentielle d'une application partielle revient alors à la dériver au sens usuel. Dans ce contexte, on utilise plutôt les notations suivantes :

$$\frac{\partial f}{\partial x_i}(a), \frac{df}{dx_i}(a), \partial_i f(a), \partial_{x_i} f(a), d_i f(a), f'_i(a), \dots$$

Remarques II.28.

1. Lorsqu'elle existe, la différentielle partielle de f par rapport à sa i^{e} variable au point a est une application linéaire et continue de E_i vers F .
2. Il n'y a aucune objection à étendre ces définitions lorsque f est définie sur le produit de plus de 2 $\mathbb{K}$ -evn.
3. **Attention :** La notation $\frac{\partial f}{\partial x_i}$ peut porter à confusion, bien qu'elle soit la plus courante : ce n'est ABSOLUMENT PAS un quotient... c'est plutôt une notation historique due à nos amis physiciens car elle "symbolise" le rapport entre l'accroissement de la fonction : ∂f et l'accroissement de la variable x_i : ∂x_i ...
4. La notation des formes différentielles : $df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy$ est une autre manière de présenter les dérivées partielles :

Sous condition de différentiabilité, on a $Df(a, b)(h, k) = \frac{\partial f}{\partial x}(a, b)h + \frac{\partial f}{\partial y}(a, b)k$.

Remarque II.29. En pratique pour les calculs :

Calculer la i^{e} différentielle partielle de f en $a = (a_1, a_2, \dots, a_n)$, c'est en fait calculer la différentielle en $h = 0_{E_i}$ de la fonction ϕ_i définie sur $\mathbb{R}$ par $\phi_i(h) = f(a_1, \dots, a_{i-1}, a_i + h, a_{i+1}, \dots, a_n)$.

Cas particulier des fonctions de $\mathbb{R}^n$ dans $\mathbb{R}$:

Lorsqu'on travaille avec des fonctions de $\mathbb{R}^n$ dans $\mathbb{R}$, les fonctions partielles ϕ_i sont des applications de $\mathbb{R}$ dans $\mathbb{R}$... leurs différentielles (et donc les différentielles partielles de f) sont des applications linéaires continues de $\mathbb{R}$ dans $\mathbb{R}$ donc entièrement déterminées par leur coefficient directeur, qui est en fait la dérivée (au sens usuel) de la fonction partielle : $\phi'_i(a_i)$. C'est en général la notation $\frac{\partial f}{\partial x_i}(a)$ ou $\partial_i f(a)$ que l'on utilise pour désigner cette dérivée.

Exercice II.30.

1. Calculer la deuxième dérivée partielle de la fonction $f : \mathbb{R}^3 \rightarrow \mathbb{R}$, définie par $f(x, y, z) = \sin(xy) - \frac{z}{y^2}$ au point $(1, 1, 1)$ relativement à la base canonique de $\mathbb{R}^3$.
2. Soit E l'espace des polynômes de degré ≤ 2 d'une seule variable, muni de la base $\mathcal{B} = (e_1 = 1, e_2 = X, e_3 = X^2)$. On considère l'application $f : E \rightarrow \mathbb{R}$ définie pour un polynôme $P(X) = aX^2 + bX + c$ par $f(P) = \cos a + a^3 \cos c + \cos b$.
Calculer la troisième dérivée partielle de f relativement à la base $\mathcal{B}$ au point $Q = 1 + X^2$.

II.3.2 Lien entre différentielles partielles et différentielle d'une application

Les résultats présentés ici pour des evn produits de 2 evn sont généralisables sans problème à des evn produit de n evn.

Théorème II.31. Soient E_1, E_2 et F trois $\mathbb{K}$ -evn, Ω un ouvert de $E_1 \times E_2$ et $a = (a_1, a_2)$ un point de Ω . On note $\Omega_1 = \{x \in E_1 \mid (x, a_2) \in \Omega\}$ et $\Omega_2 = \{x \in E_2 \mid (a_1, x) \in \Omega\}$.

Si $f : \Omega \rightarrow F$ est une application différentiable en a , alors :

Les applications partielles associées à f en a sont toutes les deux différentiables, respectivement au point $a_1 \in \Omega_1$ et $a_2 \in \Omega_2$ et on a :

$$\forall (h_1, h_2) \in E_1 \times E_2, Df(a)(h_1, h_2) = \partial_1 f(a)(h_1) + \partial_2 f(a)(h_2).$$

Preuve : Si on définit les applications linéaires continues $u_1 : E_1 \rightarrow E_1 \times E_2$ et $u_2 : E_2 \rightarrow E_1 \times E_2$ par $u_1(x) = (x, a_2)$ et $u_2(x) = (a_1, x)$, les applications partielles sont données par $f \circ u_1$ et $f \circ u_2$. Par conséquent, si f est différentiable en a , alors les applications partielles le sont respectivement en a_1 et a_2 . D'autre part, on a $Du_1 = (u_1, 0)$ et $Du_2 = (0, u_2)$.

Ainsi, $\partial_i f(a) = D(f \circ u_i)(a) = Df(u_i(a)) \circ Du_i$. Pour tout $h = (h_1, h_2) \in E_1 \times E_2$, on a donc

$$\partial_1 f(a)(h_1) = Df((a_1, a_2)) \circ Du_1(h_1, h_2) = Df(a)(h_1, 0),$$

$$\text{et } \partial_2 f(a)(h_2) = Df((a_1, a_2)) \circ Du_2(h_1, h_2) = Df(a)(0, h_2).$$

Or, par linéarité, $Df(a)(h_1, h_2) = Df(a)(h_1, 0) + Df(a)(0, h_2)$, on a donc bien :

$$\forall (h_1, h_2) \in E_1 \times E_2, Df(a)(h_1, h_2) = \partial_1 f(a)(h_1) + \partial_2 f(a)(h_2),$$

ce qui conclut la preuve.

Attention : La proposition précédente est fondamentale car en pratique, dans les espaces de dimension finie (et les espaces de Banach), SOUS CONDITION DE DIFFÉRENTIABILITÉ, la différentielle va s'exprimer en fonction des différentielles partielles.

L'existence des différentielles partielles N'IMPLIQUE PAS la différentiabilité ! Voir par exemple l'exercice ci-dessous (dû à Cauchy).

Exercice II.32. On considère l'application $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par :

$$f(x, y) = \begin{cases} \frac{xy}{x^2+y^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } (x, y) = (0, 0) \end{cases}$$

1. Que valent les fonctions partielles de f en $(0, 0)$?
2. En déduire l'existence et la valeur des dérivées partielles de f en $(0, 0)$.
3. Montrer que f n'est pas continue en $(0, 0)$ et que, par conséquent, elle n'est pas différentiable en $(0, 0)$.

Remarque : La “réciproque” est uniquement vraie si les dérivées/différentielles partielles sont continues au point $a \in \Omega$. Ci dessous, l'énoncé du théorème, sur lequel on reviendra en détails car c'est une des grosses conséquences de l'inégalité de la moyenne, traité au chapitre suivant :

Théorème II.33. Soient $E_1, E_2, \dots, E_n$ et F des $\mathbb{K}$ -ev, Ω un ouvert de $E_1 \times E_2 \times \dots \times E_n$ et $f : \Omega \rightarrow F$ une application continue.

L'application f est de classe $\mathcal{C}^1$ sur Ω si et seulement si les n différentielles partielles de f $\frac{\partial f}{\partial x_i}$, $i = 1, \dots, n$ existent en tout point de Ω et sont des applications continues de Ω dans $\mathcal{L}(E_i, F)$.

En pratique : Ce théorème permet souvent de calculer des différentielles : On calcule les dérivées partielles au voisinage d'un point “critique” et on passe à la limite... Si cette limite existe et est finie, alors l'application est différentiable (et même $\mathcal{C}^1$).

II.3.3 Cas particulier des applications de $\mathbb{R}^n$ dans $\mathbb{R}^p$

Comme on l'a déjà mentionné à la Remarque II.29, le calcul des différentielles partielles d'une application de $\mathbb{R}^n$ dans $\mathbb{R}$ peut s'assimiler à n calculs simples de dérivées.

Les résultats de ce paragraphe sont en fait transposables pour des applications définies et à valeurs dans des evn de dimension finie.

Définition II.34. Matrice jacobienne et jacobien d'une application de $\mathbb{R}^n$ dans $\mathbb{R}^p$ Soit une application $f : \mathbb{R}^n \rightarrow \mathbb{R}^p$ différentiable en $a = (a_1, a_2, \dots, a_n)$. On appelle **matrice jacobienne de f en a** la matrice $J_f(a)$ associée à sa différentielle en a dans les bases canoniques de $\mathbb{R}^n$ et $\mathbb{R}^p$ (hé oui ! toute application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^p$ a une représentation matricielle !).

Si les f_i pour $i = 1, \dots, p$ sont les applications coordonnées de f , on a :

$$J_f(a) = \begin{pmatrix} \partial_1 f_1(a) & \dots & \partial_n f_1(a) \\ \vdots & & \vdots \\ \partial_1 f_p(a) & \dots & \partial_n f_p(a) \end{pmatrix}$$

où le coefficient $\partial_j f_i(a)$ représente la j -ième dérivée partielle de f_i .

Lorsque $n = p$, le **jacobien de f** est le déterminant de la matrice jacobienne de f . On le voit apparaître (en valeur absolue) lorsqu'on effectue des changements de variables dans les intégrales.

Proposition II.35. Matrices Jacobiennes et applications composées

Si $f : \mathbb{R}^n \rightarrow \mathbb{R}^p$ est une application différentiable en $a = (a_1, a_2, \dots, a_n)$ et que $g : \mathbb{R}^p \rightarrow \mathbb{R}^q$ est différentiable en $f(a)$, alors $g \circ f$ est différentiable en a , et on a, grâce à la formule donnant la différentielle d'une application composée :

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

Exercice II.36. Application/preuve dans un cas particulier :

Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ et $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ deux applications différentiables.

1. Décrire $Df(a, b)$ en fonction des dérivées partielles de f en $(a, b) : \frac{\partial f}{\partial x}(a, b)$ et $\frac{\partial f}{\partial y}(a, b)$ et donner sa jacobienne $J_f(a, b)$.
2. Décrire $Dg(f(a, b))$ en fonction des dérivées partielles de ϕ en $f(a, b) : \frac{\partial g}{\partial u}(f(a, b))$ et $\frac{\partial g}{\partial v}(f(a, b))$ et donner sa jacobienne $J_g(f(a, b))$.
3. En utilisant la formule de la différentielle d'une application composée, en déduire la différentielle de $g \circ f$ en (a, b) et sa jacobienne $J_{g \circ f}(a, b)$. Vérifier qu'on a bien $J_{g \circ f}(a, b) = J_g(f(a, b)) \times J_f(a, b)$.

Dérivées partielles et changement de bases :

Les dérivées partielles dépendent bien évidemment de la base choisie (et donc la jacobienne), mais il y a un lien entre les dérivées partielles dans des bases différentes.

Soit E et F deux $\mathbb{K}$ -evn. On suppose que E est de dimension finie.

Soient $\mathcal{B} = (e_1, e_2, \dots, e_n)$ et $\tilde{\mathcal{B}} = (\alpha_1, \alpha_2, \dots, \alpha_n)$ deux bases de E .

Soit $f : E \rightarrow F$ différentiable en $a = \sum b_i e_i = \sum a_i \alpha_i$ de dérivées partielles $D_{e_i} f(a)$. Changer de base, c'est comme si on changeait d'espace en fait...

Si A est l'application linéaire et continue permettant de passer de la base $\tilde{\mathcal{B}}$ à la base $\mathcal{B}$ ($\forall i, A(\alpha_i) = e_i$), alors $f(\sum b_i e_i) = f(A(\sum a_i \alpha_i)) = f \circ A(\sum a_i \alpha_i)$.

Ainsi, la i^e dérivée partielle de f dans la base $\mathcal{B}$ en $a = \sum b_i e_i$ sont les dérivées partielles de $f \circ A$ dans la base $\tilde{\mathcal{B}}$ au point $a = \sum a_i \alpha_i$:

$$D_{e_i} f(a) = D_{\alpha_i} f \circ A (A^{-1}(a)) .$$

En effet, puisque A est linéaire et continue, on a $DA = A$ et donc

$$D_{\alpha_i} f \circ A (A^{-1}(a)) = Df \circ A (A^{-1}(a)) (\alpha_i) = Df (A(A^{-1}(a)) \circ A(\alpha_i) = Df(a)(A(\alpha_i)) = Df(a)(e_i).$$

Proposition II.37. Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}^p$ est une application différentiable en $a = (a_1, a_2, \dots, a_n)$.

Soit $\mathcal{B}_n$ la base canonique de $\mathbb{R}^n$ et $\tilde{\mathcal{B}}_n$ une autre base de $\mathbb{R}^n$ et soit A_n la matrice de passage de la base $\tilde{\mathcal{B}}_n$ à la base $\mathcal{B}_n$.

Soit $\mathcal{B}_p$ la base canonique de $\mathbb{R}^p$ et $\tilde{\mathcal{B}}_p$ une autre base de $\mathbb{R}^p$ et soit A_p la matrice de passage de la base $\tilde{\mathcal{B}}_p$ à la base $\mathcal{B}_p$.

La jacobienne $\tilde{J}_f(a)$ de f en a dans les nouvelles bases $\tilde{\mathcal{B}}_n$ et $\tilde{\mathcal{B}}_p$ est donnée, en fonction de A_n , A_p et de la jacobienne $J_f(a)$ de f en a dans les bases $\mathcal{B}_n$ et $\mathcal{B}_p$ par la formule :

$$\tilde{J}_f(a) = A_p^{-1} \times J_f(a) \times A_n.$$

Exercice II.38. Soient f et g deux fonctions de $\mathbb{R}^2$ dans $\mathbb{R}$. Retrouver la forme des dérivées partielles du produit $f \times g$ en fonction des dérivées partielles de f et de g .

E, F evn	$Df(a) \in \mathcal{L}(E, F)$ $Df(a) : E \rightarrow F$ $h \mapsto Df(a)(h)$
$E = \mathbb{R}, F = \mathbb{R}$	$Df(a) \in \mathcal{L}(\mathbb{R}, \mathbb{R}) \simeq \mathbb{R}$ $Df(a) : \mathbb{R} \rightarrow \mathbb{R}$ $h \mapsto f'(a) \cdot h$ Dérivée de f en $a : f'(a)$
$E, F = F_1 \times F_2$ $f = (f_1, f_2)$	$Df(a) \in \mathcal{L}(E, F) \simeq \mathcal{L}(E, F_1) \times \mathcal{L}(E, F_2)$ $Df(a) : E \rightarrow F$ $h \mapsto (Df_1(a)(h), Df_2(a)(h))$
$E = \mathbb{R}, F = \mathbb{R}^p$ $f = (f_1, f_2, \dots, f_p)$	$Df(a) \in \mathcal{L}(\mathbb{R}, \mathbb{R}^p) \simeq \mathbb{R}^p$ $Df(a) : E \rightarrow \mathbb{R}^p$ $h \mapsto (f'_1(a) \cdot h, f'_2(a) \cdot h, \dots, f'_p(a) \cdot h)$ $= (f'_1(a), f'_2(a), \dots, f'_p(a)) \cdot h$ $= \overrightarrow{\text{Grad} f(a)} \cdot h$ Gradient de f en a : $\overrightarrow{\text{Grad} f(a)} = (f'_1(a), f'_2(a), \dots, f'_p(a))$
$E = E_1 \times E_2, F$	$Df(a) \in \mathcal{L}(E, F) \simeq \mathcal{L}(E_1, F) \times \mathcal{L}(E_2, F)$ $Df(a) : E_1 \times E_2 \rightarrow F$ $(h, k) \mapsto D_1 f(a)(h) + D_2 f(a)(k)$
$E = \mathbb{R}^n, F = \mathbb{R}$	$Df(a) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}) \simeq \mathbb{R}^n$ $Df(a) : \mathbb{R}^n \rightarrow \mathbb{R}$ $h = (h_1, h_2, \dots, h_n) \mapsto \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) \cdot h_i$
$E = \mathbb{R}^n, F = \mathbb{R}^p$ $f = (f_1, f_2, \dots, f_p)$	$Df(a) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p) \simeq \mathbb{R}^{n \times p}$ $Df(a) : \mathbb{R}^n \rightarrow \mathbb{R}^p$ $h \mapsto \left(\sum_{i=1}^n \frac{\partial f_1}{\partial x_i}(a) \cdot h_i, \sum_{i=1}^n \frac{\partial f_2}{\partial x_i}(a) \cdot h_i, \dots, \sum_{i=1}^n \frac{\partial f_p}{\partial x_i}(a) \cdot h_i \right)$ $= J_f(a) \cdot \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{pmatrix}$ où $h = (h_1, h_2, \dots, h_n)$. Matrice jacobienne de f : $J_f(a) = \left(\frac{\partial f_i}{\partial x_j}(a) \right)_{i \in [1, p], j \in [1, n]}$

Le tableau ci-dessus résume un peu les différents cas de figures et les cas particuliers qui ont été rencontrés dans cette section.

II.3.4 Dérivées directionnelles

Définition II.39. Soient E et F deux $\mathbb{K}$ -ev, Ω un ouvert de E et $f : \Omega \rightarrow F$ une application.

On dit que f est **dérivable en $a \in \Omega$ dans la direction $v \in E$** si la limite

$$\lim_{t \rightarrow 0} \frac{f(a + tv) - f(a)}{t}$$

existe et est finie, i.e. l'application vectorielle de la variable réelle $t \mapsto f(a + tv)$ est dérivable en 0.

Lorsque c'est le cas, cette limite est appelée **différentielle/dérivée de f en a dans la direction v** . On trouve différentes notations dans la littérature pour la dérivée de f en a dans la direction v . En voici quelques unes :

$$D_v f(a), \quad D_v f_a, \quad \partial_v f(a), \quad \partial_v f_a, \quad d_v f(a), \quad d_v f_a, \quad f'_v(a), \quad f'(a; d), \dots$$

Attention : L'existence des différentielles/dérivées directionnelles n'implique pas la différentiabilité, notamment car elle n'implique pas la continuité :

Exercice II.40.

1. Calculer les dérivées directionnelles de $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par :

$$f(x, y) = \begin{cases} \frac{y^3}{x} & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases}$$

2. Que vaut $\lim_{t \rightarrow 0} f(t^3, t)$? en déduire que f n'est pas différentiable au sens de Fréchet.

Liens entre différentielles directionnelles, différentielles partielles et différentielle :

Soient E et F deux $\mathbb{K}$ -evn, Ω un ouvert de E et $f : \Omega \rightarrow F$ une application.

On suppose que E est de dimension finie et $\mathcal{B} = (e_1, e_2, \dots, e_n)$ est une base de E .

Si f est différentiable en a et de différentielle $Df(a)$, on a grâce à la proposition II.43 :

$$D_v f(a) = Df(a)(v).$$

Et puisque tout vecteur v de E peut s'écrire $v = \sum_{i=1}^n v_i e_i$, et que $Df(a)$ est linéaire et continue, on a donc :

$$Df(a)(v) = \sum_{i=1}^n v_i Df(a)(e_i) = \sum_{i=1}^n v_i D_{e_i} f(a).$$

Définition II.41. Soient E et F deux $\mathbb{K}$ -ev, Ω un ouvert de E et $f : \Omega \rightarrow F$ une application.

Si f est dérivable en $a \in \Omega$ dans toutes les directions $v \in E$ et que l'application $v \in E \mapsto D_v f(a)$ est linéaire et continue, alors on dit que f est **différentiable au sens de Gâteaux** ou **Gâteaux-différentiable en a** . Autrement dit, f est différentiable au sens de Gâteaux en a s'il existe une application linéaire continue $\Lambda_a : E \rightarrow \mathcal{L}(E; F)$ telle que, pour tout $v \in E$, on ait :

$$\lim_{t \rightarrow 0} \frac{f(a + tv) - f(a)}{t} = \Lambda_a(v).$$

Attention : Notez bien qu'être différentiable au sens de Gâteaux est plus fort que l'admettre des dérivées directionnelles... Voici une illustrations de cette remarque :

Exercice II.42.

1. Calculer la dérivée directionnelle $D_{(u,v)}f(0,0)$ de $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ en $(0,0)$ définie par :

$$f(x, y) = \begin{cases} \frac{x(x^2-3y^2)}{x^2+y^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } (x, y) = (0, 0) \end{cases}$$

selon une direction arbitraire $(u, v) \in \mathbb{R}^2$.

2. L'application $(u, v) \rightarrow D_{(u,v)}f(0,0)$ est-elle linéaire et continue ? En déduire que f n'est pas Gâteaux-différentiable en $(0,0)$.

II.3.5 Liens entre différentiabilité et Gâteaux-différentiabilité

Et alors ? c'est quoi la différence entre les différentiabilité au sens de Fréchet et au sens de Gâteaux ?

On a d'abord l'implication suivante :

Proposition II.43. Soient E et F deux $\mathbb{K}$ -ev, Ω un ouvert de E et $f : \Omega \rightarrow F$ une application.

Si f est différentiable en a et de différentielle $Df(a)$, alors f elle est Gâteaux-différentiable, et on a :

$$D_v f(a) = Df(a)(v).$$

Ce qui se traduit, avec les notations de II.41, par $\Lambda_a = Df(a)$.

Remarque II.44. La réciproque est fausse, essentiellement car, contrairement à la différentiabilité, le fait d'être Gâteaux-différentiable n'implique pas la continuité de la fonction elle-même comme on l'a vu pour l'existence des différentielles partielles... Voici une illustrations de cette remarque :

Exercice II.45. 1. Calculer les dérivées directionnelles de $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par :

$$f(x, y) = \begin{cases} \frac{x^6}{y^4+(y-x^2)^2} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{si } (x, y) = (0, 0) \end{cases}$$

2. En déduire que f est Gâteaux-différentiable en $(0,0)$.
3. Que vaut $\lim_{x \rightarrow 0} f(x, x^2)$? en déduire que f n'est pas différentiable au sens de Fréchet.

II.4 Remarque sur la $\mathbb{R}$ -différentiabilité et la $\mathbb{C}$ -différentiabilité

Ce paragraphe traite de l'importance du corps de base... Si E et F sont des $\mathbb{K}$ -evn et que f est une fonction de E dans F , la différentiabilité de f dépend intimement du corps $\mathbb{K}$, puisque la recherche de la différentielle se fait parmi les applications $\mathbb{K}$ -linéaires continues de E dans F ... On

peut illustrer ce fait avec l'exemple suivant.

On considère une fonction $f : \mathbb{C} \rightarrow \mathbb{C}$. Le corps $\mathbb{C}$ peut être muni de sa structure de $\mathbb{R}$ -evn ($\mathbb{C}$ est en fait indentifié à $\mathbb{R}^2$), mais il peut être muni également de sa structure de $\mathbb{C}$ -evn. Lorsque se pose le problème de la différentiabilité de f en $a \in \mathbb{C}$, on peut l'envisager dans les 2 cadres :

— $\mathbb{R}$ -différentiabilité : $\mathbb{C}$ est muni de sa structure de $\mathbb{R}$ -evn :

On cherche alors une application $\mathbb{R}$ -linéaire continue de $\mathbb{C}$ dans $\mathbb{C}$: $\phi_a \in \mathcal{L}_{\mathbb{R}}(\mathbb{C}; \mathbb{C})$ telle que :
 $f(z) = f(a) + \phi_a(z - a) + o(\|z - a\|)$.

Si ϕ existe, on l'appelle la $\mathbb{R}$ -différentielle de f en a .

— $\mathbb{C}$ -différentiabilité : $\mathbb{C}$ est muni de sa structure de $\mathbb{C}$ -evn :

On cherche alors une application $\mathbb{C}$ -linéaire continue de $\mathbb{C}$ dans $\mathbb{C}$: $\psi_a \in \mathcal{L}_{\mathbb{C}}(\mathbb{C}; \mathbb{C})$ telle que :
 $f(z) = f(a) + \psi_a(z - a) + o(\|z - a\|)$.

Si ψ existe, on l'appelle la $\mathbb{C}$ -différentielle de f en a .

Or $\mathcal{L}_{\mathbb{C}}(\mathbb{C}; \mathbb{C})$ et $\mathcal{L}_{\mathbb{R}}(\mathbb{C}; \mathbb{C})$ ne sont pas les mêmes ensembles !!!

On a bien sur $\mathcal{L}_{\mathbb{C}}(\mathbb{C}; \mathbb{C}) \subset \mathcal{L}_{\mathbb{R}}(\mathbb{C}; \mathbb{C})$ (c'est immédiat car $\mathbb{R} \subset \mathbb{C}$, donc une application $\mathbb{C}$ -linéaire est également $\mathbb{R}$ -linéaire) donc :

Proposition II.46. *Soit une application $f : \mathbb{C} \rightarrow \mathbb{C}$.*

Si l'application f est $\mathbb{C}$ -différentiable en $a \in \mathbb{C}$, alors elle est aussi $\mathbb{R}$ -différentiable.

Dans ce cas, la $\mathbb{R}$ -différentielle de f en a est exactement la $\mathbb{C}$ -différentielle de f en a .

Démonstration laissée en exo... presque immédiat.

Cependant, certaines applications $\mathbb{R}$ -linéaires de $\mathbb{C}$ dans $\mathbb{C}$ ne sont pas $\mathbb{C}$ -linéaires... Par exemple, l'application f définie par $f(x + iy) = x - iy$ est $\mathbb{R}$ -linéaire mais n'est pas $\mathbb{C}$ -linéaire :

— On a bien, pour $\lambda \in \mathbb{R}$, $f(\lambda(x + iy)) = \lambda x - i\lambda y = \lambda(x - iy) = \lambda f(x + iy)$, donc f est $\mathbb{R}$ -linéaire.

— $f((1+i)(x+iy)) = f(x-y+i(x+y)) = x-y-i(x+y)$ et $(1+i)f(x+iy) = (1+i)(x-iy) = x+y+i(x-y)$... donc $f((1+i)(x+iy)) \neq (1+i)f(x+iy)$ (sauf pour $x+iy=0$).

Pour une fonction $f : \mathbb{C} \rightarrow \mathbb{C}$, se peut donc qu'on puisse écrire $f(z) = f(a) + \phi_a(z - a) + o(\|z - a\|)$ avec ψ $\mathbb{R}$ -linéaire mais non $\mathbb{C}$ linéaire. Dans ce cas, au point a , f est $\mathbb{R}$ -différentiable mais n'est pas $\mathbb{C}$ -différentiable.

Ainsi, **Pour qu'une fonction $f : \mathbb{C} \rightarrow \mathbb{C}$ $\mathbb{R}$ -différentiable en a soit aussi $\mathbb{C}$ -différentiable en a , il faut rajouter des conditions...** Ces conditions doivent exprimer le fait que la $\mathbb{R}$ -différentielle de f est une application $\mathbb{C}$ -linéaire, en plus d'être $\mathbb{R}$ -linéaire.

Proposition II.47. *Soit $\phi : \mathbb{C} \rightarrow \mathbb{C}$ une application $\mathbb{R}$ -linéaire, définie par*

$$\phi(x + iy) = (ax + by) + i(cx + dy).$$

L'application ϕ est $\mathbb{C}$ -linéaire si et seulement si $a = d$ et $b = -c$.

Preuve : L'application ϕ est $\mathbb{C}$ -linéaire si et seulement si $(\alpha + i\beta)\phi(x + iy) = \phi((\alpha + i\beta)(x + iy))$ pour tout $\alpha + i\beta \in \mathbb{C}$. Par $\mathbb{R}$ -linéarité, on a :

$$(\alpha + i\beta)\phi(x + iy) = \alpha\phi(x + iy) + i\beta\phi(x + iy) \quad \text{et} \quad \phi((\alpha + i\beta)(x + iy)) = \alpha\phi(x + iy) + \beta\phi(i(x + iy)).$$

Ainsi, l'application ϕ est $\mathbb{C}$ -linéaire si et seulement si $i\phi(x+iy) = \phi(i(x+iy))$, c'est-à-dire que ϕ est $\mathbb{C}$ -linéaire si et seulement si $-cx - dy + i(ax + by) = bx - ay + i(dx - cy)$.

Par identification, cette condition est équivalente à : $-c = b$ et $a = d$.

Bon, d'accord, mais **en pratique ?**

Soit $f : \mathbb{C} \rightarrow \mathbb{C}$ une application $\mathbb{R}$ -différentiable en $z_0 = a+ib$, définie par $f(z) = X(z) + iY(z)$ où X et Y sont des application à valeurs réelles de la variable complexe. Cette application f peut-être identifiée avec l'application $\hat{f} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, définie par $\hat{f}(x, y) = (X(x+iy), Y(x+iy))$.

La matrice de la $\mathbb{R}$ -différentielle de $\hat{f}$ en (a, b) est donnée par :

$$J_{\hat{f}}(a) = \begin{pmatrix} \frac{\partial X}{\partial x}(a, b) & \frac{\partial X}{\partial y}(a, b) \\ \frac{\partial Y}{\partial x}(a, b) & \frac{\partial Y}{\partial y}(a, b) \end{pmatrix}.$$

C'est-à-dire que, pour tout $(h_1, h_2) \in \mathbb{R}^2$,

$$D\hat{f}(a, b)(h_1, h_2) = \left(\frac{\partial X}{\partial x}(a, b)h_1 + \frac{\partial X}{\partial y}(a, b)h_2, \frac{\partial Y}{\partial x}(a, b)h_1 + \frac{\partial Y}{\partial y}(a, b)h_2 \right).$$

En utilisant la proposition précédente, on a : $\hat{f}$ est $\mathbb{C}$ -différentiable en (a, b) (et donc f est $\mathbb{C}$ -différentiable en $z_0 = a + ib$) si et seulement si :

$$\frac{\partial X}{\partial x}(a, b) = \frac{\partial Y}{\partial y}(a, b) \quad \text{et} \quad \frac{\partial X}{\partial y}(a, b) = -\frac{\partial Y}{\partial x}(a, b).$$

Ces conditions sont connues sous le nom de **conditions de Cauchy**.

Les fonctions $\mathbb{C}$ -différentiables de $\mathbb{C}$ dans $\mathbb{C}$ sont également connues sous le nom de **fonctions holomorphes** (ce qui est quand même plus court !) et font l'objet d'une autre partie du cours de Mathématiques.

Chapitre III

Inégalité de la moyenne

L'inégalité de la moyenne, c'est l'extension du théorème des accroissements finis. On le trouve d'ailleurs dans certains bouquins sous l'appellation théorème des accroissements finis, même dans le cadre général de fonctions différentiables d'un espace de Banach dans un autre. Plusieurs "variantes" de ce résultats existent d'ailleurs.

D'autre part, la terminologie "théorème des accroissements finis" s'explique d'une manière historique, par le désir de mettre en opposition les accroissements finis et les accroissements infinitésimaux dont on parlait autrefois en calcul différentiel.

III.1 Enoncé(s) du théorème

Comme mentionné en introduction, il existe plusieurs version de ce qu'on appelle théorème des accroissements finis ou inégalité de la moyenne. On va en présenter 2 : un pour les applications vectorielles de la variable réelle et un pour les applications d'un $\mathbb{K}$ -evn quelconque dans un autre $\mathbb{K}$ -evn.

III.1.1 Rappel : le cas des fonctions réelles de la variable réelle

Théorème III.1. Théorème des accroissements finis

Soit $f : [a, b] \rightarrow \mathbb{R}$ une application continue et dérivable sur $]a, b[$.

$$\exists c \in]a, b[, f(b) - f(a) = (b - a)f'(c)$$

Pour donner une image, on peut illustrer ainsi le théorème : "Si un véhicule parcourt une distance à la vitesse moyenne de 60km/h, alors son compteur a indiqué au moins une fois la vitesse précise de 60 km/h."

Théorème III.2. Inégalité de la moyenne

Soient $f : [a, b] \rightarrow \mathbb{R}$ et $g : [a, b] \rightarrow \mathbb{R}$ deux applications continues et dérivable sur $]a, b[$. Si pour tout $t \in]a, b[$, on a $f'(t) \leq g'(t)$, alors on a :

$$f(b) - f(a) \leq g(b) - g(a).$$

En particulier, s'il existe un nombre $M > 0$ tel que, pour tout $t \in]a, b[$, on a $f'(t) \leq M$, alors :

$$f(b) - f(a) < M(b - a).$$

Pour donner une image, on peut illustrer ainsi le théorème : “Si un premier véhicule roule toujours moins vite qu'un deuxième véhicule, à temps égal, le premier aura tout parcouru moins que le second.”

Remarque/Exercice : Expliquer pourquoi, dans ce théorème on aurait pu mettre ces conclusions plus fortes :

$$\forall (x, y) \in [a, b]^2, x < y \implies f(y) - f(x) < g(y) - g(x).$$

Et s'il existe un nombre $M > 0$ tel que, pour tout $t \in]a, b[$, on a $f'(t) \leq M$, alors :

$$\forall (x, y) \in [a, b]^2, x < y \implies f(y) - f(x) < M(y - x),$$

mais cela nuit un peu à la compréhension du problème et ne rajoute finalement pas grand chose.

Remarque/Exercice : Ce théorème n'est plus valable pour des fonctions de $\mathbb{R}$ dans $\mathbb{R}^2$...

Montrer que $f : [0, 2\pi] \rightarrow \mathbb{R}^2$ la fonction définie par $f(x) = (\cos x, \sin x)$ ne vérifie le théorème des accroissements finis dans sa version usuelle.

III.1.2 Fonctions vectorielles de la variable réelle

Théorème III.3. Soient F un espace vectoriel normé et $a < b$ deux nombres réels.

Soit $f : [a, b] \rightarrow F$ et $g : [a, b] \rightarrow \mathbb{R}$ deux applications continues et différentiables en tout point x de $]a, b[$, de différentielles $f'(x)$ et $g'(x)$ et que :

$$\forall x \in]a, b[, \|f'(x)\| \leq g'(x).$$

On a alors :

$$\|f(b) - f(a)\| \leq g(b) - g(a).$$

Remarque : C'est exactement le même résultat que l'inégalité de la moyenne pour les fonctions réelles de la variable réelle.

Ce théorème est en fait une version affaiblie d'un théorème plus fort, qui met en jeu non pas les différentielles, mais les différentielles à droite (ou à gauche) des fonctions vectorielles de la variable réelle :

Définition III.4. Soit F un $\mathbb{R}$ -evn.

On dit que $f : [a, b] \rightarrow F$ est **dérivable à gauche** (resp. **à droite**) **en un point** $a \in \mathbb{R}$ si et seulement si la limite $\lim_{x \rightarrow a, x < a} \frac{1}{x-a}(f(x) - f(a))$ (resp. $\lim_{x \rightarrow a, x > a} \frac{1}{x-a}(f(x) - f(a))$) existe et est finie. Cette limite quand elle existe est appelée **dérivée à gauche** (resp. **à droite**) **de f en $a \in \mathbb{R}$.**

Théorème III.5. Soient F un espace vectoriel normé et $a < b$ deux nombres réels.

Soit $f : [a, b] \rightarrow F$ et $g : [a, b] \rightarrow \mathbb{R}$ deux applications continues et dérivables à droite (resp. à gauche) en tout point x de $]a, b[$, de dérivées à droite (resp. à gauche) $f'_d(x)$ et $g'_d(x)$ (resp. $f'_g(x)$ et $g'_g(x)$). On suppose que :

$$\forall x \in]a, b[, \|f'_d(x)\| \leq g'_d(x) \text{ (resp. } \|f'_g(x)\| \leq g'_g(x)).$$

On a alors :

$$\|f(b) - f(a)\| \leq g(b) - g(a).$$

Démonstration : Soit $\epsilon > 0$. On va montrer que

$$\forall x \in [a, b], \|f(x) - f(a)\| \leq g(x) - g(a) + \epsilon(x - a) + \epsilon, \quad (\text{III.1})$$

puis cette inégalité sera appliquée à $x = b$ et on fera tendre ϵ vers 0 pour obtenir le résultat voulu.

Soit U_ϵ l'ensemble des $x \in [a, b]$ pour lesquels l'équation (III.1) est fautive :

$$U_\epsilon = \{x \in [a, b], \|f(x) - f(a)\| > g(x) - g(a) + \epsilon(x - a) + \epsilon\}.$$

On va montrer que U_ϵ est vide...

Tout d'abord, U_ϵ est un ensemble ouvert de $[a, b]$ car s'est l'image réciproque de l'ouvert $\mathbb{R}_+^*$ par l'application continue $x \mapsto \|f(x) - f(a)\|_F - (g(x) - g(a) + \epsilon(x - a) + \epsilon)$.

Raisonnons maintenant par l'absurde, en supposant U_ϵ non vide. Puisque U_ϵ est un ouvert non vide de $[a, b]$, il a une borne inférieure : c . On peut dire 3 choses sur c :

1. $c > a$ En effet, l'équation (III.1) est vraie pour a et, par continuité, sur un voisinage de a .
2. $c \notin U_\epsilon$, puisque U_ϵ est ouvert.
3. $c < b$, car si $c = b$, alors $U_\epsilon \subset \{b\}$ et puisque U_ϵ est ouvert, il serait forcément vide...

On a également, par hypothèse : $\|f'_d(c)\| \leq g'_d(c)$ et par définition de la dérivée à droite, on a :

$$\lim_{x \rightarrow a, x > a} \frac{f(x) - f(c)}{x - c} = f'_d(c) \quad \text{et} \quad \lim_{x \rightarrow a, x > a} \frac{g(x) - g(c)}{x - c} = g'_d(c).$$

Ainsi, il existe un intervalle $[c, c + \eta]$ tel que, pour tout $x \in [c, c + \eta]$, on ait :

$$\left\| \frac{f(x) - f(c)}{x - c} \right\| \leq \|f'_d(c)\| + \frac{\epsilon}{2} \quad \text{et} \quad \left\| \frac{g(x) - g(c)}{x - c} \right\| \geq \|g'_d(c)\| - \frac{\epsilon}{2}.$$

On déduit de ces trois dernières inégalités que :

$$\forall x \in [c, c + \eta], \|f(x) - f(c)\| \leq g(x) - g(c) + \epsilon(x - c).$$

D'autre part, puisqu'on a vu que $c \notin U_\epsilon$, on a :

$$\|f(c) - f(a)\| \leq g(c) - g(a) + \epsilon(c - a) + \epsilon$$

Ces deux dernières inégalités implique donc, pour tout $x \in [c, c + \eta]$:

$$\begin{aligned} \|f(x) - f(a)\| &\leq \|f(x) - f(c)\| + \|f(c) - f(a)\|, \\ &\leq g(x) - g(c) + g(c) - g(a) + \epsilon(x - c) + \epsilon(c - a) + \epsilon, \\ &\leq g(x) - g(a) + \epsilon(x - a) + \epsilon. \end{aligned}$$

Ainsi, pour tout $x \in [c, c + \eta]$, $x \notin U_\epsilon$... cela contredit le fait que c est la borne inférieure de U_ϵ . L'ensemble U_ϵ est donc vide. On a :

$$\forall x \in [a, b], \|f(x) - f(a)\| \leq g(x) - g(a) + \epsilon(x - a) + \epsilon,$$

En particulier, pour $x = b$ et en faisant tendre ϵ vers 0 on obtient la conclusion du théorème.

Corollaire III.6. Soient F un espace vectoriel normé et $a < b$ deux nombres réels.

Soit $f : [a, b] \rightarrow F$ une application continue et dérivable à droite (resp. à gauche) en tout point x de $]a, b[$, de dérivée à droite (resp. à gauche) $f'_d(x)$ (resp. $f'_g(x)$). On suppose que :

$$\exists M > 0, \forall x \in]a, b[, \|f'_d(x)\| \leq M \text{ (resp. } \|f'_g(x)\| \leq M).$$

On a alors :

$$\|f(b) - f(a)\| \leq M(b - a).$$

III.1.3 Fonctions d'un $\mathbb{K}$ -evn dans un autre $\mathbb{K}$ -evn

Nous avons vu jusqu'à présent que le théorème des accroissements finis nous dit essentiellement (dans le cas vectoriel et réel) que si la dérivée est bornée, la différence de déplacement induite par f entre les 2 extrémités d'un segment ($f(b) - f(a)$) est majorée par une quantité qui dépend linéairement de la longueur du segment ($b - a$).

Pour pouvoir avoir un analogue du théorème des accroissements finis dans le cas des applications d'un $\mathbb{K}$ -evn quelconque dans un autre $\mathbb{K}$ -evn, il faut donc déjà être sûr que, si f est définie en a et b , elle se sera sur tout le *segment* $[a, b]$...

Définitions III.7. Soient E un evn et a et b deux points de E . On appelle **segment d'extrémités a et b** , l'ensemble noté $[a, b]$ défini par :

$$[a, b] = \{a + t(b - a), t \in [0, 1]\}.$$

Un ensemble Ω de E est dit **convexe** si pour tout $(a, b) \in \Omega^2$, $[a, b] \subset \Omega$.

Notez que, avec cette définition, $[a, b] = [b, a]$... l'ordre usuel avec lequel on note les segments de $\mathbb{R}$ ($a < b$) n'est en fait qu'une convention bien pratique et est applicable parce que $\mathbb{R}$ est ordonné totalement. Si sur un espace plus gros, on a un ordre total, on peut appliquer une convention similaires aux segments.

Théorème III.8. Soient E et F deux $\mathbb{K}$ -evn, Ω un ouvert de E et une application différentiable $f : \Omega \rightarrow F$.

Soient a et b deux points de Ω tels que $[a, b] \subset \Omega$.

L'application f vérifie :

$$\|f(b) - f(a)\|_F \leq \left(\sup_{x \in [a,b]} \|Df(x)\| \right) \|b - a\|_E.$$

En particulier, si Ω est convexe et qu'il existe un $M > 0$ tel que, pour tout $x \in \Omega$, $\|Df(x)\| \leq M$, alors :

$$\forall (a, b) \in \Omega^2, \|f(b) - f(a)\|_F \leq M \|b - a\|_E.$$

Démonstration en exercice : On se ramène à l'application du Théorème III.3 et du Corollaire III.6, grâce à la fonction $\phi : t \mapsto f(a + t(b - a))$.

Corollaire III.9. Une application différentiable d'un ouvert connexe Ω d'un $\mathbb{K}$ -evn dans un autre $\mathbb{K}$ -evn dont la différentielle est identiquement nulle est constante.

Si Ω n'est pas connexe, alors sous les mêmes conditions, l'application est constante sur chacune des composantes connexes de Ω .

Preuve : Soit $a \in \Omega$ sur lequel est définie l'application considérée f . Tout point x de Ω est le centre d'une boule ouverte B contenue dans Ω . Or les boules ouvertes sont des parties convexes de Ω . Ainsi, on peut y appliquer le corollaire III.6 pour montrer que f est constante sur B . Ainsi, pour tout point x de Ω tel que $f(x) = f(a)$, f est constante égale à $f(a)$: l'ensemble des éléments x de Ω pour lesquels $f(x) = f(a)$ est donc un ouvert de Ω .

D'autre part, l'ensemble des éléments x de Ω pour lesquels $f(x) = f(a)$ est un fermé de Ω (image réciproque du singleton $\{f(a)\}$ par une application continue). Cet ensemble est donc $\emptyset$ ou Ω lui-même. Or cet ensemble contient au moins a donc f est constante égale à $f(a)$ sur Ω tout entier.

III.2 Application : Retour sur les liens entre différentiabilité et dérivées partielles

Soient E et F deux $\mathbb{K}$ -ev, Ω un ouvert de E et $f : \Omega \rightarrow F$ une application.

On a vu au chapitre précédent (Proposition II.43) que si f est différentiable en a et de différentielle $Df(a)$, alors f elle est Gâteaux-différentiable, et on a :

$$D_v f(a) = Df(a)(v).$$

mais on a également vu que la réciproque était fautive. On a cependant une conditions sur les différentielles partielles qui permet de décider de la différentiabilité d'une fonction en un point :

Théorème III.10. Soient $E_1, E_2, \dots, E_n$ et F des $\mathbb{K}$ -ev, Ω un ouvert de $E_1 \times E_2 \times \dots \times E_n$ et $f : \Omega \rightarrow F$ une application continue.

f est de classe $\mathcal{C}^1$ sur Ω si et seulement si les n différentielles partielles de f $\frac{\partial f}{\partial x_i}$, pour $i = 1, \dots, n$ existent en tout point de Ω et sont des applications continues de Ω dans $\mathcal{L}(E_i, F)$.

Attention : Notez bien que la **continuité de f fait partie des hypothèses!!!**

Démonstration : Preuve pour $n = 2$ suffit. On généralise très facilement (en utilisant un raisonnement par récurrence).

Soient E_1, E_2 et F des $\mathbb{K}$ -ev, Ω un ouvert de $E_1 \times E_2$ et $f : \Omega \rightarrow F$ une application continue.

On notera par $p_i : E \rightarrow E_i$ pour $i = 1$ et 2 les projections canoniques de E sur ses composantes et par $u_i : E_i \rightarrow E$ pour $i = 1$ et 2 les injections canoniques des E_i dans E de sorte que :

$$\forall (h, k) \in E_1 \times E_2, p_1(h, k) = h, p_2(h, k) = k, u_1(h) = (h, 0), \text{ et } u_2(k) = (0, k).$$

Première partie : On suppose f de classe $\mathcal{C}^1$, i.e. l'application $x \mapsto Df(x)$ est continue de Ω dans $\mathcal{L}(E_1 \times E_2; F)$.

Soit $x = (x_1, x_2) \in \Omega$. Puisque f est de classe $\mathcal{C}^1$, ses différentielles partielles existent et on a :

$$\forall (h, k) \in E_1 \times E_2, Df(x)(h, k) = \frac{\partial f}{\partial x_1}(x)(h) + \frac{\partial f}{\partial x_2}(x)(k).$$

Ainsi, on a les égalités suivantes :

$$\forall h \in E_1, \frac{\partial f}{\partial x_1}(x)(h) = Df(x)(h, 0) = Df(x) \circ u_1(h)$$

et

$$\forall k \in E_2, \frac{\partial f}{\partial x_2}(x)(k) = Df(x)(0, k) = Df(x) \circ u_2(k).$$

Ainsi, pour $i = 1$ ou 2 , l'application $\frac{\partial f}{\partial x_i} : x \mapsto \frac{\partial f}{\partial x_i}(x)$ de Ω dans $\mathcal{L}(E_i; F)$ peut donc s'écrire comme la composée l'application constante (donc continue) de Ω dans $\mathcal{L}(E_i; E_2 \times E_2)$ définie par $x \mapsto u_i$ et de l'application $Df : x \mapsto D(x)$ continue par hypothèse. L'application $\frac{\partial f}{\partial x_i}$ est donc continue sur Ω .

Deuxième partie : On suppose que les différentielles partielles $\frac{\partial f}{\partial x_1}$ et $\frac{\partial f}{\partial x_2}$ existent et sont continues sur Ω . On pose,

$$\forall x \in E_1 \times E_2, \forall (h, k) \in E_1 \times E_2, g(x)(h, k) = \frac{\partial f}{\partial x_1}(x)(h) + \frac{\partial f}{\partial x_2}(x)(k),$$

L'application $g : \Omega \rightarrow \mathcal{L}(E_1 \times E_2; F)$ ainsi définie est somme de deux termes. Chacun de ces deux termes sont continus sur Ω car ils sont composés de l'application constante (donc continue) de Ω dans $\mathcal{L}(E_2 \times E_2; E_i)$ définie par $x \mapsto p_i$ et de l'application $\frac{\partial f}{\partial x_i} : x \mapsto \frac{\partial f}{\partial x_i}(x)$ continue par hypothèse.

L'application g est donc continue. Il reste donc à montrer que cette fonction est bien la différentielle de f ...

Soient $x = (x_1, x_2) \in \Omega$ et V un voisinage ouvert convexe de x contenu dans Ω . Pour tout $y = (y_1, y_2) \in V$, on pose :

$$\delta(x, y) = f(y) - f(x) - g(x)(y - x).$$

Soit $\epsilon > 0$. on va montrer qu'il existe un voisinage de x sur lequel $\delta(x, y) = o(\|y - x\|)$. En rajoutant et enlevant $f(x_1, y_2)$ à $\delta(x, y)$, on a

$$\delta(x, y) = f(y_1, y_2) - f(x_1, x_2) + f(x_1, y_2) - f(x_1, y_2) - \frac{\partial f}{\partial x_1}(x_1, x_2)(y_1 - x_1) - \frac{\partial f}{\partial x_2}(x_1, x_2)(y_2 - x_2),$$

de sorte que :

$$\begin{aligned} \|\delta(x, y)\| \leq & \left\| f(y_1, y_2) - f(x_1, y_2) - \frac{\partial f}{\partial x_1}(x_1, x_2)(y_1 - x_1) \right\| \\ & + \left\| f(x_1, y_2) - f(x_1, x_2) - \frac{\partial f}{\partial x_2}(x_1, x_2)(y_2 - x_2) \right\|, \end{aligned}$$

On va montrer que le second membre de l'inégalité est un $o(\|y - x\|)$, ce qui entraînera que $\delta(x, y) = o(\|y - x\|)$.

En utilisant la définition même de la différentielle partielle $\frac{\partial f}{\partial x_2}(x)$, on a :

$$\left\| f(x_1, y_2) - f(x_1, x_2) - \frac{\partial f}{\partial x_2}(x_1, x_2)(y_2 - x_2) \right\| = o(\|y - x\|).$$

Pour le premier terme : $\left\| f(y_1, y_2) - f(x_1, y_2) - \frac{\partial f}{\partial x_1}(x_1, x_2)(y_1 - x_1) \right\|$, on va utiliser l'inégalité de la moyenne.

On considère x_1 , x_2 et y_2 comme fixés et on applique l'inégalité de la moyenne à l'application $\phi : E_1 \rightarrow F$ définie sur E_1 par :

$$\phi(z) = f(z, y_2) - \frac{\partial f}{\partial x_1}(x_1, x_2)(z),$$

de sorte que, $D\phi(z) = \frac{\partial f}{\partial x_1}(z, y_2) - \frac{\partial f}{\partial x_1}(x_1, x_2)$ et que, par linéarité de $\frac{\partial f}{\partial x_1}(x)$, on ait :

$$f(y_1, y_2) - f(x_1, y_2) - \frac{\partial f}{\partial x_1}(x_1, x_2)(y_1 - x_1) = \phi(y_1) - \phi(x_1).$$

L'inégalité de la moyenne appliquée à ϕ sur l'intervalle $[x_1, y_1]$ implique alors que

$$\left\| f(y_1, y_2) - f(x_1, y_2) - \frac{\partial f}{\partial x_1}(x_1, x_2)(y_1 - x_1) \right\| \leq \sup_{z \in [x_1, y_1]} \left\| \frac{\partial f}{\partial x_1}(z, y_2) - \frac{\partial f}{\partial x_1}(x_1, x_2) \right\| \|y - x\|.$$

L'application $\frac{\partial f}{\partial x_1}$ étant continue sur Ω et donc en (x_1, x_2) , on a :

$$\lim_{(z, y_2) \rightarrow (x_1, x_2)} \left\| \frac{\partial f}{\partial x_1}(z, y_2) - \frac{\partial f}{\partial x_1}(x_1, x_2) \right\| = 0$$

donc :

$$\left\| \frac{\partial f}{\partial x_1}(z, y_2) - \frac{\partial f}{\partial x_1}(x_1, x_2) \right\| = o(\|y - x\|).$$

Corollaire III.11. *Applications de $\mathbb{R}^n$ dans un $\mathbb{R}$ -evn.*

Soit F un $\mathbb{R}$ -evn.

Une application f continue d'un ouvert Ω de $\mathbb{R}^n$ vers F est de classe $\mathcal{C}^1$ si et seulement si les n -dérivées partielles de $f : \frac{\partial f}{\partial x_i}$ existent en tout point de Ω et sont continues sur Ω .

III.3 Application : Convergence de suites de fonctions différentiables

Théorème III.12. Soient Ω un ouvert d'un $\mathbb{K}$ -evn E et F un espace de Banach.

Soit $(f_n)_{n \geq 0}$ une suite d'applications différentiables de E dans F telle que :

1. Il existe $a \in \Omega$ tel que la suite $(f_n(a))_{n \geq 0}$ converge ;
2. la suite des différentielles $(Df_n)_{n \geq 0}$ converge uniformément vers l'application $g : E \rightarrow \mathcal{L}(E; F)$.

Alors, pour tout $x \in \Omega$, la suite $(f_n(x))_{n \geq 0}$ converge vers un élément de E noté $f(x)$ et l'application $f : \Omega \rightarrow F$ définie ainsi est différentiable sur Ω , de différentielle $Df = g$.

De plus, la convergence de la suite $(f_n)_{n \geq 0}$ vers f est uniforme sur toute partie bornée de Ω .

Remarque III.13. Ce théorème existe en version “locale” : la propriété de convergence uniforme dans la condition 2. et dans la conclusion est remplacée par la propriété de convergence uniforme locale : Tout point de Ω possède un voisinage sur lequel la convergence est uniforme.

Démonstration :

La démonstration se scinde en deux parties. Dans un premier temps, on montre grâce au fait que $(f_n(a))_{n \geq 0}$ converge et au théorème des accroissements finis, que la suite $(f_n)_{n \geq 0}$ est de Cauchy sur toute partie bornée de Ω , ce qui donnera la convergence uniforme de $(f_n)_{n \geq 0}$ sur toute partie bornée de Ω . Ensuite, on va montrer que la fonction limite f est différentiable et donner la valeur de Df en utilisant de nouveau le théorème des accroissements finis.

Soient $M > 0$ et $x \in \Omega$ un point vérifiant $\|x - a\| \leq M$. Pour tous entiers n et m , on a d'après le théorème des accroissements finis :

$$\|f_m(x) - f_n(x) - f_m(a) + f_n(a)\| \leq \sup_{z \in [a, x]} (Df_m(z) - Df_n(z)) \cdot \|x - a\|.$$

d'où :

$$\|f_m(x) - f_n(x)\| \leq \|f_m(a) - f_n(a)\| \sup_{z \in [a, x]} (Df_m(z) - Df_n(z)) \cdot \|x - a\|.$$

Soit $\epsilon > 0$. Puisque la suite $(f_n(a))_{n \geq 0}$ converge, elle est de Cauchy. Il existe donc N_1 tel que, pour tous entiers $n \geq N_1$ et $m \geq N_1$, on a :

$$\|f_m(a) - f_n(a)\| \leq \frac{\epsilon}{2}.$$

D'autre part, la suite $(Df_n)_{n \geq 0}$ étant uniformément convergente, elle est également de Cauchy. Il existe donc N_2 tel que, pour tous entiers $n \geq N_2$ et $m \geq N_2$, on a :

$$\sup_{z \in \Omega} (Df_m(z) - Df_n(z)) \leq \frac{\epsilon}{2M}.$$

(Attention, cela n'est vrai que si la convergence de $(Df_n)_{n \geq 0}$ est uniforme... c'est une hypothèse fondamentale!)

Ainsi, puisque $\|x - a\| \leq M$, on a :

$$\sup_{z \in [a, x]} (Df_m(z) - Df_n(z)) \cdot \|x - a\| \leq \frac{\epsilon}{2}.$$

On en déduit donc que lorsque n et m sont plus grands que $\max(N_1, N_2)$, on a :

$$\|f_m(x) - f_n(x)\| \leq \epsilon.$$

La suite $(f_n(x))_{n \geq 0}$ est donc une suite de Cauchy pour la norme uniforme sur $\Omega \cap B(a, M)$. Puisque F est complet et le réel $M > 0$ arbitraire, ceci prouve que pour tout $x \in \Omega$, la suite $(f_n(x))_{n \geq 0}$ converge vers un élément $f(x) \in F$ et que la convergence est uniforme sur toute partie bornée de Ω .

Intéressons-nous maintenant à la différentiabilité de f . Pour tous $x, y \in \Omega$ et pour tous entiers n et m , on a :

$$\begin{aligned} f_m(y) - f_m(x) - g(x)(y - x) &= (f_m - f_n)(y) - (f_m - f_n)(x) \\ &\quad + f_n(y) - f_n(x) - Df(x)(y - x) + (Df(x) - g)(y - x). \end{aligned}$$

On en déduit que :

$$\begin{aligned} \|f_m(y) - f_m(x) - g(x)(y - x)\| &\leq \|(f_m - f_n)(y) - (f_m - f_n)(x)\| \\ &\quad + \|f_n(y) - f_n(x) - Df(x)(y - x)\| + \|(Df(x) - g)\| \|y - x\|. \end{aligned} \quad (\text{III.2})$$

Nous allons majorer les différents termes à droite de cette dernière inégalité.

Le premier terme à droite dans l'équation III.2 se majore grâce au théorème des accroissements finis appliqué à la fonction $(f_m - f_n)$:

$$\|(f_m - f_n)(y) - (f_m - f_n)(x)\| \leq \sup_{z \in [x, y]} \|Df_m(z) - Df_n(z)\| \cdot \|y - x\|.$$

On considère maintenant que le point x est fixé et on choisit arbitrairement $\epsilon > 0$. Puisque la suite $(Df_n)_{n \geq 0}$ converge uniformément, il existe N_3 tel que, pour tous entiers $n \geq N_3$ et $m \geq N_3$, on a :

$$\sup_{z \in [x, y]} \|Df_m(z) - Df_n(z)\| \leq \frac{\epsilon}{3}.$$

Ainsi,

$$\|(f_m - f_n)(y) - (f_m - f_n)(x)\| \leq \frac{\epsilon}{3} \|y - x\|.$$

Puisque la suite $(Df_n(x))_{n \geq 0}$ converge vers $g(x)$, il existe N_4 tel que, pour tous entiers $n \geq N_4$ et $m \geq N_4$, on a :

$$\|Df_n(x) - g(x)\| \leq \frac{\epsilon}{3}.$$

Fixons maintenant un entier $n \geq \max(N_3, N_4)$. D'après la définition de la différentielle de f_n au point x , il existe $\eta > 0$ tel que, si $\|y - x\| < \eta$, alors

$$\|f_n(y) - f_n(x) - Df_n(x)(y - x)\| \leq \frac{\epsilon}{3} \|y - x\|.$$

En injectant ces derniers résultats dans l'équation III.2, on obtient :

$$\|f_m(y) - f_m(x) - g(x)(y - x)\| \leq \epsilon \|y - x\|.$$

Si on fait tendre alors m vers l'infini, on obtenons alors l'inégalité :

$$\|f(y) - f(x) - g(x)(y - x)\| \leq \epsilon \|y - x\| ,$$

et on en déduit que f est différentiable en tout point $x \in \Omega$ et que sa différentielle au point x est $g(x)$, ce qui termine la preuve.

Chapitre IV

Différentielles d'ordres supérieurs et extrema locaux

Lorsqu'on étudie une fonction réelle de plusieurs variables, il peut être intéressant, comme pour les fonctions de la variable réelle, de pouvoir avoir une idée de la représentation de la fonction. Comme pour les fonction de $\mathbb{R}$ dans $\mathbb{R}$, une étape importante consiste à connaître les points extrémaux. Cette problématique apparait souvent dans des problèmes physiques. Par exemple, on cherche à minimiser des coûts de production qui dépends de plusieurs facteurs (nb pieces totales produites, cout de la main d'oeuvre, taille de l'installation, etc...) ou maximiser des quantités....

La recherche des points extrémaux passe par l'étude des différentielles successives d'une fonction.

IV.1 différentielle seconde

Définition IV.1. Soient E et F des evn, Ω un ouvert de E et $f : \Omega \rightarrow F$ une application différentiable sur Ω . On considère la différentielle de f :

$$Df : \Omega \rightarrow \mathcal{L}(E; F)$$

On dit que f est **deux fois différentiable en** $a \in \Omega$ si :

- f est différentiable sur un voisinage de a ,
- Df est différentiable en a .

Dans ce cas, on note $D^2f(a)$ la différentielle de Df en a et on l'appelle la **différentielle seconde de f en a** .

Si f est deux fois différentiable en tout point de Ω et si l'application $D^2f : \Omega \rightarrow \mathcal{L}(E; \mathcal{L}(E; F))$ est continue sur Ω , on dit que f est **de classe $\mathcal{C}^2$ sur Ω** .

Vu sous cet angle, il n'est pas très facile de comprendre ce qu'est $D^2f(a)$... on va donc utiliser la remarque suivante :

La différentielle seconde d'une application est une application de Ω dans $\mathcal{L}(E; \mathcal{L}(E; F))$ Or l'espace $\mathcal{L}(E; \mathcal{L}(E; F))$ est en fait isomorphe à l'ensemble des applications bilinaire continues de E dans $F : \mathcal{L}^2(E; F)$. En effet, l'application

$$\Phi : \mathcal{L}(E; \mathcal{L}(E; F)) \rightarrow \mathcal{L}^2(E; F)$$

définie pour tout $u \in \mathcal{L}(E; \mathcal{L}(E; F))$ par $\Phi(u)(h, k) = u(h)(k)$ pour tout $(h, k) \in E^2$ est un isomorphisme d'espaces vectoriels (une application linéaire bijective, quoi...).

Ainsi, pour tout $a \in \Omega$ où f est deux fois différentiable, $D^2f(a)$ est essentiellement une application bilinéaire continue de E dans F , d'où la reformulation suivante de la définition :

Définition IV.2. Soient E et F deux evn, Ω un ouvert de E et $f : \Omega \rightarrow F$ une application différentiable sur un voisinage ouvert de $a \in \Omega$.

On dit que f est **deux fois différentiable en** a s'il existe une application bilinéaire continue $B_a \in \mathcal{L}^2(E; F)$ telle que, pour tout $(h, k) \in E^2$ avec $\|h\|$ assez petite, on a :

$$Df(a+h)(k) = Df(a)(k) + B_a(h, k) + o(\|h\| \|(h, k)\|).$$

Dans ce cas, on dit que B_a est la différentielle de f en a est elle est notée $D^2f(a)$.

A priori, dans la définition même de la différentielle seconde en un point a , le rôle des éléments de E (h et k) ne sont pas symétriques... cependant, le résultat suivant permet de dire qu'en fait, leur rôle est parfaitement symétrique :

Théorème IV.3. Lemme de Schwarz

Soient E et F deux evn, Ω un ouvert de E et $f : \Omega \rightarrow F$ une application deux fois différentiable en $a \in \Omega$.

L'' application bilinéaire continue $D^2f(a)$ est symétrique :

$$\forall (h, k) \in E^2, D^2f(a)(h, k) = D^2f(a)(k, h).$$

Démonstration du Théorème IV.3

Soient E et F deux evn, Ω un ouvert de E et $f : \Omega \rightarrow F$ une application deux fois différentiable en $a \in \Omega$.

Au voisinage de $(0, 0) \in \mathbb{R}^2$, on définit la fonction g à valeurs dans F de la manière suivante :

$$g(x, y) = f(a + x \cdot h + y \cdot k) - f(a + x \cdot h) - f(a + y \cdot k) + f(a).$$

La fonction g est deux fois différentiable puisque f et les applications de $\mathbb{R}^2$ dans E définies par $(x, y) \mapsto a + x \cdot h + y \cdot k$, $(x, y) \mapsto a + y \cdot k$ et $(x, y) \mapsto a + x \cdot h$ le sont. Sur le même voisinage de $(0, 0)$, les dérivées partielles de g sont notées de la manière suivante :

$$g_1(x, y) = \frac{\partial g}{\partial x}(x, y) \quad \text{et} \quad g_2(x, y) = \frac{\partial g}{\partial y}(x, y).$$

On a les égalités suivantes :

$$g_1(x, y) = Df(a + x \cdot h + y \cdot k)(h) - Df(a + x \cdot h)(h),$$

$$g_2(x, y) = Df(a + x \cdot h + y \cdot k)(k) - Df(a + y \cdot k)(k).$$

On a alors $g_1(x, 0) = 0$ et $g_2(0, y) = 0$ et de plus :

$$\frac{\partial g_1}{\partial y}(x, y) = D^2f(a + x \cdot h + y \cdot k)(h)(k) = D^2f(a + x \cdot h + y \cdot k)(h, k),$$

$$\frac{\partial g_2}{\partial x}(x, y) = D^2 f(a + x \cdot h + y \cdot k)(k)(h) = D^2 f(a + x \cdot h + y \cdot k)(k, h).$$

Ainsi,

$$\frac{\partial g_1}{\partial y}(0, 0) = D^2 f(a)(h, k) \quad \text{et} \quad \frac{\partial g_2}{\partial x}(0, 0) = D^2 f(a)(k, h).$$

On va donc montrer que ces deux quantités sont les mêmes...

En fait, on va montrer que, si on fixe $\epsilon > 0$, alors la différence entre ces deux quantités est toujours plus petite que ϵ .

Soit $\epsilon > 0$. On pose $A = \frac{\partial g_1}{\partial y}(0, 0)$ et $B = \frac{\partial g_2}{\partial x}(0, 0)$.

Par définition des dérivées partielles, on a pour y suffisamment proche de 0 :

$$\left\| g_1(x, y) - g_1(x, 0) - \frac{\partial g_1}{\partial y}(0, 0)y \right\| \leq \frac{\epsilon}{2}|y|,$$

d'où :

$$\|g_1(x, y) - Ay\| \leq \frac{\epsilon}{2}|y|. \quad (\text{IV.1})$$

Si on pose $f_y(x) = g(x, y) - Axy$, alors $f'_y(x) = g_1(x, y) - Ay$. On peut donc, grâce à l'inégalité IV.1 appliquer le théorème de l'inégalité de la moyenne à f_y au voisinage de 0 et donc, pour x suffisamment petit, on a :

$$\|f_y(x) - f_y(0)\| = \|g(x, y) - Axy\| \leq \frac{\epsilon}{2}|xy|.$$

En utilisant exactement le même raisonnement autour de g_2 , par définition des dérivées partielles, on a pour x suffisamment proche de 0 :

$$\left\| g_2(x, y) - g_2(0, y) - \frac{\partial g_2}{\partial x}(0, 0)x \right\| \leq \frac{\epsilon}{2}|x|,$$

d'où :

$$\|g_2(x, y) - Bx\| \leq \frac{\epsilon}{2}|x|. \quad (\text{IV.2})$$

Si on pose $f_x(y) = g(x, y) - Bxy$, alors $f'_x(y) = g_2(x, y) - Bx$. On peut donc, grâce à l'inégalité IV.2 appliquer le théorème de l'inégalité de la moyenne à f_x au voisinage de 0 et donc, pour y suffisamment petit, on a :

$$\|f_x(y) - f_x(0)\| = \|g(x, y) - Bxy\| \leq \frac{\epsilon}{2}|xy|.$$

Puisque qu'on a, pour x et y suffisamment proche de 0 les deux inégalités :

$$\|g(x, y) - Axy\| \leq \frac{\epsilon}{2}|x||y| \quad \text{et} \quad \|g(x, y) - Bxy\| \leq \frac{\epsilon}{2}|x||y|,$$

on en déduit que $\|Axy - Bxy\| \leq \epsilon|xy|$ et donc que $\|A - B\| \leq \epsilon$, et cela, pour tout $\epsilon > 0$. On a donc $A = B$, i.e. $D^2 f(a)(h, k) = D^2 f(a)(k, h)$.

C'est bien joli, mais en pratique, c'est le corollaire suivant dont on se sert le plus, qui permet de donner l'expression de la différentielle seconde en fonction des dérivées partielles secondes :

Corollaire IV.4. Soient $E_1, E_2, \dots, E_n$ et F des evn, Ω un ouvert de $E = E_1 \times E_2 \times \dots \times E_n$.
Si $f : \Omega \rightarrow F$ est une application deux fois différentiable en $a = (a_1, a_2, \dots, a_n) \in \Omega$, alors :

$$\forall (i, j) \in \{1, \dots, n\}^2, \frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a).$$

Ainsi, quelque soient les éléments $h = (h_1, \dots, h_n)$ et $k = (k_1, \dots, k_n)$ de E , on a la **formule explicite de la différentielle seconde en fonction des dérivées partielles** :

$$D^2 f(a)(h, k) = \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i \partial x_i}(a)(h_i k_i) + \sum_{1 \leq i < j \leq n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a)(h_i k_j + h_j k_i).$$

La matrice $H(f)(a) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right)_{i, j \in \{1, \dots, n\}}$ est appelée **matrice Hessienne de f en a** . C'est une matrice symétrique si f est deux fois différentiable.

IV.2 différentielles d'ordre supérieur

De la même manière qu'on a procédé pour définir la notion de fonction deux fois différentiable, on procède par récurrence pour définir la notion de fonction n fois différentiable à partir d'une fonction $n - 1$ fois différentiable.

Pour tout n , on note $\mathcal{L}^n(E; F)$ l'ensemble des applications n -linéaires de E dans F . De la même manière qu'au paragraphe précédent, on a un isomorphisme d'espaces vectoriels entre $\mathcal{L}^{n+1}(E; F)$ et $\mathcal{L}^n(E; \mathcal{L}^n(E; F))$.

Définition IV.5. Soit $n \geq 2$. On dit que f est n fois différentiable en $a \in \Omega$ si

- il existe un ouvert $\Omega' \subset \Omega$ contenant a sur lequel f est $n - 1$ fois différentiable,
- $D^n f : \Omega' \rightarrow \mathcal{L}^{n-1}(E; F)$ est différentiable en a .

Dans ce cas, on note $D^n f(a) \in \mathcal{L}^n(E; F)$ sa différentielle et on l'appelle la **différentielle n -ième de f en a** .

Si f est n fois différentiable en tout point de Ω et si l'application $D^n f : \Omega \rightarrow \mathcal{L}(E; \mathcal{L}^n(E; F))$ est continue sur Ω , on dit que f est **de classe $\mathcal{C}^n$ sur Ω** .

Remarque : On peut évidemment avoir l'expression de la différentielle p -ième d'une application $f : \mathbb{R}^n \rightarrow \mathbb{R}$ en fonction des p -ièmes dérivées partielles de f :

Pour tout $(h^{(1)}, h^{(2)}, \dots, h^{(p)}) \in (\mathbb{R}^n)^p$:

$$D^p f(a)(h^{(1)}, h^{(2)}, \dots, h^{(p)}) = \sum_{1 \leq i_1 < i_2 < \dots < i_p \leq n} \frac{\partial^p f}{\partial x_{i_1} \dots \partial x_{i_p}}(a) h_{i_1}^{(1)} h_{i_2}^{(2)} \dots h_{i_p}^{(p)}.$$

Cette formule peut être modifiée en tenant compte du fait que certaines des n -ièmes dérivées partielles sont identiques, grâce au Lemme de Schwarz.

IV.3 formule de Taylor-Young

Théorème IV.6. Soient E et F deux evn, Ω un ouvert de E et $f : \Omega \rightarrow F$ une application n fois différentiable au point $a \in \Omega$.

Pour $h \in E$, on note $h^{(p)}$ le vecteur de E^p formé de copies de $h : \forall i \in \{1, \dots, p\}, h_i^{(p)} = h$.

$$\forall x \in \Omega, f(x) = f(a) + \sum_{p=1}^n \frac{1}{p!} D^p f(a) \left((x-a)^{(p)} \right) + o \left(\left\| (x-a)^{(n)} \right\| \right).$$

Ce théorème se montre par récurrence sur n . Pour $n = 1$ c'est la définition de la différentielle en a . Le passage du rang n au rang $n+1$ se fait en montrant que la différence $\delta(h) = f(a+h) - f(a) - \sum_{p=1}^n \frac{1}{p!} D^p f(a) \left((h)^{(p)} \right)$ est un $o \left(\left\| (x-a)^{(n+1)} \right\| \right)$. On applique l'hypothèse de récurrence à $D\delta$ et le théorème de l'inégalité de la moyenne à δ .

IV.4 Points critiques et extrema locaux

Définition IV.7. Soient E et F deux evn, Ω un ouvert de E et $f : \Omega \rightarrow F$ une application différentiable sur Ω .

On appelle **point critique** un point $a \in \Omega$ pour lequel $Df(a) = 0$.

Définition IV.8. Soient E un evn, Ω un ouvert de E et $f : \Omega \rightarrow \mathbb{R}$ différentiable sur Ω .

On dit que f admet un **minimum local** (resp. **maximum local**) en $a \in \Omega$ s'il existe un ouvert $U \subset \Omega$ contenant a et tel que :

$$\forall x \in U, f(x) \geq f(a) \text{ (resp. } f(x) \leq f(a) \text{)}.$$

Si les inégalités sont strictes, on parle de minimum (resp. maximum) local strict.

Un minimum ou un maximum local est appelé **extrema local**.

Théorème IV.9. Soient E un evn, Ω un ouvert de E et $f : \Omega \rightarrow \mathbb{R}$ différentiable sur Ω .

Si $a \in \Omega$ est un extrema local, alors a est un point critique : $Df(a) = 0$.

Démonstration Laissée en exercice.

Définition IV.10. Soient E un evn, Ω un ouvert de E et $f : \Omega \rightarrow \mathbb{R}$ différentiable sur Ω .

Un point critique qui n'est pas un extrema local est appelé **point-selle**.

Ci dessous, un résultat qui permet d'identifier certains points critiques comme étant des extrema locaux.

Théorème IV.11. Soient E un evn, Ω un ouvert de E et $f : \Omega \rightarrow \mathbb{R}$ différentiable sur Ω .

On suppose que a est un point critique de f et on note la forme quadratique $q : E \rightarrow \mathbb{R}$ définie par

$$\forall h \in E, q(h) = D^2 f(a)(h, h).$$

— Si q est définie positive (i.e. $\forall h \in E, h \neq 0 \implies q(h) > 0$), alors f admet en a un minimum local strict.

- Si q est définie négative (i.e. $\forall h \in E, h \neq 0 \implies q(h) < 0$), alors f admet en a un maximum local strict.
- S'il existe $(h, k) \in E^2$ tels que $q(h) > 0$ et $q(k) < 0$, alors a est un point-selle.

Remarque 1 : Si la forme quadratique q est non nulle et juste positive ($\forall h \in E, q(h) \leq 0$) alors f admet en a un minimum local et si la forme quadratique q est non nulle et juste négative ($\forall h \in E, q(h) \geq 0$) alors f admet en a un maximum local.

Remarque 2 : Si $\text{Diff}^2 f(a) = 0$, alors on ne peut pas conclure et il est nécessaire d'aller regarder les différentielles d'ordres supérieurs pour décider si oui ou non a est un extrema local...

Corollaire IV.12. Cas $E = \mathbb{R}^2$: Formule de Lagrange (1759)

Soit Ω un ouvert de $\mathbb{R}^2$ et soit $f : \Omega \rightarrow \mathbb{R}$ une application deux fois différentiable en $a \in \Omega$ telle que $\text{D}f(a) = 0$. On pose :

$$r = \frac{\partial^2 f}{\partial x \partial x}, \quad s = \frac{\partial^2 f}{\partial x \partial y}, \quad t = \frac{\partial^2 f}{\partial y \partial y}.$$

- Si $rt - s^2 > 0$, alors f admet au point a un extremum local strict. si $r > 0$ alors c'est un minimum et si $r < 0$, c'est un maximum.
- Si $rt - s^2 < 0$, alors f n'admet pas d'extremum local en a .

Chapitre V

Inversion locale et théorème des fonctions implicites

V.1 Problématique et isomorphismes topologiques...

L'idée de ce chapitre est d'étendre le résultat classique qui dit que si $f : \mathbb{R} \rightarrow \mathbb{R}$ est une application continuellement dérivable sur un voisinage ouvert de $a \in \mathbb{R}$ et que $f'(a) \neq 0$, alors f est inversible au voisinage de a : elle admet une fonction réciproque f^{-1} , définie au voisinage de $f(a)$, dérivable sur un voisinage de $f(a)$ et telle que : $(f^{-1})'(f(a)) = \frac{1}{f'(a)}$.

La recherche d'une application réciproque (au moins au voisinage d'un point) est un problème récurrent en physique, pour lesquels on a régulièrement des informations sur les "observations" (les images de la fonctions) plutôt que sur les paramètres (les variables). Il serait donc bien utile d'avoir l'équivalent de ce résultat dans le cadre plus large que nous avons fixé.

Une première question se pose naturellement : comment étendre la condition $f'(a) \neq 0$ que l'on a pour les fonctions réelles de la variable réelle à des fonctions d'un evn dans un autre... suffit-il de supposer de $Df(a) \neq 0$ pour que f^{-1} existe et soit différentiable en $f(a)$?? eh bien non... ca serait un peu trop facile...

Par exemple, la fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ définie simplement par $f(x, y) = (x^3, y)$ est inversible sur $\mathbb{R}^2$: $f^{-1} : (x, y) \mapsto (x^{\frac{1}{3}}, y)$ et f admet une différentielle Df non nulle pour tout $(a, b) \in \mathbb{R}^2$:

$$Df(a, b) : (h, k) \mapsto \begin{pmatrix} 3a^2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} h \\ k \end{pmatrix} = (3a^2h, k).$$

Si $(a, b) \neq (0, 0)$, f^{-1} est différentiable et on a

$$D(f^{-1})(a, b) : (h, k) \mapsto \begin{pmatrix} \frac{1}{3}a^{-\frac{2}{3}} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} h \\ k \end{pmatrix} = (\frac{1}{3}a^{-\frac{2}{3}}h, k),$$

mais lorsque a tends vers 0, le premier terme tends vers l'infini et donc f^{-1} n'est pas différentiable en $(0, 0)$ même si $Df(0, 0) \neq 0$.

Il faut donc une condition plus forte... en fait, **dans le cas des fonctions réelles de la variable réelle, il y a 2 choses implicites à la condition $f'(a) \neq 0$: d'une part, la bijectivité de f qui permet bien évidemment la définition de la fonction inverse, mais elle permet aussi l'existence de $\frac{1}{f'(a)}$, qui est la valeur de la dérivée de f^{-1} en $f(a)$.**

Supposons que l'on soit dans le cadre que nous avons fixé jusqu'à présent : E et F deux $\mathbb{K}$ -evn, Ω un ouvert de E et une application $f : \Omega \rightarrow F$ pour laquelle tout se passe bien : elle est inversible sur Ω , $f(\Omega) = V$ et son inverse $f^{-1} : V \rightarrow \Omega$ est sympathique (on mettra des vraies conditions plus tard pour remplacer les mots en italique!). On aurait alors :

$$f \circ f^{-1} = Id_{\Omega} \quad \text{et} \quad f^{-1} \circ f = Id_V.$$

Comme tout est sensé bien se passer (parcequ'on l'a décidé...), on peut différentier ces relations (en utilisant la règle de différentiation des applications composées), et on obtiendrait, pour tout $x \in \Omega$:

$$D(f^{-1})(f(x)) \circ Df(x) = Id_{\Omega} \quad \text{et} \quad Df(x) \circ D(f^{-1})(f(x)) = Id_V. \quad (\text{V.1})$$

Voici donc le constat : Hormis le fait que l'on doit garantir l'existence de la fonction réciproque (via la bijectivité de f), si la différentielle de f^{-1} existe en $b = f(a)$, c'est l'inverse de Df au point $a = f^{-1}(b)$... Il faut donc que $Df(a)$ **soit un élément inversible de $\mathcal{L}(E; F)$** , c'est-à-dire un isomorphisme continu de E dans F **dont l'inverse est également continu** (pour que ce soit aussi une différentielle) : c'est ce qu'on appelle un **isomorphisme topologique**...

Définition V.1. Soient E et F deux $\mathbb{K}$ -evn.

Une application $L \in \mathcal{L}(E; F)$ est un **isomorphisme topologique** si L est bijective et $L^{-1} \in \mathcal{L}(F; E)$.

On note $\mathcal{I}som(E; F)$ l'ensemble des isomorphismes topologiques de E dans F .

Si $\mathcal{I}som(E; F) \neq \emptyset$, on dira que E et F , munis des topologies associées à leurs normes sont **topologiquement isomorphes**.

Remarque : La notion d'isomorphisme topologique traduit mathématiquement le fait que E et F sont tellement "semblables" qu'en plus de pouvoir passer de l'un à l'autre en le déformant de manière continue, la déformation "transporte" la topologie (l'image d'un voisinage d'un point est un voisinage de l'image du point, l'image d'un ouvert est un ouvert, etc...).

Exemple : L'isomorphisme topologique le plus utilisé (et tellement utilisé qu'on l'oublie presque...) c'est celui qui existe entre $\mathbb{R}^2$ muni de la norme euclidienne et $\mathbb{C}$ muni du module.

Il serait donc agréable d'avoir un cadre où, une application linéaire bijective et continue est nécessairement d'inverse continue. C'est le cas des espaces de Banach : lorsque les evn qui entrent en jeu sont complets, on a le **Théorème de l'isomorphisme de Banach** :

Théorème V.2. Soient E et F deux espaces de Banach (evn complets).

Si $L \in \mathcal{L}(E; F)$ est bijective alors $L^{-1} \in \mathcal{L}(F; E)$.

De plus, l'application $\Phi : L \mapsto L^{-1}$ est une application continue de $\mathcal{L}(E; F)$ dans lui-même.

Attention : Evidemment, en dimension finie, on aura aucun problème car les evn de dimension finie sont complets : toute application linéaire continue bijective sera d'inverse continu.

L'hypothèse de complétude joue vraiment son rôle lorsque l'on est en dimension infinie. Certains problèmes peuvent apparaître si les espaces dans lesquels on travaille ne sont pas complets. Ci-dessous, un exemple d'application linéaire continue bijective dont l'inverse est linéaire mais pas continue.

Exercice V.3. Soit S_{00} l'ev des suites nulles à partir d'un certain rang, muni de la norme 1 : $\|(u_n)_{n \geq 0}\|_1 = \sum_{n \geq 0} |u_n|$. Soit S_0 l'evn des suites dont la limite est nulle, pour lesquels $\|(u_n)_{n \geq 0}\|_1 < \infty$.

1. Pour tout entier $p \geq 0$, on définit la suite $\mathbf{u}_p = (1, \frac{1}{2^2}, \frac{1}{3^2}, \dots, \frac{1}{p^2}, 0, 0, \dots)$. Montrer que la suite $(\mathbf{u}_p)_{p \geq 0}$ converge dans S_0 vers : $\mathbf{u} = (1, \frac{1}{2^2}, \frac{1}{3^2}, \dots, \frac{1}{n^2}, \dots)$.
En déduire que S_{00} n'est pas un espace vectoriel normé complet lorsqu'il est muni de la norme 1.
2. Soit $L : S_{00} \rightarrow S_{00}$ l'application définie par $L((u_n)_{n \geq 0}) = \left(\frac{u_n}{n+1}\right)_{n \geq 0}$. Montrer que L est une application linéaire, bijective et continue. En déduire sa différentielle.
3. Montrer que L^{-1} n'est pas continue.

Dans la suite de ce chapitre, on travaillera donc uniquement avec des espaces de Banach.

V.2 Homéomorphismes et difféomorphismes de classe $\mathcal{C}^1$

Définition V.4. Soient E et F deux $\mathbb{K}$ -evn, Ω un ouvert de E et une application $f : \Omega \rightarrow F$.

On dit que f est un **homéomorphisme de Ω vers un ouvert V de F** si :

1. f est continue sur Ω ;
2. f est une bijection de Ω vers V ;
3. l'application réciproque $f^{-1} : V \rightarrow \Omega$ est continue sur V .

Définition V.5. Soient E et F deux $\mathbb{K}$ -evn, Ω un ouvert de E et une application $f : \Omega \rightarrow F$.

On dit que f est un **difféomorphisme de Ω vers un ouvert V de F** si :

1. f est différentiable sur Ω ;
2. f est une bijection de Ω vers V ;
3. l'application réciproque $f^{-1} : V \rightarrow \Omega$ est différentiable sur V .

On dit également que f est un **$\mathcal{C}^1$ -difféomorphisme** ou **difféomorphisme de classe $\mathcal{C}^1$** si f est un difféomorphisme et que f et f^{-1} sont toutes les deux de classe $\mathcal{C}^1$.

Attention : Un difféomorphisme de Ω vers V est un homéomorphisme (bijection continue, de fonction réciproque continue). Cependant, un homéomorphisme différentiable de Ω vers V n'est pas forcément un difféomorphisme : Il faut que la fonction réciproque soit également différentiable.

Par exemple : $f : x \mapsto x^3$ est un homéomorphisme de $\mathbb{R}$ dans $\mathbb{R}$ différentiable, mais ce n'est pas un difféomorphisme : $f^{-1} : x \mapsto x^{\frac{1}{3}}$ n'est pas différentiable en 0...

Proposition V.6. Soient E et F deux $\mathbb{K}$ -evn complets, Ω un ouvert de E et V un ouvert de F et $f : \Omega \rightarrow V$ un homéomorphisme différentiable en $a \in \Omega$.

f^{-1} est différentiable en a si et seulement si $Df(a) \in \mathcal{I}som(E; F)$.

Dans ce cas, on a

$$D(f^{-1})(f(a)) = (Df(a))^{-1}.$$

Remarque : Cette proposition étend de manière simple la proposition usuelle pour les fonctions de $\mathbb{R}$ dans $\mathbb{R}$ suivante :

Si $f : \mathbb{R} \rightarrow \mathbb{R}$ est dérivable en a , bijective et de réciproque f^{-1} continue en a , alors $(f^{-1})'(f(a))$ existe si et seulement si $f'(a) \neq 0$ et dans ce cas, on a : $(f^{-1})'(f(a)) = \frac{1}{f'(a)}$.

La proposition V.6 admet une version globale suivante :

Proposition V.7. Soient E et F deux $\mathbb{K}$ -evn complets, Ω un ouvert de E et V un ouvert de F et $f : \Omega \rightarrow V$ un homéomorphisme de classe $\mathcal{C}^1$.

f est un $\mathcal{C}^1$ -difféomorphisme de Ω sur V si et seulement si pour tout $x \in \Omega$, $Df(x)$ est un isomorphisme de E sur F .

Dans ce cas, on a

$$\forall x \in \Omega, D(f^{-1})(f(x)) = (Df(x))^{-1}.$$

A propos de ces deux dernières propositions... Dans les hypothèses de ces deux résultats, on suppose que f est un homéomorphisme. C'est une condition assez forte, puisqu'elle force la bijectivité et suppose donc l'existence et la continuité de l'application inverse.

Preuve de la propriété V.6 :

On peut tenir le raisonnement effectué au paragraphe précédent pour montrer une des implications. Si f^{-1} est différentiable en a on a, sur un voisinage de a :

$$f \circ f^{-1} = Id \quad \text{et} \quad f^{-1} \circ f = Id.$$

Lorsqu'on différencie ces relations (en utilisant la règle de différentiation des applications composées), on obtient :

$$D(f^{-1})(f(a)) \circ Df(a) = Id \quad \text{et} \quad Df(a) \circ D(f^{-1})(f(a)) = Id,$$

donc $Df(a)$ est un isomorphisme de E sur F et son inverse est $D(f^{-1})(f(a))$.

Montrons que la condition $Df(a)$ est un isomorphisme de E sur F est suffisante pour que f^{-1} soit différentiable en a .

Par hypothèse, on a : $f(a+h) = f(a) + Df(a)(h) + \|h\| \epsilon(h)$ avec $\lim_{h \rightarrow 0} \epsilon(h) = 0$. On peut réécrire cela ainsi :

$$Df(a)(h) = f(a+h) - f(a) - \|h\| \epsilon(h).$$

En utilisant notre hypothèse, on a donc :

$$h = (Df(a))^{-1}(f(a+h) - f(a)) - \|h\| (Df(a))^{-1}(\epsilon(h)).$$

En posant $b = f(a)$ et $y = f(a + h)$, on a $h = f^{-1}(y) - f^{-1}(b)$ et donc

$$f^{-1}(y) - f^{-1}(b) = (Df(a))^{-1} (y - b) - \|f^{-1}(y) - f^{-1}(b)\| (Df(a))^{-1} (\epsilon (f^{-1}(y) - f^{-1}(b))). \quad (\text{V.2})$$

Pour montrer la différentiabilité de f^{-1} en b , on doit donc montrer que :

$$\delta(y - b) = \|f^{-1}(y) - f^{-1}(b)\| (Df(a))^{-1} (\epsilon (f^{-1}(y) - f^{-1}(b))) = o(\|y - b\|)$$

D'une part, puisque f^{-1} est continue (f est un homéomorphisme), on a

$$\lim_{y \rightarrow b} \epsilon (f^{-1}(y) - f^{-1}(b)) = 0$$

D'autre part, on tire de l'égalité V.2 l'inégalité suivante :

$$\|f^{-1}(y) - f^{-1}(b)\| \left(1 - \|(Df(a))^{-1}\| \|\epsilon (f^{-1}(y) - f^{-1}(b))\|\right) \leq \|(Df(a))^{-1}\| \|y - b\|$$

On en déduit alors qu'il existe une constante K telle que, pour y suffisamment proche de b , on a :

$$\|f^{-1}(y) - f^{-1}(b)\| \leq K \|y - b\|$$

. En effet, pour $\epsilon > 0$ fixé et y suffisamment proche de b , $\|\epsilon (f^{-1}(y) - f^{-1}(b))\| \leq \epsilon$, donc

$$\|f^{-1}(y) - f^{-1}(b)\| \leq \frac{\|(Df(a))^{-1}\|}{1 - \epsilon} \|y - b\|.$$

On a donc montré que $\delta(y - b) = o(\|y - b\|)$, ce qui permet de conclure à la différentiabilité de f^{-1} en $b = f(a)$ et donne l'égalité $(Df(a))^{-1} = D(f^{-1})(f(a))$.

Preuve de la propriété V.7 :

Si f est un $\mathcal{C}^1$ difféomorphisme, alors, f^{-1} est $\mathcal{C}^1$ et on a forcément $D(f^{-1})(f(x)) = (Df(x))^{-1}$ quelque soit $x \in \Omega$ puisque $f \circ f^{-1} = Id_F$ et $f^{-1} \circ f = Id_E$.

Pour la réciproque, on utilise la propriété V.6 pour montrer que f^{-1} est différentiable en tout point de Ω . Il reste à montrer que $D(f^{-1}) : V \rightarrow \mathcal{L}(F; E)$ est continue. Pour cela, il suffit de voir que $D(f^{-1})$ est la composition de 3 applications continues :

Des équations V.1 permettent de déduire que $D(f^{-1})(f(x)) = (Df(x))^{-1}$. Ainsi, pour tout $y \in V$,

$$D(f^{-1})(y) = (Df(f^{-1}(y)))^{-1}.$$

Donc $D(f^{-1})$ est la composée de f^{-1} qui est continue, Df qui l'est aussi, et enfin de l'application d'inversion $\Phi : \mathcal{I}som(E; F) \rightarrow \mathcal{I}som(E; F)$ définie par $\Phi(L) = L^{-1}$ qui est également continue (Théorème de l'isomorphisme de Banach).

En pratique, pour les fonctions de $\mathbb{R}^n$ dans $\mathbb{R}^p$:

Si la matrice jacobienne J d'un difféomorphisme $f : \mathbb{R}^n \rightarrow \mathbb{R}^p$ en un point x est inversible (son déterminant, le Jacobien, est non nul...), alors la jacobienne de f^{-1} en $f(a)$ est donnée par J^{-1} .

V.3 Le théorème d'inversion locale

V.3.1 Enoncé du théorème

Dans la propriété V.6, la continuité de $a \mapsto Df(a)$ et l'inversibilité de $Df(a)$ pour tout $a \in \Omega$ ne sont pas des conditions suffisantes pour faire d'une application différentiable un $\mathcal{C}^1$ -difféomorphisme... **la bijectivité est vraiment une condition nécessaire**, et (évidemment) il existe des applications remplissant les premières conditions qui ne sont pas bijectives :

Exemple V.8. Soit $f : \mathbb{C}^* \rightarrow \mathbb{C}^*$ définie par $f(z) = z^2$. ($\mathbb{C}$ est muni de sa structure de $\mathbb{R}$ -espace vectoriel.) La différentielle de $f : Df(z_0) : h \mapsto 2z_0h$ est inversible, d'inverse $(Df(z_0))^{-1} : h \mapsto \frac{h}{2z_0}$ qui est aussi inversible mais f n'est pas bijective.

Néanmoins, on peut dire des choses sur f , notamment que si $Df(a)$ est inversible et que f est de classe $\mathcal{C}^1$, alors "localement", f est un difféomorphisme... C'est le propos du théorème d'inversion locale.

Théorème V.9. (Théorème d'inversion locale)

Soient E et F deux espaces de Banach, Ω un ouvert de E et $a \in \Omega$.

Si $f : \Omega \rightarrow F$ est une application de classe $\mathcal{C}^1$ telle que $Df(a) \in \mathcal{I}som(E, F)$, alors il existe un voisinage ouvert V de a , inclus dans Ω tel que $W = f(V)$ soit un ouvert de F et que la restriction de f à V soit un $\mathcal{C}^1$ -difféomorphisme de V sur W .

Reformulation sympathique : En gros, cet énoncé dit que, si la différentielle d'une fonction f en un point a est inversible ($Df(a) \in \mathcal{I}som(E, F)$), alors lorsqu'on est proche de a , la fonction f (supposée $\mathcal{C}^1$) est inversible, et que son inverse est aussi $\mathcal{C}^1$ au voisinage de $f(a)$.

Exemple V.10. Voici un exemple d'application : Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ définie par $f(x, y) = (x + y^4, y + x^3y)$. La fonction f est un difféomorphisme local en $(0, 0)$ car la jacobienne de f en $(0, 0)$:

$$J(f)(0, 0) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

est inversible. Ainsi, si (a, b) est proche de $f(0, 0) = (0, 0)$, alors le système d'équations :

$$\begin{cases} x + y^4 &= a \\ y + x^3y &= b \end{cases}$$

admet une unique solution $(x(a, b), y(a, b))$ qui dépend différentiablement de (a, b) et telle que $(x(0, 0), y(0, 0)) = (0, 0)$.

A défaut de permettre de résoudre totalement la question, le théorème d'inversion locale permet de dire qu'une solution existe et est unique, ce qui est déjà pas mal.... En attaquant de front le système d'équation, on a à résoudre l'équation du treizième degré $y + (a - y^4)^3y = b...$

V.3.2 Le théorème du point fixe et la preuve du théorème d'inversion locale

La preuve du théorème d'inversion locale utilise le théorème du point fixe dans les espaces métriques complets (d'où le souci d'être dans des espaces de Banach...)

Théorème V.11. Théorème du point fixe de Banach

Soient E un espace métrique complet, B un fermé de E et $f : B \rightarrow B$ une application contractante :

$$\exists k \in]0, 1[, \forall (x, y) \in B^2, \|f(x) - f(y)\| \leq k \|x - y\|$$

Il existe un unique point fixe $x \in B$ tel que $f(x) = x$.

Preuve du Théorème d'inversion locale (Théorème V.9) :

Soient E et F deux espaces de Banach, Ω un ouvert de E , $a \in \Omega$, et $f : \Omega \rightarrow F$ est une application de classe $\mathcal{C}^1$ telle que $Df(a) \in \mathcal{I}som(E, F)$.

Quitte à considérer la fonction $x \mapsto f(x) - f(a)$, on peut supposer que $f(a) = 0_F$ (cela permet de simplifier les calculs).

Pour montrer que f est localement un $\mathcal{C}^1$ -difféomorphisme, on va utiliser le théorème du point fixe et la Proposition V.7 sur un voisinage V de a sur lequel f est un homéomorphisme de classe $\mathcal{C}^1$ et tel que, pour tout $x \in V$, $Df(x) \in \mathcal{I}som(E, F)$.

On doit donc trouver un voisinage de a sur lequel f est bijective et d'inverse continue. Puisque f est $\mathcal{C}^1$, il suffira ensuite de vérifier qu'il existe un voisinage de a sur lequel on a $Df(x) \in \mathcal{I}som(E, F)$ pour conclure, grâce à la Proposition V.7 que f est un difféomorphisme de classe $\mathcal{C}^1$.

Ainsi, on doit montrer que, pour $y \in F$ fixé proche de $0_F = f(a)$, l'équation $f(x) = y$ admet une unique solution, qui donne la bijectivité de f (et donc l'existence de f^{-1}) sur un voisinage ouvert de a . On vérifiera ensuite que l'application f^{-1} définie ainsi est bien continue. Une fois cela fait, on vérifiera que la différentielle de f est un isomorphisme topologique en tout point de ce voisinage.

Première étape : Il existe un voisinage de a sur lequel l'application f est bijective

Soit $y \in F$ proche de $f(a) = 0_F$.

On considère la fonction $g_y : \Omega \rightarrow E$ définie comme suit :

$$\forall x \in E, g_y(x) = x - Df(a)^{-1} (f(x) - y).$$

Montrer l'existence et l'unicité d'un élément $x \in E$ tel que $f(x) = y$ revient à montrer l'existence et l'unicité d'un élément $x \in E$ tel que $g_y(x) = x$ car on a :

$$g_y(x) = x \iff x - Df(a)^{-1} (f(x) - y) = x \iff -Df(a)^{-1} (f(x) - y) = 0 \iff f(x) = y.$$

La dernière équivalence est due au fait que $Df(a)$ est un isomorphisme.

On va donc trouver un voisinage de a dans lequel on peut appliquer le théorème du point fixe de Banach à la fonction g_y ...

Puisque f est de classe $\mathcal{C}^1$, g_y l'est aussi. D'autre part, $Dg_y(x) = \text{Id}_E - Df(a)^{-1} \circ Df(x)$, donc $Dg_y(a) = 0_{\mathcal{L}(E, E)}$.

Puisque Dg_y est continue, pour tout $\epsilon > 0$ fixé, il existe donc un boule ouverte $B(a, R) \subset \Omega$, tel que

$$\forall x \in B(a, R), \|Dg_y(x)\| \leq \epsilon. \quad (\text{V.3})$$

On peut d'ailleurs choisir R suffisamment petit pour que $\overline{B}(a, R) \subset \Omega$. Notons que $Dg_y(x)$ ne dépend pas de y et donc que R non plus.

En appliquant l'inégalité de la moyenne à g_y dans $\overline{B}(a, R)$, on a :

$$\forall (x, x') \in \overline{B}(a, R), \|g_y(x) - g_y(x')\| \leq \epsilon \|x - x'\|. \quad (\text{V.4})$$

Donc l'application g_y est contractante sur $\overline{B}(a, R)$. En particulier, on a, pour tout $x \in \overline{B}(a, R)$:

$$\|g_y(x) - g_y(a)\| \leq \epsilon \|x - a\|.$$

Par conséquent, puisque $\|g_y(x) - a\| \leq \|g_y(x) - g_y(a)\| + \|g_y(a) - a\|$ et que $g_y(a) - a = Df(a)^{-1}(y)$, on a :

$$\|g_y(x) - a\| \leq \epsilon \|x - a\| + \|Df(a)^{-1}(y)\|,$$

En choisissant par exemple y dans la boule $\mathcal{W}_R = B\left(0_F, \frac{(1-2\epsilon)R}{\|Df(a)^{-1}\|}\right)$, alors y est suffisamment proche de $f(a) = 0_F$ pour que, par continuité, on ait $\|Df(a)^{-1}(y)\| \leq (1-2\epsilon)R$. On obtient donc :

$$\forall x \in \overline{B}(a, R), \|g_y(x) - a\| < (1-2\epsilon)R + \epsilon R \leq (1-\epsilon)R. \quad (\text{V.5})$$

Ainsi, Pour $x \in \overline{B}(a, R)$, on a $g_y(x) \in \overline{B}(a, R)$. On est donc dans les hypothèses du théorème du point fixe de Banach lorsque y est dans la boule $\mathcal{W}_R$. Il existe donc, pour tout $y \in \mathcal{W}$, un unique $\bar{x}_y \in \overline{B}(a, R)$ tel que $g_y(\bar{x}_y) = \bar{x}_y$.

En fait, on a même mieux car l'équation (V.5) nous assure que $\|g_y(\bar{x}_y) - a\| = \|\bar{x}_y - a\| \leq (1-\epsilon)R$. donc $\bar{x}_y$ est dans la boule ouverte $B(a, R)$. On a donc montré que pour tout $y \in \mathcal{W}$, il existe un unique $\bar{x}_y \in B(a, R)$ tel que $f(\bar{x}_y) = y$.

On peut donc définir f^{-1} sur $\mathcal{W}_R$ et donc la fonction f est bijective sur le voisinage ouvert de a défini par $\mathcal{V} = B(a, R) \cap f^{-1}(\mathcal{W}_R)$ (on se restreint à $B(a, R)$ pour être sûr d'être dans Ω).

Deuxième étape : Sur le voisinage ouvert $\mathcal{W}_R$ de $f(a) = 0_F$, l'application f^{-1} est continue.

Pour $(x, y) \in \mathcal{V} \times \mathcal{W}_R$, on note $G(x, y) = g_y(x) = x - Df(a)^{-1}(f(x) - y)$.

Soient y et y' deux éléments de $\mathcal{W}_R$. On pose $f^{-1}(y) = \bar{x}$ et $f^{-1}(y') = \bar{x}'$, de sorte que $G(\bar{x}, y) = \bar{x}$ et $G(\bar{x}', y') = \bar{x}'$. On a donc

$$\begin{aligned} \bar{x} - \bar{x}' &= G(\bar{x}, y) - G(\bar{x}', y') \\ &= G(\bar{x}, y) - G(\bar{x}', y) + G(\bar{x}', y) - G(\bar{x}', y') \\ &= g_y(\bar{x}) - g_y(\bar{x}') + Df(a)^{-1}(y - y') \end{aligned}$$

En appliquant l'inégalité de la moyenne à g_y sur $\mathcal{V}$ (ouvert inclus dans $\overline{B}(a, R)$) comme au début de la première étape (voir équation (V.4)), on obtient donc :

$$\|\bar{x} - \bar{x}'\| \leq \epsilon \|\bar{x} - \bar{x}'\| + \|Df(a)^{-1}\| \|y - y'\|,$$

c'est-à-dire :

$$(1 - \epsilon) \|\bar{x} - \bar{x}'\| \leq \|Df(a)^{-1}\| \|y - y'\|.$$

On en conclut donc que

$$\|f^{-1}(y) - f^{-1}(y')\| \leq \frac{\|Df(a)^{-1}\|}{(1 - \epsilon)} \|y - y'\|,$$

et donc que f^{-1} est lipschitzienne et donc continue sur $\mathcal{W}_R$.

Troisième étape : Pour tout élément $x \in \mathcal{V}$, $Df(x) \in \mathcal{I}som(E, F)$

Ce résultat est en fait une conséquence d'un résultat classique d'algèbre linéaire : $\mathcal{I}som(E, E)$ est un ouvert de $\mathcal{L}(E, E)$. On peut cependant le montrer assez vite à l'aide des séries d'applications linéaires (c'est d'ailleurs cet outil qui est utilisé pour montrer que $\mathcal{I}som(E, E)$ est un ouvert de $\mathcal{L}(E, E)$).

On a vu que $Dg_y(x) = \text{Id}_E - Df(a)^{-1} \circ Df(x)$, donc $Dg_y(a) = 0_{\mathcal{L}(E, E)}$. La continuité de Dg_y avait permis de définir $B(a, R) \subset \Omega$ sur laquelle $\|Dg_y(x)\| \leq \epsilon$ (voir équation (V.3)).

Puisque $\mathcal{V} \subset B(a, R)$, alors pour tout $x \in \mathcal{V}$, $\|Dg_y(x)\| \leq \epsilon$.

Pour simplifier la suite, on note $D = Dg_y(x) = \text{Id}_E - Df(a)^{-1} \circ Df(x)$. On a les égalités d'applications suivantes :

$$(\text{Id}_E - D) \circ \left(\sum_{n=0}^N D^n \right) = \left(\sum_{n=0}^N D^n \right) \circ (\text{Id}_E - D) = \text{Id}_E - D^{N+1}$$

Comme $\|D\| \leq \epsilon$, la série d'applications linéaires continues $\sum_{k \geq 0} D^k$ converge uniformément et sa somme S est donc une application linéaire continue qui a pour inverse $(\text{Id}_E - D) = Df(a)^{-1} \circ Df(x)$. Ainsi, $Df(a)^{-1} \circ Df(x)$ est inversible, d'inverse S et donc $Df(x)$ l'est aussi, d'inverse $S \circ Df(a)$. On a donc montré que pour tout $x \in \mathcal{V}$, $Df(x) \in \mathcal{I}som(E, F)$.

Conclusion :

La fonction f est $\mathcal{C}^1$ par hypothèse. Les deux premières étapes permettent d'affirmer que f est un homéomorphisme de classe $\mathcal{C}^1$ d'un voisinage ouvert $\mathcal{V}$ de a vers un voisinage ouvert $\mathcal{W}$ de $f(a) = 0_F$. La troisième étape affirme ensuite que, pour tout $x \in \mathcal{V}$, $Df(x) \in \mathcal{I}som(E, F)$ et la proposition V.7 permet donc de conclure que f est un difféomorphisme de classe $\mathcal{C}^1$ de $\mathcal{V}$ vers $\mathcal{W}_R$. $\square$

V.4 Le théorème des fonctions implicites

Supposons que l'on ait une fonction f définie sur un ouvert $\Omega \subset \mathbb{R}^n \times \mathbb{R}^p$ à valeurs dans $\mathbb{R}^p$ et telle que $f(a, b) = 0_F$, pour un certain couple $(a, b) \in \mathbb{R}^n \times \mathbb{R}^p$. Le théorème des fonctions implicites tente de répondre à la question suivante : On a $f(a, b) = 0_{\mathbb{R}^p}$ mais si j'ai x proche de a , puis-je choisir y proche de b pour avoir encore $f(x, y) = 0_{\mathbb{R}^p}$? Quelles conditions dois-je imposer à f pour qu'un tel résultat existe ?

Théorème V.12. (Théorème des fonctions implicites)

Soit f une fonction définie sur un ouvert $\Omega \subset \mathbb{R}^n \times \mathbb{R}^p$ à valeurs dans $\mathbb{R}^p$.

On suppose que $f(a, b) = 0_{\mathbb{R}^p}$, pour un certain couple $(a, b) \in \mathbb{R}^n \times \mathbb{R}^p$.

Si $D_2 f(a, b)$ est inversible ($D_2 f(a, b) \in \text{Isom}(\mathbb{R}^p)$), alors :

Il existe un voisinage ouvert $\mathcal{V}$ de a inclu dans $\mathbb{R}^n$, un voisinage ouvert $\mathcal{W}$ de b inclu dans $\mathbb{R}^p$, et une application $\phi : \mathcal{V} \rightarrow \mathcal{W}$ telle que :

$$\forall (x, y) \in \mathcal{V} \times \mathcal{W}, \quad f(x, y) = 0_{\mathbb{R}^p} \iff y = \phi(x).$$

De plus ϕ est différentiable sur $\mathcal{V}$ et on a :

$$\forall x \in \mathcal{V}, \quad D\phi(x) = - (D_2 f(x, \phi(x)))^{-1} \circ D_1 f(x, \phi(x)).$$

Preuve :

On considère l'application $\psi : \Omega \rightarrow \mathbb{R}^n \times \mathbb{R}^p$ définie par :

$$\psi(x, y) = (x, f(x, y)).$$

Pour tout $(h, k) \in \mathbb{R}^n \times \mathbb{R}^p$, on a :

$$D\psi(a, b)(h, k) = (h, Df(a, b)(h, k)) = (h, D_1 f(a, b)(h) + D_2 f(a, b)(k)).$$

$D\psi(a, b)$ est inversible. En effet :

$$D\psi(a, b)(h, k) = (z, t) \iff h = z \text{ et } k = (D_2 f(a, b))^{-1} (t - D_1 f(a, b)(z)).$$

On applique alors le théorème d'inversion locale à ψ en (a, b) :

Il existe $\mathcal{U}_1$ et $\mathcal{V}_1$ voisinages respectifs de (a, b) et de $(a, f(a, b))$ tels que la restriction $\psi : \mathcal{U}_1 \rightarrow \mathcal{V}_1$ est un $\mathcal{C}^1$ -difféomorphisme.

Son inverse ψ^{-1} peut s'écrire $\psi^{-1}(u, v) = (u, g(u, v))$ où $g : \mathcal{V}_1 \rightarrow \mathbb{R}^p$ est une application $\mathcal{C}^1$ et on a :

$$f(x, y) = 0_{\mathbb{R}^p} \iff \psi(x, y) = (x, 0_{\mathbb{R}^p}) \iff (x, y) = \psi^{-1}((x, 0_{\mathbb{R}^p})) = (x, g(x, 0_{\mathbb{R}^p})) \iff y = g(x, 0_{\mathbb{R}^p}).$$

Soient maintenant un voisinage ouvert $\mathcal{V}$ de a inclu dans $\mathbb{R}^n$, un voisinage ouvert $\mathcal{W}$ de b inclu dans $\mathbb{R}^p$, tels que $\mathcal{V} \times \mathcal{W} \subset \mathcal{V}_1$.

On pose, pour $x \in \mathcal{V}$, $\phi(x) = g(x, 0_{\mathbb{R}^p})$. On a bien $\phi : \mathcal{V} \rightarrow \mathcal{W}$ et

$$\forall (x, y) \in \mathcal{V} \times \mathcal{W}, \quad f(x, y) = 0_{\mathbb{R}^p} \iff y = \phi(x).$$

Pour trouver la différentielle de ϕ en un point x de $\mathcal{V}$, il suffit alors de différentier l'égalité

$$f(x, \phi(x)) = 0_{\mathbb{R}^p}.$$

On obtient donc, pour tout $h \in \mathbb{R}^n$:

$$D_1 f(x, \phi(x))(h) + D_2 f(x, \phi(x)) \circ D\phi(x)(h) = 0_{\mathbb{R}^p},$$

D'où :

$$D\phi(x) = - (D_2 f(x, \phi(x)))^{-1} \circ D_1 f(x, \phi(x)).$$

Chapitre VI

Equations différentielles

Lorsqu'on tente de modéliser des phénomènes physiques, biologiques, économiques et plus généralement des problèmes temporels, il n'est pas rare de tomber sur des fonctions f dont l'évolution à un instant t donné dépend de la valeur de f elle-même et éventuellement de l'instant t aussi. Ces fonctions suivent des équations différentielles... Par exemple,

- En physique, lorsqu'on considère un pendule simple (sans frottements), l'angle θ formé par la verticale et le fil du pendule vérifie l'équation $\theta''(t) + w^2\theta(t) = 0$, pour $t \in \mathbb{R}_+$.
- En biologie, la concentration C en carbone 14 dans les os est régie par l'équation différentielle $C'(t) + \lambda C(t) = 0$, pour $t \in \mathbb{R}_+$ et $\lambda = 1,21 \cdot 10^{-4}$.

VI.1 Généralités

Définition VI.1. Soient E et F deux $\mathbb{R}$ -ev de dimension finie.

Une **équation différentielle ordinaire d'ordre $n \in \mathbb{N}^*$** est une relation du type

$$G(x, y, y', \dots, y^n) = 0$$

où :

- $x \in \mathbb{R}$,
- $y : \mathbb{R} \rightarrow E$ est une application n fois dérivable,
- $G : \Omega \rightarrow F$ est une application avec Ω ouvert de $\mathbb{R} \times E^{n+1}$.

Remarque VI.2. le terme “ordinaire” est utilisé lorsque la variable de base est réelle. On peut très bien adapter la définition lorsque x appartient à une $\mathbb{R}$ -ev quelconque.

Définition VI.3. Quelques types d'équations différentielles particuliers

On dit qu'une équation différentielle est :

- **sous forme résolue** (ou **résolue**) si elle peut s'écrire sous la forme : $y^n = F(x, y, y', \dots, y^{n-1})$.
- **autonome** si elle peut s'écrire sous la forme $H(y, y', \dots, y^n) = 0$.
- **à variables séparables** si elle peut s'écrire sous la forme $H(y, y', \dots, y^n) = A(x)$.
- **linéaire** si elle peut s'écrire sous la forme $L(x, y, y', \dots, y^n) = 0$, où L est linéaire en les variables $y, y', \dots, y^n$.

Remarque VI.4. En théorie, on peut très souvent passer du cas non résolu au cas résolu en utilisant le théorème des fonctions implicites. De ce fait, la plupart des énoncés sont donnés pour des équations différentielles sous forme résolue... Face à une équation différentielle non résolue, il suffira juste de vérifier que l'on se trouve dans le cadre d'application du théorème des fonctions implicites pour pouvoir utiliser les énoncés... En pratique, il arrive fréquemment de modifier légèrement l'équation différentielle pour pouvoir la mettre sous forme résolue mais cela impose quelques "restrictions à effacer plus tard", par exemple, l'équation $xy' - 2y = 0$ avec $x \in \mathbb{R}$ peut être transformée en un premier temps par $y' = \frac{2y}{x}$ avec $x \in \mathbb{R}^*$ pour la résoudre mais il faudra ensuite tenir compte du fait que x peut être nul dans l'équation initiale, donc qu'a priori, on doit la résoudre sur $\mathbb{R}$ et pas seulement sur $\mathbb{R}^*$.

Remarque VI.5. Pour les énoncés et les preuves, on se restreint souvent au cas des équations résolues (cf remarque précédente) mais aussi de degré 1. En effet, quitte à modifier E , on peut toujours se ramener à une équation différentielle de degré 1 via la méthode suivante : Si l'équation résolue est du type $y^n = G(x, y, y', \dots, y^{n-1})$ avec $G : \Omega \rightarrow F$ une application avec Ω ouvert de $\mathbb{R} \times E^n$, on pose $\vec{Y} = (y, y', \dots, y^{n-1})$. De ce fait, $\vec{Y}' = (y', \dots, y^n)$ et on définit $\vec{G} : \Omega \rightarrow E^n$ par

$$\vec{G}(x, z_0, z_1, \dots, z_{n-1}) = (z_1, z_2, \dots, z_{n-1}, G(x, y, y', \dots, y^{n-1})),$$

de sorte que $\vec{Y}' = \vec{G}(x, \vec{Y})$ est équivalente à $y^n = G(x, y, y', \dots, y^{n-1})$.

Définition VI.6. Soit deux $\mathbb{R}$ -ev E et F , Ω un ouvert de $\mathbb{R} \times E^{n+1}$ et $G : \Omega \rightarrow F$.

On considère l'équation ordinaire ($\mathcal{E}$) : $G(x, y, y', \dots, y^n) = 0$.

On appelle **solution de l'équation différentielle** ($\mathcal{E}$) un couple (I, ϕ) où :

- I est un intervalle de $\mathbb{R}$ non réduit à un singleton,
- $\phi : I \rightarrow E$ est une fonction n -fois dérivable et vérifie :

$$\forall x \in I, G(x, \phi(x), \phi'(x), \dots, \phi^n(x)) = 0$$

On dit alors que ϕ est une I -solution de l'équation ($\mathcal{E}$).

On note en général :

- $S_I(\mathcal{E})$ l'ensemble des I -solutions de l'équation ($\mathcal{E}$).
- $S(\mathcal{E})$ l'ensemble des solutions de l'équation ($\mathcal{E}$).

Intégrer une équation différentielle c'est déterminer l'ensemble de ses solutions.

Remarque VI.7. Dans la donnée d'une solution d'équation différentielle, l'intervalle de définition est très important... en particulier car on doit avoir implicitement $\forall x \in I, (x, \phi(x), \phi'(x), \dots, \phi^n(x)) \in \Omega$. L'intégration d'une équation différentielle se fera TOUJOURS sur un intervalle.

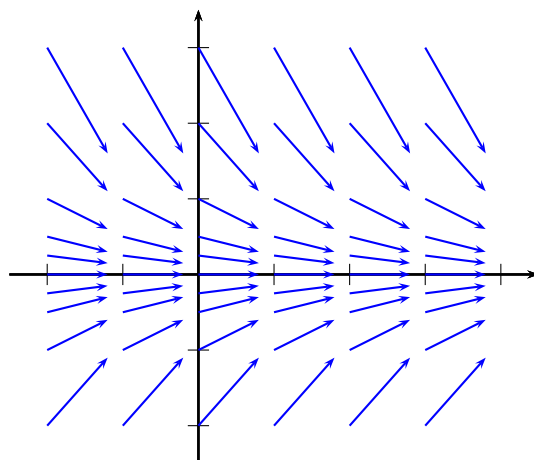
Remarque VI.8. Si G est de classe $\mathcal{C}^k$, alors il en va de même pour toute solution de ($\mathcal{E}$). Cela n'a évidemment d'impact que si $k > n$... cela se montre par récurrence.

Remarque VI.9. Interprétation graphique des solutions

Dans le cas d'une équation résolue $y' = G(x, y)$, à tout point $(a, b) \in \mathbb{R}^2$, on peut associer un vecteur $\vec{v}_{(a,b)}$ de pente $G(a, b)$ et de longueur fixe.

Si l'on trace le champs de vecteurs obtenu, une solution de l'équation différentielle est une courbe $\mathcal{C}$ du plan qui est tangente aux vecteurs $\vec{v}_{(a,b)}$ en tout point de son ensemble de définition.

Ici, le champs de vecteur est associé à l'équation $y' = -\frac{1}{2}y$.



Définitions VI.10. Si (I, ϕ) est une solution d'une équation différentielle $(\mathcal{E})$, on appelle **prolongement de la solution** (I, ϕ) toute solution (J, ψ) de $(\mathcal{E})$ telle que $I \subset J$ et $\psi|_I = \phi$.

On appelle **solution maximale**, une solution de $(\mathcal{E})$ qui n'admet pas de prolongement.

Remarque VI.11. Toute solution peut être prolongée en une solution maximale. D'autre part, si (I, ϕ) est une solution maximale, alors I est un intervalle ouvert.

Méthode VI.12. Raccordement de 2 solutions

Souvent dans le cas des équations différentielles présentées de façon non résolues, la mise sous forme résolue entraîne l'exclusion d'une ou plusieurs valeurs de la variable de base. Par exemple, pour résoudre $xy' - 2y = 0$ pour $x \in \mathbb{R}$, on passe par la forme résolue $y' = \frac{2y}{x}$ pour $x \in \mathbb{R}^*$.

Les solutions d'une équations différentielles étant toujours définies sur un intervalle (et pas une réunion d'intervalles), on va donc résoudre sur $\mathbb{R}_+$ ou sur $\mathbb{R}_-$ dans un premier temps, et ensuite voir quelles solutions se prolongent sur $\mathbb{R}$.

Sur $\mathbb{R}_+$, les fonctions $\phi_k : x \mapsto kx^2$ pour $k \in \mathbb{R}$ sont solutions.

Sur $\mathbb{R}_-$, les fonctions $\phi_p : x \mapsto px^2$ pour $p \in \mathbb{R}$ sont solutions.

Or quels que soient les réels k et p , on a $\lim_{x \rightarrow 0} \phi_k(x) = \lim_{x \rightarrow 0} \phi_p(x) = 0$ et $\lim_{x \rightarrow 0} \phi'_k(x) = \lim_{x \rightarrow 0} \phi'_p(x) = 0$ donc étendre toutes les fonctions ϕ en 0... ce ne sont donc pas des solutions maximales. On peut en fait raccorder n'importe quelles fonctions ϕ_k et ϕ_p pour former une solution ϕ définie sur $\mathbb{R}$ de l'équation. Ainsi, pour tout couple $(k, p) \in \mathbb{R}^2$, la fonction :

$$\begin{aligned} \phi_{(k,p)} : \mathbb{R} &\rightarrow \mathbb{R} \\ x &\mapsto \begin{cases} kx^2 & \text{si } x \geq 0, \\ px^2 & \text{si } x < 0. \end{cases} \end{aligned}$$

Ces solutions $(\mathbb{R}, \psi_{(k,p)})$ sont des solutions maximales. Ainsi, l'équation $xy' - 2y = 0$ a donc une infinité de solutions maximales.

Cela pose entre autres la question de savoir dans quels cas a-t-on unicité de la solution à une équation différentielle sur un intervalle donné...

VI.2 Le problème de Cauchy

Graphiquement (voir remarque VI.9), on peut observer que 2 solutions ϕ_1 et ϕ_2 d'une même équation différentielle résolue d'ordre 1 $y' = G(x, y)$ sur un même intervalle ne peuvent jamais se croiser. En effet, si elle se croisaient, leurs pentes respectives (y') au point de croisement (x_0, y_0) doivent être différentes, ce qui n'est pas possible car $\phi_1'(x_0) = \phi_2'(x_0) = G(x_0, y_0)$.

L'idée de Cauchy est d'assurer l'unicité de la solution d'une équation différentielle résolue d'ordre 1 en fixant (en plus de certaines restrictions sur les fonctions $G...$) un point par lequel notre solution doit passer, ce qui correspond en fait à fixer les conditions initiales d'un phénomène physique ou biologique modélisé. Si l'on reprend les exemples donnés en introduction du chapitre :

- En physique, on peut s'intéresser à l'évolution de l'angle θ formé par la verticale et le fil du pendule lorsqu'on le lâche avec un angle initial de $\frac{\pi}{4}$:

$$(\mathcal{C}) : \begin{cases} \theta''(t) + w^2\theta(t) = 0, & t \in \mathbb{R}, \\ \theta(0) = \frac{\pi}{4}. \end{cases}$$

- En biologie, l'évolution de la concentration C en carbone 14 dans les os est régie par l'équation différentielle $C'(t) + \lambda C(t) = 0$, pour $t \in \mathbb{R}_+$ et $\lambda = 1,21 \cdot 10^{-4}$ et sur les os vivants, on a en général $C = 10^{-12}$.

$$(\mathcal{C}) : \begin{cases} C'(t) + \lambda C(t) = 0, & t \in \mathbb{R}, \\ C(0) = 10^{-12}. \end{cases}$$

Définitions VI.13.

On appelle **condition initiale** d'une équation différentielle $(\mathcal{E}) : G(x, y, y', \dots, y^n) = 0$ avec G définie sur un ouvert Ω , un n -uplet $(x_0, y_0, y_1, \dots, y_{n-1}) \in \Omega$ tel que toute solution (I, ϕ) de $(\mathcal{E})$ telle que $x_0 \in I$ doit vérifier $\forall i \in \{0, 1, \dots, n-1\}$, $\phi^{(i)}(x_0) = y_i$.

Un **problème de Cauchy** est la donnée d'une équation différentielle avec une condition initiale $(x_0, y_0, y_1, \dots, y_{n-1}) \in \Omega$.

$$(\mathcal{C}) : \begin{cases} G(x, y, y', \dots, y^n) = 0 \\ \forall i \in \{0, 1, \dots, n-1\}, \phi^{(i)}(x_0) = y_i \end{cases}$$

Notez que cela forme à chercher des solutions définies sur un intervalle contenant x_0 . Lorsqu'un problème admet une unique solution maximale, on dit qu'il y a **unicité au problème de Cauchy**.

Attention : il n'y a pas unicité (ni même existence) à tous les problèmes de Cauchy, notamment lorsque l'équation différentielle n'est pas déjà mise sous forme résolue.

Si on prend le problème de Cauchy suivant :

$$(\mathcal{C}_a) : \begin{cases} xy' - 2y = 0, \\ y(0) = a. \end{cases}$$

Ce problème n'a pas de solutions si $a \neq 0$, et pour $a = 0$, on a une infinité de solutions (Cf. Méthode VI.12).

Heureusement, dans certains cas on a l'existence et l'unicité qui sont assurées ! notamment lorsque l'équation différentielle est sous forme résolue $y' = G(x, y)$ et que la fonction G est localement lipschitzienne en sa seconde variable uniformément en la première variable...

Définition VI.14. Soit E un $\mathbb{R}$ -ev et Ω un ouvert de $\mathbb{R} \times E$. Une fonction $G : \Omega \rightarrow E$ est dite **localement lipschitzienne en sa seconde variable uniformément en la première variable sur Ω** si pour tout $(x_0, y_0) \in \Omega$, il existe un voisinage ouvert $\mathcal{V}_{(x_0, y_0)}$ de (x_0, y_0) , inclus dans Ω tel que :

$$\exists C > 0, \forall (x, y), (x, z) \in \mathcal{V}_{(x_0, y_0)}^2, \|G(x, y) - G(x, z)\| \leq C \|z - y\|.$$

Attention ! La constante C est uniforme en la première variable : elle ne dépend pas de x . C'est une condition plus forte que de dire “ G localement lipchitzienne en y ”.

Notez aussi que ce n'est pas parce que l'équation différentielle est autonome (G est constante en x) que cette propriété est vérifiée au voisinage de tout point... (il faut quand même que G soit lipschitzienne en y). Par exemple, si $G(x, y) = \sqrt{y}$ n'est pas lipschitzienne au voisinage de $(0, 0)$.

En pratique... Les cas de figures suivants assure à une fonction G le caractère localement lipschitzien en la seconde variable uniformément en la première variable :

- G $\mathcal{C}^1$ -différentiable sur un voisinage de (x_0, y_0) ,
- G admet une deuxième dérivée partielle continue.

Dans le second cas, cela implique que $(x, y) \mapsto D_2 G(x, y)$ est bornée sur un voisinage de y_0 et on peut donc utiliser le théorème de la moyenne à $y \mapsto G(x, y)$ pour obtenir le caractère localement lipschitzien en la seconde variable uniformément en la première variable.

Théorème VI.15. Théorème de Cauchy-Lipschitz - Cas des ED sous forme résolues

Soient E un espace de Banach, Ω un ouvert de $\mathbb{R} \times E^n$ et $G : \Omega \rightarrow E$, continue sur Ω et localement lipschitzienne en sa seconde variable (de E^n) uniformément en la première variable (de $\mathbb{R}$) sur Ω .

Pour toute condition initiale $(x_0, y_0, y_1, \dots, y_{n-1}) \in \Omega$, le problème de Cauchy :

$$(\mathcal{C}) : \begin{cases} y^{(n)} = G(x, y, y', \dots, y^{n-1}) \\ \forall i \in \{0, 1, \dots, n-1\}, y^{(i)}(x_0) = y_i \end{cases}$$

admet une unique solution maximale (I_0, ϕ) avec $x_0 \in I_0$.

Autrement dit, il existe un intervalle ouvert I_0 contenant x_0 et une fonction $\phi : I_0 \rightarrow E$ telle que :

- ϕ est une I_0 solution de l'équation différentielle $y^{(n)} = G(x, y, y', \dots, y^{n-1})$,
- ϕ vérifie $\forall i \in \{0, 1, \dots, n-1\}, \phi^{(i)}(x_0) = y_i$,
- Si J est un intervalle inclu dans I_0 et ψ une J -solution du problème de Cauchy $(\mathcal{C})$, alors $\psi = \phi|_J$.

Remarque VI.16. Pour les équations différentielles non résolues, on peut utiliser le théorème des fonctions implicites pour obtenir G . Puisque le théorème des fonctions implicites donne également accès à la différentielle de G , on peut, au cas par cas, vérifier si on est bien dans les hypothèses du théorème de Cauchy-Lipschitz.

Remarque VI.17. Comme mentionné dans la remarque VI.5, on peut restreindre la preuve au cas des équations différentielles d'ordre 1, quitte à passer au multidimensionnel. pour les équations différentielles d'ordre 1 sous forme résolue $y' = G(x, y)$, la démonstration repose sur le **théorème du point fixe de Banach** :

Si $F : C \rightarrow C$ est une application contractante d'un espace métrique complet C dans lui-même, alors F admet un unique point fixe dans C : $\exists! x^* \in C, F(x^*) = x^*$.

Notez qu'une application F est contractante si elle est k -lipschitzienne pour un certain $k \in]0, 1[$: $\exists k \in]0, 1[, \forall (x, y) \in C^2, d(F(x), F(y)) \leq k d(x, y)$.

Preuve du théorème de Cauchy-Lipschitz (ED sous forme résolue, ordre 1)

Soient $G : \Omega \rightarrow E$, où E est un espace de Banach, Ω est un ouvert de $\mathbb{R} \times E$ et $(x_0, y_0) \in \Omega$. On considère le problème de Cauchy suivant :

$$(C) : \begin{cases} y' = G(x, y) \\ y(x_0) = y_0 \end{cases}$$

On suppose que G est localement lipschitzienne en sa seconde variable uniformément en la première variable sur un voisinage de (x_0, y_0) .

et on va appliquer le théorème du point fixe de Banach à une fonction F du type :

$$\begin{aligned} F : \mathcal{C}_0(I, E) &\rightarrow \mathcal{C}_0(I, E) \\ y &\mapsto F(y) : x \mapsto y_0 + \int_{x_0}^x G(t, y(t)) dt \end{aligned}$$

où $\mathcal{C}_0(I, E)$ est l'ensemble des applications continues d'un intervalle I contenant x_0 (que nous définirons plus loin) à valeurs dans E .

En effet, $F(y) = y$ revient à dire que pour tout $x \in I$, on a : $y(x) = y_0 + \int_{x_0}^x G(t, y(t)) dt$. Cela est équivalent à dire que y est la primitive de $G(x, y(x))$ qui vaut y_0 en x_0 , i.e. :

$$F(y) = y \iff \begin{cases} y' = G(x, y) \\ y(x_0) = y_0 \end{cases}.$$

Prouver l'existence d'une fonction du type de F qui est contractante (et donc admet un unique point fixe) revient donc à prouver l'existence d'une unique solution au problème de Cauchy.

Puisque G est localement lipschitzienne en sa seconde variable uniformément en la première variable sur un voisinage de (x_0, y_0) , il existe $\mathcal{V}$ voisinage ouvert de (x_0, y_0) contenu dans Ω et une constante $C > 0$ tels que :

$$\forall (x, y), (x, z) \in \mathcal{V}^2, \|G(x, y) - G(x, z)\| \leq C \|y - z\|. \quad (\text{VI.1})$$

Soient $\alpha > 0$ et $\beta > 0$ tels que $[x_0 - \alpha, x_0 + \alpha] \times \overline{B}(y_0, \beta) \subset \mathcal{V}$.

On notera $I_\alpha = [x_0 - \alpha, x_0 + \alpha]$ et $\overline{B}_\beta = \overline{B}(y_0, \beta)$.

On remarquera que $\overline{B}_\beta$ est une partie fermée d'un espace complet (E) , et donc que $\overline{B}_\beta$ est donc un espace métrique complet (où la métrique dérive de la norme de E).

L'application $\begin{matrix} I_\alpha & \rightarrow & \mathbb{R} \\ x & \mapsto & \|G(x, y_0)\| \end{matrix}$ est continue sur I_α , donc bornée. Soit $m = \sup_{x \in I_\alpha} \|G(x, y_0)\|$.

On a alors :

$$\begin{aligned} \forall (x, y) \in I_\alpha \times \overline{B}_\beta, \|G(x, y)\| &\leq \|G(x, y) - G(x, y_0)\| + \|G(x, y_0)\| \\ &\leq C \|y - y_0\| + m \end{aligned}$$

Ainsi,

$$\forall (x, y) \in I_\alpha \times \overline{B}_\beta, \|G(x, y)\| \leq C\beta + m \quad (\text{VI.2})$$

Pour tout intervalle ouvert $J \subset I_\alpha$, on note $\mathcal{C}_0(J, \overline{B}_\beta)$ l'ensemble des applications continues sur J à valeurs dans $\overline{B}_\beta$. Puisque $\overline{B}_\beta$ est complet (boule fermée d'un espace de Banach), $\mathcal{C}_0(J, \overline{B}_\beta)$ est un espace métrique complet, lorsqu'il est muni de la distance $d(f, g) = \sup_{x \in J} \|f(x) - g(x)\|$.

On considère alors la fonction F_J définie comme suit :

$$\begin{aligned} F_J : \mathcal{C}_0(J, \overline{B}_\beta) &\rightarrow \mathcal{C}_0(J, E) \\ y &\mapsto F_J(y) : x \mapsto y_0 + \int_{x_0}^x G(t, y(t)) dt \end{aligned}$$

On va montrer que, quitte à diminuer un peu I_α , F_J est à valeurs dans $\mathcal{C}_0(J, \overline{B}_\beta)$ et que F_J est contractante. On pourra ensuite appliquer le théorème du point fixe à F_J .

Pour tout $x \in J$, on a, grace à l'inéquation VI.2 :

$$\|F_J(y)(x) - y_0\| = \left\| \int_{x_0}^x G(t, y(t)) dt \right\| \leq \int_{x_0}^x \|G(t, y(t))\| dt \leq \|x - x_0\| (C\beta + m).$$

Ainsi,

$$\forall x \in J, \forall y \in \mathcal{C}_0(J, \overline{B}_\beta), \|F_J(y)(x) - y_0\| \leq \alpha (C\beta + m).$$

et on peut, quitte à choisir α plus petit, on peut supposer que $\alpha (C\beta + m) \leq \beta$, ce qui permet d'avoir : $\forall x \in J, \forall y \in \mathcal{C}_0(J, \overline{B}_\beta), \|F_J(y)(x) - y_0\| \leq \beta$, i.e., $F_J(y)(x) \in \overline{B}_\beta$. On peut donc supposer que $F_J(y)$ est bien à valeurs dans $\overline{B}_\beta$ quelque soit $y \in \mathcal{C}_0(J, \overline{B}_\beta)$ et donc F_J est à valeurs dans $\mathcal{C}_0(J, \overline{B}_\beta)$.

Montrons que F_J est contractante. Soient $x \in J$ et $(y, z) \in \mathcal{C}_0(J, \overline{B}_\beta)^2$.

$$\begin{aligned} \|F_J(y)(x) - F_J(z)(x)\| &= \left\| \int_{x_0}^x G(t, y(t)) - G(t, z(t)) dt \right\|, \\ &\leq \int_{x_0}^x \|G(t, y(t)) - G(t, z(t))\| dt, \\ &\leq \int_{x_0}^x C \|y(t) - z(t)\| dt \quad \text{d'après l'inégalité VI.1,} \\ &\leq C \|x - x_0\| \sup_{t \in J} \|y(t) - z(t)\|. \\ &\leq C\alpha \sup_{t \in J} \|y(t) - z(t)\|. \end{aligned}$$

Ainsi $\forall x \in J, \|F_J(y)(x) - F_J(z)(x)\| \leq C\alpha d(y, z)$, on obtient donc :

$$\forall (y, z) \in \mathcal{C}_0(J, \overline{B}_\beta)^2, d(F_J(y), F_J(z)) \leq C\alpha d(y, z)$$

Quitte à réduire une fois de plus la valeur de α en supposant par exemple que $\alpha < \frac{1-\epsilon}{C}$ pour un $\epsilon > 0$ fixé, on peut donc forcer F_J à être une application contractante de $\mathcal{C}_0(J, \overline{B}_\beta)$ dans lui-même.

Ainsi, pour tout intervalle ouvert $J \subset I_\alpha$, il existe donc une unique solution $\phi_J \in \mathcal{C}_0(J, \overline{B}_\beta)$ telle que $F_J(\phi_J) = \phi_J$, donc pour laquelle on a, pour tout $x \in J$, $\phi_J(x) = \int_{x_0}^x G(t, \phi_J(t)) dt$. En

particulier, cela impose que ϕ_J est de classe $\mathcal{C}^1$ puisqu'elle est primitive d'une application continue et on a bien sûr

$$F(\phi_J) = \phi_J \iff \begin{cases} \phi'_J = G(x, \phi_J) \\ \phi_J(x_0) = y_0 \end{cases} \iff \phi_J \text{ solution du problème de Cauchy } (\mathcal{C}).$$

A priori, cette méthode fournit une unique solution au problème de Cauchy $(\mathcal{C})$ seulement une fois J fixée... mais cela n'est pas vraiment un problème. Pour $J = \overset{\circ}{I}_\alpha =]x_0 - \alpha, x_0 + \alpha[$, on obtient une solution $\phi_{\overset{\circ}{I}_\alpha}$ qui est non seulement solution au problème de Cauchy sur $\overset{\circ}{I}_\alpha$ mais sa restriction à n'importe quel intervalle ouvert $J \subset I_\alpha$, est la solution du problème de Cauchy sur J . $\square$

Remarque VI.18. Sous les hypothèses du théorème de Cauchy-Lipschitz, on peut montrer que :

- Si (I, ϕ) est une solution maximale de $(\mathcal{C})$, alors I est un intervalle ouvert,
- Les graphes des solutions maximales aux problèmes de Cauchy

$$(\mathcal{C}_{(y_0, y_1, \dots, y_{n-1})}) : \begin{cases} G(x, y, y', \dots, y^n) = 0 \\ \forall i \in \{0, 1, \dots, n-1\}, \phi^{(i)}(x_0) = y_i \end{cases}$$

pour $(y_0, y_1, \dots, y_{n-1})$ in E forment une partition de l'ensemble Ω de définition de G . On dit qu'ils forment un feuilletage de Ω ou qu'ils feuilletent Ω .

Remarque VI.19. Le théorème de Cauchy-Lipschitz est valable en dimension infinie dès lors que l'on reste dans le cadre des espaces de Banach. En dimension infinie, il existe des variantes du théorème de Cauchy-Lipschitz qui permettent d'affaiblir les hypothèses sur la fonction G en conservant l'existence d'une solution au problème mais qui ne garantissent pas l'unicité de la solution maximale. C'est par exemple le cas pour le théorème de Cauchy-Arzela :

Théorème VI.20. Théorème de Cauchy-Arzela - Cas des ED sous forme résolues

Soient E un espace de Banach, Ω un ouvert de $\mathbb{R} \times E^n$ et $G : \Omega \rightarrow E$, continue sur Ω .

Pour toute condition initiale $(x_0, y_0, y_1, \dots, y_{n-1}) \in \Omega$, le problème de Cauchy :

$$(\mathcal{C}) : \begin{cases} y^{(n)} = G(x, y, y', \dots, y^{n-1}) \\ \forall i \in \{0, 1, \dots, n-1\}, y^{(i)}(x_0) = y_i \end{cases}$$

admet une solution (I_0, ϕ) avec I_0 intervalle ouvert tel que $x_0 \in I_0$.