Calcul différentiel et systèmes dynamiques

Raphael Lachieze-Rey[1]

L3 Mathématiques, UFR Maths-Info, Université Paris Cité, 2022-2023

[1] raphael.lachieze-rey@parisdescartes.fr,raphael.lachieze-rey@math.cnrs.fr

Contents

Part I

Calcul différentiel

Rappels

Définition 1. *Un espace vectoriel normé est un couple* $(E, \|\cdot\|_E)$ *où* $\|\cdot\|_E$*, parfois notée simplement* $\|\cdot\|$*, est une norme sur* E*, c'est-à-dire:*

- $\|\lambda x\|_E = |\lambda|\|x\|_E$ *pour* $\lambda \in \mathbb{R}, x \in E$
- $\|x+y\|_E \leqslant \|x\|_E + \|y\|_E$ *pour* $x, y \in E$
- $\|x\|_E = 0 \Rightarrow x = 0$ *pour* $x \in E$.

Les éléments d'un e.v.n. seront appelés "vecteurs", même si la dimension est infinie.

On considère un e.v.n. $(E, \|\cdot\|_E)$. Pour $x \in E$, on note pour $x \in E, r \geqslant 0$ la boule de centre x de rayon r par

$$B_E(x,r) = \{y \in E : \|x-y\|_E \leqslant r\}.$$

On note parfois cette boule $B_{\|\cdot\|_E}(x,r)$, ou juste $B(x,r)$ s'il n'y a pas d'ambiguités. Si $E = \mathbb{R}^d$ pour un certain $d \in \mathbb{N}^*$, on note pour $p \geqslant 1$

$$\|x\|_p = (\sum_{i=1}^d |x_i|^p)^{1/p}, x \in \mathbb{R}^d,$$

la norme p, pour $p \in [1, +\infty[$, $\|x\|_\infty = \max(|x_i|, i = 1, \ldots, d)$. On rappelle que sur $\mathbb{R}^d$, toutes les normes sont équivalentes. On note aussi $(e_1, \ldots, e_d)$ la base canonique de $\mathbb{R}^d$.

Définition 2. *Soit* $(E, \|\cdot\|_E), (F, \|\cdot\|_F)$ *deux e.v.n. . Une application linéaire* $f : E \to F$ *est continue ssi il existe* $C \geqslant 0$ *tel que* $\forall x \in E$*,*

$$\|f(x)\|_F \leqslant C\|x\|_E.$$

La plus petite constante C *telle que cette inégalité soit vraie pour tout* x *est notée* $C = |||f||| = |||f|||_{E,F}$*. De manière équivalente,*

$$|||f|||_{E,F} = \sup_{x \in B_E(0,1)} \|f(x)\|_F = \sup_{x \in E \setminus \{0\}} \frac{\|f(x)\|_F}{\|x\|_E}.$$

L'espace des applications linéaires continues de $(E, \|\cdot\|_E)$ vers $(F, \|F\|_F)$ est noté $\mathcal{L}(E,F)$.

Dans le cas où E et F sont de dimension finie, disons n et p, $\mathcal{L}(E,F)$ est isomorphe à $\mathcal{M}_{n,p}(\mathbb{R})$. On appelle norme matricielle sur $\mathcal{M}_{n,p}(\mathbb{R})$ une norme $\|\cdot\|$ telle que pour tout $A \in \mathcal{M}_{n,p}(\mathbb{R}), B \in \mathcal{M}_{k,n}(\mathbb{R})$,

$$\|A \cdot B\| \leqslant \|A\| \cdot \|B\|.$$

Dans le cas où E et F sont de dimension finie, alors toutes les applications linéaires sont continues et $\dim(\mathcal{L}(E,F)) = \dim(E)\dim(F)$ (c'est aussi la dimension de l'espace des matrices associé).

Manipulation des "$o(\cdot)$" dans un e.v.n.

Notation 1. *Pour E, F la notation $f(u) = \underset{u\to 0}{o}(u)$ a du sens dès lors que f est définie sur une petite boule $B_E(0,r)$, et peut s'écrire plus précisément $o(\|u\|)$,*

$$\underset{u\to 0}{o}(u) = \underset{\|u\|\to 0}{o}(\|u\|) = \|u\|\varepsilon(\|u\|)$$

où ε est une application implicite $\varepsilon : [-r, r] \to F$ telle que $\lim_{t\to 0} \|\varepsilon(t)\|_F = 0$. On écrit parfois $\varepsilon(u)$ au lieu de $\varepsilon(\|u\|)$, dans ce cas $\varepsilon : B_E(0,r) \to F$ vérifie $\lim_{u\to 0_E} \varepsilon(u) = 0$.

On peut aussi formuler $f(u) = o(u)$ de manière équivalente par

$$\lim_{u\to 0} \frac{f(u)}{\|u\|} = 0.$$

On note aussi pour $k \in \mathbb{N}$

$$\underset{\|u\|\to 0}{o}(u^k) = \underset{\|u\|\to 0}{o}(\|u\|^k) = \|u\|^k \varepsilon_k(u) \in F$$

et donc $f(u) = o(u^k)$ signifie $\lim_{u\to 0} \|u\|^{-k} f(u) = 0$.

Utilisation avec les normes triples: pour $f : E \to F, g : F \to G$ linéaires continues

$$o(g(f(u))) = o(f(u)) = o(u) \text{ car } \|g(f(u))\|_G \leqslant |||g|||_{F,G}\|f(u)\|_F \leqslant |||g|||_{F,G}|||f|||_{E,F}\|u\|_E$$

Chapter 1

Application différentielle

1.1 Définitions

Dans tout le chapitre, $(E, \|\cdot\|_E)$ est un e.v.n. , Ω est un ouvert de E et x_0 est un point quelconque de Ω. $(F, \|\cdot\|_F)$ est un autre e.v.n. .

En bleu: rectifications par rapport au poly papier

Définition 3. *Soit $f : \Omega \to F, u \in E$. On appelle dérivée partielle de f en x_0 dans la direction u la limite dans F, si elle existe,*

$$\partial_u f(x_0) = \lim_{t \to 0} \frac{f(x_0 + tu) - f(x_0)}{t} \in F.$$

On dit aussi "dérivée directionnelle" de f dans la direction u. S'il existe une base telle que $u = e_i$ est le i-ème vecteur de la base , on note parfois

$$\partial_{u_i} f(x_0) = \partial_i f(x_0)$$

On utilise aussi les notations

$$\frac{\partial f}{\partial u}(x_0) \text{ ou } \frac{\partial f}{\partial_u}|_{x_0} \text{ ou encore } f_u(x_0).$$

Remarque 1. *Si $d = 1$, pour $u \in \mathbb{R} \setminus \{0\}$, $f : \mathbb{R} \to \mathbb{R}$ est dérivable dans la direction u en x_0 ssi f est dérivable en x_0 et*

$$\partial_u f(x_0) = u f'(x_0).$$

Définition 4. *Soit $f : \Omega \to F$. On dit que f est différentiable en x_0 ssi il existe une application linéaire continue $\ell \in \mathcal{L}(E, F)$ telle que*

$$\lim_{u \to 0} \frac{f(x_0 + u) - f(x_0) - \ell(u)}{\|u\|_E} = 0$$

ou de manière équivalente

$$f(x_0 + u) = f(x_0) + \ell(u) + o(u).$$

L'application ℓ est notée df_{x_0}. Cette application est unique. Si f est différentiable en tout point $x_0 \in \Omega$, on dit que f est différentiable sur Ω, et on note df son application différentielle.

Remarquons que $x_0 \mapsto df_{x_0}$ est une application de $\Omega \to \mathcal{L}(E,F)$. On peut aussi voir $(x_0,u) \mapsto df_{x_0}(u)$ comme une application de $\Omega \times E \to F$.

La différentiabilité est équivalente à l'existence de $\ell \in \mathcal{L}(E,F)$ telle que

$$f(x_0+u) = f(x_0) + \ell(u) + \underset{u \to 0_F}{o}(u).$$

Il faut donc faire trois choses pour montrer la différentiabilité:

1. Trouver une application $\ell : E \to F$ telle que

$$f(x_0+u) = f(x_0) + \ell(u) + \underbrace{o(u)}_{\|u\|_E \varepsilon_{x_0,F}(u)}.$$

   On peut aussi trouver cette application en fixant u et trouvant $\ell(u)$ telle que pour $t \in \mathbb{R}$,

$$f(x_0+tu) = f(x_0) + t\ell(u) + o(t).$$

   On remarque que dans ce cas $\ell(u)$ coincide avec $\partial_u f(x_0)$.

   Conséquence: preuve de l'unicité. Si f admet une différentielle ℓ en x_0,

$$\ell(u) = \lim_{t \to 0} \frac{f(x_0+tu) - f(x_0)}{t},$$

   La valeur de ℓ est donc unique par unicité de la limite.

   □

2. Si ça n'est pas déjà le cas, il faut montrer que $u \to \ell(u)$ est linéaire.

3. Il faut montrer que $u \to \ell(u)$ est continue, c'est-à-dire qu'il existe C telle que

$$\ell(u) \leqslant Cu, u \in E.$$

   Si on est en dimension finie, cette étape est inutile.

Pour faire le lien avec la dérivée partielle dans une direction $u \in E$, on a pour $t \in \mathbb{R}$,

$$df_{x_0}(tu) = t\partial_u f(x_0)$$

Exercice 1. Traiter le cas où

$$E = L^2([0,1]) = \{x : [0,1] \to \mathbb{R} : \int x^2(t)dt < \infty\}$$

avec $f(x) = \|x\|^2 = \int_0^1 x(t)^2 dt, x \in E$ ou plus général d'un espace Hilbertien muni d'un produit scalaire $\langle \cdot, \cdot \rangle$ avec $f(x) = \langle x, x \rangle$.

Attention: La différentiabilité de f en x implique l'existence pour chaque $u \in E$ de $\partial_u f(x_0)$, mais la réciproque est fausse. Montrer l'existence de $\partial_u f(x_0)$ pour chaque $u \in \mathbb{R}$ n'est que la première étape, est n'est pas du tout suffisant pour montrer la différentiabilité. Il faut encore montrer la linéarité et la continuité de $u \to d_u f(x_0)$.

Traitons quelques cas faciles. Le cas de la dimension 1:

Proposition 1. *On suppose $d = 1$. Soit $f : \mathbb{R} \to \mathbb{R}$. f est différentiable en x_0 ssi f est dérivable en x, et dans ce cas $df_{x_0}(u) = uf'(x_0)$.*

Proof. Supposons f dérivable en x_0. On a donc le $DL_1(x_0)$

$$f(x_0 + u) = f(x_0) + uf'(x_0) + \underset{u \to 0}{o}(u).$$

L'application $u \mapsto f'(x_0)u$ est bien linéaire continue, donc f est différentiable et $df_{x_0}(u) = f'(x_0)u, u \in \mathbb{R}$.

Si réciproquement il existe ℓ linéaire continue $\mathbb{R} \to \mathbb{R}$ telle que

$$f(x_0 + u) = f(x_0) + \ell(u) + o(u)$$

alors comme les seules applications linéaires $\mathbb{R} \to \mathbb{R}$ sont de la forme $\ell(u) = cu$ pour un certain $c \in \mathbb{R}$, on a

$$\frac{f(x_0 + u) - f(x_0)}{u} = c + o(1)$$

ce qui est bien la définition du nombre dérivé $c = f'(x_0)$. $\square$

On peut aisément généraliser au cas où l'espace de départ est $\mathbb{R}$, mais pas forcément l'espace d'arrivée:

Définition 5. *Si $E = \mathbb{R}$ et $I = \Omega$ est un intervalle ouvert de $\mathbb{R}$, alors si $f : I \to F$ est différentiable en un point $t \in I$, on a $df_t \in \mathcal{L}(\mathbb{R}, F)$. Donc, en notant $f'(t) = df_t(1)$, par linéarité,*

$$\forall u \in \mathbb{R}, df_t(u) = uf'(t).$$

On appelle $f'(t)$ la dérivée de f au point t.

- *Si $F = \mathbb{R}$, il s'agit de la notion classique d'application dérivée.*
- *Si $F = \mathbb{R}^n$, $f(t) = (f_i(t))_{i=1,\ldots,n}$, et $f'(t) = (f_i'(t))_{i=1}^n$.*

Le dernier cas "simple" est le cas linéaire:

Proposition 2. *Soit f une application linéaire de E vers F. Alors f est différentiable en x_0 ssi f est continue, et $df_{x_0} = f$ pour tout $x_0 \in E$ dans ce cas.*

Proof. On a pour $u \in E$

$$f(x_0 + u) - f(x_0) = f(u).$$

Donc, en posant $df_{x_0} = f$, le membre de droite est bien $df_{x_0}(u) + 0 = df_{x_0}(u) + o(u)$ avec $df_{x_0} = f \in \mathcal{L}(E, F)$.

Si réciproquement f est différentiable,

$$f(x_0 + u) - f(x_0) = df_{x_0}(u) + o(u),$$

où $df_{x_0} \in \mathcal{L}(E,F)$. Donc f est différentiable ssi

$$f(u) = df_{x_0}(u) + o(u).$$

En particulier, $f(u) \to df_{x_0}(0) = 0$ quand $u \to 0$ car df_{x_0} est continue. Comme $o(u) \to 0$ aussi, $f(u) \to 0 = f(0)$, donc f est continue en 0, donc (comme f est linéaire), f est continue. □

Proposition 3. *Soit f, g deux applications différentiables en x_0, et $\lambda \in \mathbb{R}$. Alors $f + \lambda g$ est différentiable et $d(f+\lambda g)_{x_0} = df_{x_0} + \lambda dg_{x_0}$.*

Découle de la linéarité de la limite.

Proposition 4. *Soit Ω' un ouvert de F, $(G, \|\cdot\|_G)$ un autre e.v.n. et $f : \Omega \to \Omega'$ et $g : \Omega' \to G$. Alors $g \circ f$ est différentiable en x_0 si f est différentiable en x_0 et g est différentiable en $f(x_0)$. On a dans ce cas pour $u \in E$,*

$$d(g\circ f)_{x_0}(u) = dg_{f(x_0)}(df_{x_0}(u)).$$

Proof. Soit $u \in E$. Alors pour u dans un voisinage de 0

$$\begin{aligned} g\circ f(x_0+u) =& g(f(x_0) + df_{x_0}(u) + o(u)) \\ =& g(f(x_0)) + dg_{f(x_0)}(df_{x_0}(u) + o(u)) + o(df_{x_0}(u) + o(u)) \\ =& g\circ f(x_0) + \ell(u) + \underbrace{dg_{f(x_0)}(o(u)) + o(df_{x_0}(u) + o(u))}_{\gamma(u)}. \end{aligned}$$

Or,

$$\ell : u \mapsto df_{x_0}(u) \mapsto dg_{f(x_0)}(df_{x_0}(u))$$

est linéaire continue en tant que composée d'applications linéaires continues. Pour montrer que $\ell = d(g\circ f)_{x_0}$, il suffit de montrer $\gamma(u) = o(u)$. On a

$$\begin{aligned} \|dg_{f(x_0)}(o(u))\|_G \leqslant |||dg_{f(x_0)}||| \|o(u)\|_F \leqslant |||dg_{f(x_0)}||| \|u\|_E \varepsilon(u) = o(u), \\ \|o(df_{x_0}(u) + o(u))\|_G = \|df_{x_0}(u) + o(u)\|_F \underbrace{\|\varepsilon'(df_{x_0}(u) + o(u))\|_G}_{=:\varepsilon''(u)} \leqslant (\|df_{x_0}(u)\|_F + \|u\|_E\varepsilon(u))\varepsilon''(u) \\ \leqslant (|||df_{x_0}||| \|u\|_E + \|u\|_E \varepsilon(u))\varepsilon''(u) = o(u). \end{aligned}$$

Donc on a bien $\gamma(u) = o(u) + o(u) = o(u)$. □

Exercice 2. (exercice de base) Soit

$$n(x) = \|x\|_2 = \sqrt{x_1^2 + x_2^2}$$

Montrez que n est différentiable sur $\Omega = \mathbb{R}^2 \setminus \{0\}$, et pas différentiable en 0.

On a $n(x) = g(f(x))$ avec $f(x) = x_1^2 + x_2^2$ de Ω vers $\Omega' := \mathbb{R}_+^*$ et $g(y) = \sqrt{y}$ de Ω' vers $\mathbb{R}$. Comme f et g sont différentiables,avec

$$df_x(u) = 2\langle x, u\rangle, x \in \Omega,$$

et

$$dg_y(u) = ug'(y) = \frac{u}{2\sqrt{y}}$$

pour $u \in \mathbb{R}, y \neq 0$, f est différentiable en tout $x \in \Omega$ et

$$dn_x(u) = dg_{f(x)}(df_x(u)) = \frac{df_x(u)}{2\sqrt{f(x)}} = \frac{\langle x, u\rangle}{\sqrt{\|x\|^2}} = \frac{\langle x, u\rangle}{\|x\|}.$$

Montrer que dans l'exemple précédent, n n'est pas différentiable en 0.

Soit $u \neq 0$. Si n est différentiable en 0,

$$n(tu) = n(0) + t\ell(u) + o(tu) = t\ell(u) + o(t)$$

et donc

$$t\ell(u) = |t|n(u) + o(t).$$

Donc pour $t > 0$, en divisant par t,

$$\ell(u) = n(u)$$

et pour $t < 0$,

$$-\ell(u) = n(u).$$

On aboutit à $n(u) = 0$, ce qui est impossible car n est une norme et $u \neq 0$.

Notation 2. *Si une fonction f est différentiable en tout point $x \in \Omega$, on dit que f est différentiable sur Ω. Dans la définition suivante, on considère l'application $x \mapsto df_x$ de Ω vers $\mathcal{L}(E, F)$.*

Définition 6. *On dit que f est continument différentiable en x_0, ou de classe $\mathcal{C}^1$ en x_0, si f est différentiable sur un voisinage U de x_0 et l'application*

$$\begin{aligned} U &\mapsto (\mathcal{L}(E,F), ||| \cdot |||_{E,F}) \\ x &\mapsto df_x \end{aligned}$$

est continue en x_0.

Comme sur $\mathbb{R}$, la différentiabilité implique la continuité: si f est différentiable au point $x_0 \in \Omega$, alors elle est continue en x_0. Comme sur $\mathbb{R}$, la réciproque est évidemment fausse.

Bien que cette notion semble complexe (continuité vers l'espace $\mathcal{L}(E, F)$), les méthodes pour montrer directement $\mathcal{C}^1$ sont parfois plus simple que pour montrer la différentiabilité, en particulier dans les espaces de dimension finie.

1.2 Cas de la dimension finie

On considère ici que $E = \mathbb{R}^n, F = \mathbb{R}$ pour $n \in \mathbb{N}$. Le cas $F = \mathbb{R}^q$ peut facilement s'en déduire avec la section précédente. Sauf mention contraire, on considère toujours la norme euclidienne

$$\|x\| = \|x\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2}.$$

Les normes étant équivalentes, les résultats sont similaires (aux constantes près) si l'on choisit une autre norme.

On prend toujours Ω un ouvert de $\mathbb{R}^n$, $f : \Omega \to F, x \in \Omega$. On rappelle qu'en dimension finie, tout application linéaire est continue, ce qui fait une chose de moins à prouver pour prouver la différentiabilité.

Proposition 5. *Si f est différentiable en x_0, alors pour $u \in \mathbb{R}^n$,*

$$df_{x_0}(u) = \nabla f(x_0) \cdot u = \sum_{i=1}^n \partial_i f(x_0) u_i.$$

En notant e_i le i-ème vecteur canonique pour $1 \leqslant i \leqslant n$,

$$\partial_i f(x_0) = df_{x_0}(e_i) = \lim_{t \to 0} \frac{f(x_0 + te_i) - f(x_0)}{t}.$$

On a de plus $|||df_{x_0}||| = \|\nabla f(x_0)\|$, de sorte qu'on peut confondre les deux objets via une isométrie.

Proof. Comme df_{x_0} est une forme linéaire, elle s'écrit sous la forme

$$df_{x_0}(u) = df_{x_0}(\sum_{i=1}^n u_i e_i) = \sum_{i=1}^n df_{x_0}(u_i e_i) = \sum_{i=1}^n u_i df_{x_0}(e_i) = \sum_{i=1}^n u_i \partial_i f(x_0).$$

On a ensuite

$$|||df_{x_0}||| = \sup_{u \neq 0} \frac{|df_{x_0}(u)|}{\|u\|} = \sup_{u \neq 0} \frac{|u \cdot \nabla f(x_0)|}{\|u\|} \leqslant \|\nabla f(x_0)\|$$

par l'inégalité de Cauchy-Schwarz, et l'égalité est atteinte pour $u = \nabla f(x_0)$. □

Attention:L'existence de $\partial_i f(x)$ pour chaque x ne garantit pas la différentiabilité de f. On peut même aller plus loin, et construire un contre exemple où pour toute direction $u \in \mathbb{R}^n$, f est dérivable dans la direction u en x, mais f n'est pas différentiable en x.

Posons pour $u \in E$

$$f(x) = \begin{cases} x_1 \text{ si } x_2 = 0 \\ 0 \text{ sinon} \end{cases}$$

On a clairement pour $u = (u_1, u_2) \neq (0, 0)$

$$\lim_{t \to 0} \frac{f(tu) - f(0)}{t} = \begin{cases} 1 \text{ si } u_2 = 0 \\ 0 \text{ sinon} \end{cases}$$

donc f admet une dérivée partielle en 0 dans chaque direction

$$\partial_u f(0) = \begin{cases} 1 \text{ si } u_2 = 0 \\ 0 \text{ sinon} \end{cases}$$

Si f était différentiable, on aurait pour $u \in E$,

$$f(u) = f(0) + df_0(u) + o(u) = \partial_1 f(0) u_1 + \partial_2 f(0) u_2 + o(u) = u_1 + o(u).$$

C'est faux car si l'on étudie la suite $u_n = (1/n, 1/n)$,

$$f(u_n) = 0 \neq \frac{1}{n} + o\left(\frac{1}{n}\right) = (u_n)_1 + o(u_n).$$

Autrement dit,

$$\frac{f(u_n) - f(0) - df_0(u_n)}{\|u_n\|} = \frac{1/n}{\|u_n\|} \to \frac{1}{\sqrt{2}} \neq 0.$$

1.3 Propriétés

Méthodes de base pour montrer qu'une application est différentiable.

Proposition 6. *Soit $f, g : \Omega \to F$, différentiables en $x_0 \in \Omega$ (resp. $\mathcal{C}^1$ en x_0).*

1. *$f + g$ est différentiable en x_0 (resp. $\mathcal{C}^1$ en x_0).*
2. *Si $f : \Omega \to \mathbb{R}, g : \Omega \to G$, fg est différentiable en x_0 (resp. $\mathcal{C}^1$ en x_0). Si de plus $f(x_0) \neq 0$, $\frac{1}{f}g$ est différentiable en x_0 (resp. $\mathcal{C}^1$ en x_0).*
3. *Si $E = \mathbb{R}^n, F = \mathbb{R}^q$ et $f = (f_i)_{i=1,\ldots,q}$ est polynomiale, i.e. $f_i(x_1, \ldots, x_n)$ est un polynome pour tout i, alors f est de classe $\mathcal{C}^1$ (elle est même $\mathcal{C}^\infty$, qu'on définira plus tard).*

Exemple 1 (Exemples de fonctions polynomiales)**.**

- $f(x) = x^3 + 2x - 1$ $(n = q = 1)$
- $f(x_1, x_2, x_3) = (x_2x_3, x_1x_3, x_1x_2, 1)$ $(n = 3, q = 4)$
- $f(M) = \det(M)$, $E = \mathcal{M}_k(\mathbb{R}), k \in \mathbb{N}^* \approx \mathbb{R}^{k^2}, q = 1$ car

$$\det(M) = \sum_{\sigma \in \Sigma_k} \varepsilon(\sigma) m_{i,\sigma(i)} \text{ où } M = (m_{i,j})_{1 \leqslant i,j \leqslant k}$$

- $f(P) = P'$, $P \in \mathbb{R}_k[X] \approx \mathbb{R}^{k+1}, P' \in \mathbb{R}_{k-1}[X] \approx \mathbb{R}^k$, avec

$$P = a_0 + a_1X + \cdots + a_kX^k \approx (a_0, \ldots, a_k) \in \mathbb{R}^{k+1}, P' \approx (a_1, 2a_2, \ldots, ka_k) \in \mathbb{R}^k$$

Proof. 1. Additivité de la limite.

2. Voir TD1, Exo2-1.
3. Si $q = 1$, tout polynome s'écrit comme succession d'additions et de multiplications des fonctions de base $x \to x_i, 1 \leqslant i \leqslant n$. Par exemple,

$$f(x_1, x_2, x_3) = x_1^2x_2^3 - 2x_1^3 + x_3x_1 = (x_1 \cdot x_1)x_2 \cdot x_2 \cdot x_2 - 2x_1 \cdot x_1 \cdot x_1 + x_3 \cdot x_1,$$

donc la différentiabilité de cette fonction s'obtient en appliquant plusieurs fois les points 1 et 2. On verra au chapitre sur les espaces produits qu'il suffit que chaque f_i soit diff. (resp. $\mathcal{C}^1$) pour que f le soit.

□

Chapter 2

Inégalité des accroissements finis

2.1 Enoncé

Pour $x, y \in E$, on note les points compris entre x et y par

$$[x,y] = \{x + t(y-x);\ t \in [0,1]\}$$
$$[x,y[= \{x + t(y-x);\ t \in [0,1[\}$$

Théorème 1. *Soit $x, y \in \Omega$ tels que $[x,y] \subset \Omega$, et f une application différentiable en tout point de $[x,y[$. Alors*

$$\|f(x) - f(y)\|_F \leqslant \sup_{z \in [x,y[} |||df_z||| \|x-y\|_E.$$

Proof. *On note $C = \sup_{z \in [x,y[} |||df_z|||$.

On commence par le cas où $E = \mathbb{R}$, qui comporte en fait toute la complexité du problème. On va montrer que pour tout $\varepsilon > 0, z \in [x,y]$,

$$\|f(x) - f(z)\|_F \leqslant (C+\varepsilon)|x-z|,$$

et on aura le résultat en prenant $y = z$ et en faisant tendre ε vers 0 dans le membre de droite.

On considère z_{max} la borne supérieure de l'ensemble

$$F := \{z \in [x,y] : \forall s \in [x,z], \|f(x) - f(s)\| \leqslant (C+\varepsilon)\|s-x\|\}$$

($z_{max} \geqslant x$ car l'inégalité est vraie pour $z = x$). On remarque que pour $z \in [x,y]$, si $z \in F, \forall z' \in [x,z], \|f(x) - f(z')\| \leqslant (C+\varepsilon)\|z-x'\|$, et donc $z' \in F$. En conséquence F est un intervalle de la forme $[x, z_{max}]$ ou $[x, z_{max}[$.

Supposons $z_{max} < y$: (on va raisonner par l'absurde)

Montrons tout d'abord que F est fermé et donc $z_{max} \in F$. Soit $z_n, n \in \mathbb{N}$ une suite d'élements de F qui converge vers un $z \in \mathbb{R}$ (Il faut donc montrer $z \in F$). Alors pour $s \in [x,z[$, il existe n_0 tel que pour $n \geqslant n_0, s < z_n$, et donc comme $z_n \in F$,

$$\|f(x) - f(s)\| \leqslant (C+\varepsilon)|x-s|.$$

C'est vrai pour chaque $s \in [x, z[$ et en faisant tendre s vers z, on montre que c'est vrai pour $s = z$ par continuité de l'application

$$s \mapsto \|f(x) - f(s)\| - C|x - s|.$$

Donc $z \in F$. Donc $z_{max} \in F$:

$$\|f(x) - f(z_{max})\| \leqslant (C + \varepsilon)|x - z_{max}|.$$

Utilisons la définition de la borne sup: il existe une suite $(z_n)_n$ dans $[x, y]$ telle que $z_n > z_{max}, z_n \to z_{max}$ et

$$\|f(x) - f(z_n)\| > (C + \varepsilon)\|x - z_n\|.$$

Utilisons la différentiabilité Comme f est différentiable en z_{max}, on a quand $n \to \infty$,

$$\|f(z_{max}) - f(z_n)\| = \|df_{z_{max}}(z_{max} - z_n) + o(z_{max} - z_n)\| \leqslant |||df_{z_{max}}||||z_{max} - z_n)| + o(z_{max} - z_n).$$

Par définition du $o(...)$, il existe $\alpha > 0$ tel que pour $|z_{max} - z_n| < \alpha$, $o(z_{max} - z_n) < \varepsilon|z_{max} - z_n|$, et comme $z_n \to z_{max}$, il existe $n_0 \in \mathbb{N}$ tel que pour $n \geqslant n_0$, $|z_{max} - z_n| < \alpha$. En définitive,

$$\|f(z_{max}) - f(z_n)\| \leqslant |||df_{z_{max}}||||z_{max} - z_n| + \varepsilon|z_{max} - z_n|$$
$$\|f(z_{max}) - f(z_n)\| < (|||df_{z_{max}}||| + \varepsilon)|z_{max} - z_n| \leqslant (C + \varepsilon)|z_{max} - z_n|.$$

On a finalement par l'inégalité triangulaire

$$\|f(z_{max}) - f(z_n)\| \geqslant \|f(z_n) - f(x)\| - \|f(z_{max}) - f(x)\| > (C + \varepsilon)\|z_n - x\| - (C + \varepsilon)\|z_{max} - x\| = (C + \varepsilon)\|z_n - z_{max}\|.$$

Contradiction. Donc $z_{max} = y$, ce qui prouve le résultat dans ce cas-là.

Pour le cas général, on se ramène au cas réel en considèrant $\varphi(t) = f(x + t(y - x))$ pour $t \in [0, 1]$ (on a $\varphi(0) = x, \varphi(1) = y$). Par composition des différentielles, φ est différentiable en $t \in [0, 1]$:

$$d\varphi_t(u) = df_{x+t(y-x)}((y - x)u), u \in \mathbb{R}.$$

On a de plus d'après le raisonnement en dimension 1 :

$$\|f(y) - f(x)\| = \|\varphi(1) - \varphi(0)\| \leqslant \sup_{t\in[0,1]} |||d\varphi|||_t \leqslant \sup_{[x,y]} |||df||||y - x\|$$

□

Pour le résultat suivant, on a besoin de la notion d'ensemble ouvert connexe. Intuitivement, il s'agit d'un ensemble Ω où chaque paire de points x, y peuvent être reliés sans sortir de l'ensemble, c'est-à-dire sans lever le stylo si on trace un chemin. Formellement, on dit que Ω est *connexe par arcs* ssi pour $x, y \in \Omega, \exists \gamma : [0, 1] \to \Omega$ continue telle que $\gamma(0) = x, \gamma(1) = y$.

Autrement dit, un ouvert est connexe par arcs ssi il est 'connexe par segments'.

Proposition 7. **Soit Ω un ouvert de E. Alors Ω est connexe par arcs ssi pour tout $x, y \in \Omega$, il existe $x_1, x_2, \ldots, x_q$ tels que $x_1 = x, x_q = y$ et $[x_i, x_{i+1}] \subset \Omega$ pour $1 \leqslant i < q$.*

Proof. Un sens est facile. Pour l'autre sens, suposons qu'un tel γ existe. Comme Ω est ouvert, pour $t \in [0, 1]$, il existe $r_t > 0$ tel que $B(\gamma(t), r_t^-) \subset \Omega$. Comme $\gamma([0, 1])$ est compact, du recouvrement d'ouverts

$$\gamma([0, 1]) \subset \cup_{t\in[0,1]} B(\gamma(t), r_t^-)$$

on peut extraire un recouvrement fini

$$\gamma([0,1]) \subset \cup_{i=1}^{q} B(\gamma(t_i), r_{t_i}^-).$$

On pose $z_i = \gamma(t_i)$, et quitte à les renuméroter on suppose $t_i < t_{i+1}$ pour chaque i. On pose $x_1 = x$. Comme $x = \gamma(0) \in B(0, z_i)$ pour un certain i_1, on a $[x, x_1] \subset \Omega$ On pose $x_2 = z_{i_1}$. On peut montrer par l'absurde qu'il existe $i_2 > i_1$ tel que $B(z_{i_1}, r_{t_{i_1}}^-) \cap B(z_{i_2}, r_{t_{i_2}}^-)$, ou que $y \in B(x_{i_1}, r_{t_{i_1}}^-)$. On peut ainsi construire par induction un chemin $x_0 = x; x_1, \ldots, x_m = y$ tel que $[x_i, x_{i+1}] \subset \Omega$. □

Théorème 2. **Soit une application f différentiable sur un ouvert connexe Ω telle que $df_x = 0$ pour $x \in \Omega$. Alors f est constante.*

Proof. Soit $x \in \Omega$. On va montrer que pour tout $y \in \Omega, f(y) = f(x)$. Si $[x, y] \subset \Omega$, c'est une conséquence immédiate de l'IAF:

$$\|f(y) - f(x)\| \leqslant \sup_{[x,y]} |||df||| \|x - y\| = 0.$$

Sinon, il existe $y_0 = x, y_1, \ldots, y_q = y$ tels que $[y_i, y_{i+1}] \subset \Omega$. On a $\sup_{[y_0,y_1]} |||df||| = 0$ donc $f(y_1) = f(y_0) = f(x_0)$. De même, $f(y_2) = f(y_1)$. En raisonnant par récurrence, on aboutit à $f(y_n) = f(x_0)$. □

Proposition 8. ***

Soient E, F deux espaces vectoriels normés, Ω un ouvert de E, $f : \Omega \to F$ une application continue et x_0 un point de Ω. Si f est différentiable sur $\Omega \setminus \{x_0\}$ et df_x admet une limite ℓ dans (E, F) quand x tend vers x_0, alors f est différentiable au point x_0 et $df_{x_0} = \ell$.

Proof. Soit $\varepsilon > 0$. On a $\ell(h) = \ell(x_0 + h) - \ell(x_0)$

$$\|f(x_0 + h) - f(x_0) - \ell(h)\| = \|(f - \ell)(x_0 + h) - (f - \ell)(x_0)\| \leqslant \sup_{x \in]x_0, x_0+h]} |||df_x - \ell||| \|h\|.$$

Par définition de la limite, il existe $r > 0$ tel que $|||df_x - \ell||| < \varepsilon$ pour $h \in B(x_0, r)$, donc $\|df_x(h) - \ell(h)\| < \varepsilon\|h\|$. On a donc pour $h \in B(0, r)$

$$\frac{\|f(x_0 + h) - f(x_0) - \ell(h)\|}{\|h\|} < \varepsilon$$

Ca implique bien que la limite est 0, et donc comme $\ell \in \mathcal{L}(E, F)$, c'est la différentielle de f au point x_0. □

Remarque 2. *La fonction*

$$f(x) = \begin{cases} x \sin(1/x) \ \textit{si } x \neq 0 \\ 0 \ \textit{sinon} \end{cases}$$

est un exemple où f est différentiable partout mais df_x n'a pas de limite en 0.

2.2 Espaces produits

Etant donné des e.v.n. $(E_i, \|\cdot\|_{E_i})_{i=1,\dots,n}$, on considère l'e.v.n. produit

$$E = E_1 \times \cdots \times E_n,$$

muni de la norme

$$\|x\|_E = \sum_{i=1}^{n} \|x_i\|_{Ei},\ x = (x_i)_{1 \leqslant i \leqslant n} \in E.$$

Définition 7. *Soit F un e.v.n. , $f : E \to F$. On généralise la notion de dérivée partielle (ou différentielle partielle): On dit que f est différentiable suivant la i-ème coordonnée en $x = (x_1, \dots, x_n) \in E$ si l'application partielle*

$$\begin{aligned} f_{[i]} : E_i &\mapsto F \\ y &\mapsto f(x_1, \dots, x_{i-1}, y, x_{i+1}, \dots, x_n) \end{aligned}$$

est différentiable au point x_i (les autres composantes $x_j, j \neq i$ étant fixées). On note la différentielle partielle $d(f_{[i]})_{x_i}$ par

$$d(f_{[i]})_{x_i} = d_i f_x \text{ ou } d_{E_i} f(x) \text{ ou } d_{x_i} f(x).$$

Remarque 3. *Cette notion se distingue de la notion de dérivée partielle dans le fait que E_i peut être de dimension quelconque (pas forcément 1).*

C'est une application linéaire continue de E_i vers F. Si $\dim(E_i) = 1$, alors E_i est généré par un vecteur que l'on nomme u_i, et on retrouve la notion classique de dérivée partielle: $\partial_i f(x) = \frac{\partial f}{\partial x_i}(x)$. Dans la notation, on a donc $d_i f_x : u \mapsto \partial_i f(x) u_i$.

Exemple 2. Soit $\mathbb{R}_n[X]$ l'espace des polynomes de degré au plus n, et l'application

$$\begin{aligned} \varphi : &\mathbb{R}_n[X] \times \mathbb{R} \to \mathbb{R} \\ &(P, x) \mapsto P(x). \end{aligned}$$

Que sont les différentielles partielles par rapport aux espaces $E_1 = E, E_2 = \mathbb{R}$?

Fixons $P \in E, x \in \mathbb{R}$. On a pour $Q \in E$

$$\varphi_{[1]}(Q) = Q(x),$$

on voit que c'est une application linéaire en Q car pour $R, Q \in E, \lambda \in \mathbb{R}$,

$$\varphi_{[1]}(R + \lambda Q) = R(x) + \lambda Q(x) = \varphi_{[1]}(R) + \lambda \varphi_{[1]}(Q).$$

$\varphi_{[1]}$ est de plus continue ($\dim(E_1) < \infty$). Donc φ admet une différentielle (elle-même) partielle par rapport à E_1

$$d_1 \varphi_{(P,x)}(Q) = d(\varphi_{[1]})_P(Q) = \varphi_{[1]}(Q) = Q(x).$$

On a pour P, x fixé

$$\varphi_{[2]}(x) = P(x)$$

est une application dérivable, donc différentiable, avec

$$d_2 \varphi_{(P,x)}(h) = d(\varphi_{[2]})_x(h) = h \varphi'_{[2]}(x) = h P'(x).$$

Théorème 3. *1. Soit un produit d'e.v.n. $F = F_1 \times \cdots \times F_q, f : \Omega \to F$. Alors f se décompose en $f(x) = (f_i(x), 1 \leqslant i \leqslant q)$, où $f_i(x) \in F_i$.*

Dans ce cas, f est différentiable (resp. $\mathcal{C}^1$) en $x \in \Omega$ ssi pour tout $1 \leqslant i \leqslant q$, f_i est différentiable (resp. $\mathcal{C}^1$) en x, et alors

$$df_x(u) = (df_i(x)(u))_{1 \leqslant i \leqslant q}, u \in E.$$

2. **Soit $E = E_1 \times \cdots \times E_n$un produit d'e.v.n. , et $\Omega \subset E$ ouvert. Soit $f : \Omega \to F$. Alors f est <u>de classe $\mathcal{C}^1$</u> <u>sur Ω</u> ssi elle l'est suivant chaque coordonnée, c'es-à-dire pour tout $1 \leqslant i \leqslant n$, l'application*

$$\begin{aligned} E &\mapsto \mathcal{L}(E, F) \\ x &\mapsto d_i f_x \end{aligned}$$

est continue en x. On a alors

$$df_x(h_1, \ldots, h_n) = \sum_{i=1}^{n} d_i f_x(h_i).$$

Proof. 1. Supposons les f_i différentiables. Soit $x \in \Omega, u \in E$.

$$f(x+u) = (f_i(x+u))_i = (f_i(x) + df_{i,x}(u) + o_{F_i}(u))_i = (f_i(x))_i + (df_{i,x}(u))_i + (o_{F_i}(u))_i.$$

L'application

$$\ell : u \mapsto (df_{i,x}(u))_i$$

est clairement linéaire ssi tous les $df_{i,x}$ sont linéaires. Elle est continue ssi

$$\sup_{u \neq 0} \frac{\|\ell(u)\|_F}{\|u\|_E} < \infty \Leftrightarrow \sup_{u \neq 0} \frac{\sum_i \|df_{i,x}(u)\|_{F_i}}{\|u\|_E} < \infty \Leftrightarrow \forall i, \sup_{u \neq 0} \frac{\|df_{i,x}(u)\|_{F_i}}{\|u\|_E} < \infty \Leftrightarrow \text{ chaque } df_{i,x} \text{ est continu}$$

On a aussi que les $x \mapsto df_{i,x}$ sont continues ssi $x \mapsto (df_{i,x})_i$ est continue.

Finalement,

$$\|(o_{F_i}(u))_i\|_F = \sum_i \|o_{F_i}(u)\|_{F_i} = \underbrace{\sum_i \|\varepsilon_{F_i}(u)\|_{F_i}}_{\varepsilon_F(u)} \|u\|_E = o_F(u).$$

ce qui prouve que f est différentiable en x, de différentiable $df_x = \ell$. Dans l'autre sens,

$$f(x+u) = f(x) + df_x(u) + \|u\|_E \varepsilon(u)$$

et il suffit de poser $df_{i,x}$ comme la i-ème composante de df_x et $\varepsilon_i(u)$ comme la i-ème composante de $\varepsilon(u)$ pour avoir l'autre implication.

2. **Implication f $\mathcal{C}^1$ en $x \Rightarrow f_{[i]}$ est $\mathcal{C}^1$ en x:** On a pour $x_i \in E_i, h_i \in E_i$

$$f_{[i]}(x_i + h_i) = f(x + [h_i])$$

où $[h_i] = (0, \dots, 0, h_i, \dots, 0)$. Remarquons que

$$\|[h_i]\|_E = 0 + \dots + 0 + \|h_i\|_{E_i} + 0 = \|h_i\|_{E_i}.$$

. Donc

$$= f(x) + df_x([h_i]) + \|[h_i]\|_E \underbrace{\varepsilon_E([h_i])}_{=:\varepsilon_{E_i}(h_i)}.$$

Remarquons que $\|h\|_E = \|h_i\|_{E_i}$ et

$$\lim_{h_i \to 0} \varepsilon_{E_i}(h) = \lim_{h_i \to 0} \varepsilon_E(h) = 0$$

car $[h_i] \to 0$ dans E quand $h_i \to 0$ dans E_i.

On conclut que $f_{[i]}$ est différentiable avec la linéarité et la continuité de

$$\begin{aligned} &E_i \to E \to F \\ &h_i \mapsto [h_i] \mapsto d_i f_x([h_i]) \end{aligned}$$

comme composition d'applications linéaires continues.

Montrons que $x \mapsto d_i f_x$ est continue: pour $y \in E$,

$$\begin{aligned} |||d_i f_x - d_i f_y||| &= \sup_{h \neq 0} \frac{\|d_i f_x(h) - d_i f_y\|}{\|h\|} = \sup_{h \neq 0} \frac{\|df_x(\|[h_i]) - df_y([h_i])\|}{\|h\|} \\ &\leqslant \sup_{h \neq 0} \frac{|||df_x - df_y||| \|[h_i]\|}{\|h\|} = |||df_x - df_y||| \to_{y \to x} 0 \end{aligned}$$

Réciproque: voir TD. Cette preuve s'appuie fortement sur l'IAF.

□

2.3 Produits d'e.v.n. en dimension finie

Proposition 9. *Supposons que E est de dimension finie. f est de classe $\mathcal{C}^1$ en x ssi pour tout $1 \leqslant i \leqslant d$, $\partial_i f$ existe sur un voisinage de x et est continue en x.*

Proof. C'est une conséquence directe du résultat sur les espaces produits. □

Remarque 4. *En dimension infinie, il est faux de dire que l'existence et la continuité des dérivées partielles dans toutes les directions implique la différentielle globale de l'application (cf. exo sur $\ell^1(\mathbb{R})$ dans le TD2).*

Définition 8. *Soit Ω un ouvert de $\mathbb{R}^n$. Soit $f = (f_k)_{k=1,\dots,p} : \Omega \to \mathbb{R}^p, p \in \mathbb{N}^*$. On a donc que f est différentiable ssi chaque $f_k, 1 \leqslant k \leqslant p$ est différentiable, et*

$$df_x(e_i) = (d(f_k)_x(e_i))_k = (\partial_i f_k(x))_{k=1,\dots,p} \in \mathbb{R}^p.$$

On note

$$J_f(x) = (\partial_i f_k(x))_{1\leqslant k\leqslant p, 1\leqslant i\leqslant n} = \begin{pmatrix} \nabla f_1(x) \\ ... \\ \nabla f_p(x) \end{pmatrix}$$

la **matrice jacobienne** *de f en x, et si $n = p$ on appelle $|\det(J_f(x))|$ son* **jacobien** *au point x.*

Chapter 3

Différentielles d'ordre supérieur

La différentielle est donc une application $x \to df_x$ de $\mathbb{R}^n$ vers $\mathcal{L}(\mathbb{R}^n, \mathbb{R}) \equiv \mathbb{R}^n$. On définit la différentielle seconde comme la différentielle de cette application, si elle existe: pour $h, k \in \mathbb{R}^n$,

$$df_{x_0+k} = df_{x_0} + \underbrace{d(df)_{x_0}(k)}_{\text{linéaire en } k} + \|k\|_{\mathbb{R}^n} \varepsilon_{\mathcal{L}(\mathbb{R}^n,\mathbb{R})}(k)$$

$$df_{x_0+k}(h) = df_{x_0}(h) + \underbrace{d(df)_{x_0}(k)(h)}_{\text{linéaire en } k \text{ et } h} + \|k\|_{\mathbb{R}^n} \times \underbrace{\varepsilon_{\mathcal{L}(\mathbb{R}^n,\mathbb{R})}(k)(h)}_{\substack{\text{linéaire continue en } h, \\ \text{donc } \to 0 \text{ quand } h \to 0}}$$

On doit en particulier introduire l'espace des formes bilinéaires continues.

Définition 9. *Pour un e.v.n. E quelconque, on note $\mathcal{L}^k(E, F)$ l'espace des applications k-linéaires continues. Pour $h \in E, \ell \in \mathcal{L}(E, F)$, on note*

$$\ell(h^k) := \ell(h, h, \dots, h).$$

La linéarité consiste simplement en la linéarité suivant chacune des variables séparément: $\forall 1 \leqslant i \leqslant k, h_1, \dots, h_n \in E, h_i' \in E, \lambda \in \mathbb{R}$,

$$\ell(h_1, \dots, h_{i-1}, h_i + \lambda h_i', \dots, h_k) = \ell(h_1, \dots, h_k) + \lambda \ell(h_1, \dots, h_i', \dots, h_k).$$

La continuité de $\ell \in \mathcal{L}^k(E, F)$ est équivalente à l'existence d'une constante notée $|||\ell||| = |||\ell|||_{\mathcal{L}^k(E,F)}$ (quand elle est la plus petite possible) telle que,

$$\|\ell(h_1, \dots, h_k)\|_F \leqslant |||\ell||| \|h_1\|_E \dots \|h_k\|_E.$$

On posera donc pour $h, k \in E$,

$$d^2 f_{x_0}(h, k) := d(df)_{x_0}(h)(k),$$

c'est un élément de $\mathcal{L}^2(\mathbb{R}^n, \mathbb{R})$, la continuité vient de

$$\|d(df)_{x_0}(h)(k)\| \leqslant |||d(df)_{x_0}(h)|||_{\mathcal{L}(E,F)} \|k\| \leqslant |||d(df)_{x_0}|||_{\mathcal{L}(E,\mathcal{L}(E,F))} \|h\| \|k\|$$

Nous allons définir rigoureusement les ordres successifs de différentiation dans ce chapitre, en commençant par la dimension finie pour éviter certaines difficultés dans un premier temps.

Définition 10. *On construit les différentielles successives par récurrence. On pose* $d^0 f = f$, *et pour* $k > 1$ *on suppose que* f *est* $k-1$ *fois différentiable sur un voisinage de* x *et qu'on a construit* $d^{k-1} f_x \in \mathcal{L}^{k-1}(E, \mathbb{R})$. *On dit que* f *est* k *fois différentiable en* x *si l'application*

$$\begin{aligned} \Omega &\to \mathcal{L}^k(\mathbb{R}^n, \mathbb{R}) \\ y &\to d^{k-1} f_y \end{aligned}$$

est différentiable en x. *On note alors*

$$d^k f_x(h_1, \dots, h_k) = d(d^{k-1} f)_x(h_1)(h_2, \dots, h_k)$$

sa différentielle d'ordre k; $d^k f_x$ *est un élément de* $\mathcal{L}^k(E, F)$.

La continuité de $d^k f_x$ vient de

$$\begin{aligned} \|d^k f_x(h_1, \dots, h_k)\|_F = \|d(d^{k-1} f)_x(h_1)(h_2, \dots, h_k)\|_F &\leqslant \|d(d^{k-1} f)_x(h_1)\|_{\mathcal{L}^{k-1}(E,F)} \|h_2\|_F \dots \|h_k\|_F \\ &\leqslant |||d(d^{k-1} f)_x|||_{E, \mathcal{L}^{k-1}(E,F)} \|h_1\|_F \|h_2\|_F \dots \|h_k\|_F. \end{aligned}$$

Avertissement: Il ne faut pas confondre la continuité d'une application linéaire sur un espace produit et la continuité d'une application bilinéaire. En général, une fonction ne peut pas être les deux en même temps. Par exemple:

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

est bilinéaire sur $\mathbb{R} \times \mathbb{R}$, mais pas linéaire en le couple (x, y), la continuité est donc caractérisée par

$$|f(x, y)| \leqslant C|x||y|$$

alors que

$$g(x, y) = (x + y)$$

n'est pas bilinéaire mais linéaire en (x, y), la continuité sur l'espace produit est alors caractérisée par

$$|g(x, y)| \leqslant C\|(x, y)\|_{\mathbb{R} \times \mathbb{R}} = C(|x| + |y|).$$

3.1 Différentiabilité multiple en dimension finie

$E = \mathbb{R}^n, F = \mathbb{R}, \Omega \subset E, x_0 \in \Omega, f : \Omega \to \mathbb{R}$ différentiable en x_0. Comme $df_{x_0} \equiv \nabla f(x_0) \in \mathbb{R}^n$,

$$\begin{gathered} df_{x_0} \equiv (\partial_i f(x_0))_{i=1,\dots,n} \\ \text{pour } 1 \leqslant j \leqslant n, d(df)_{x_0}(e_j) \equiv (\partial_j (df_{x_0})_i)_{i=1,\dots,n} = (\partial_j(\partial_i f(x_0)))_{i=1}^n \end{gathered}$$

en voyant df_{x_0} comme la forme linéaire $u \mapsto u \cdot \nabla f(x_0)$. On a en particulier pour $1 \leqslant i, j \leqslant n$,

$$d^2 f_{x_0}(e_j)(e_i) = d(df_{x_0})(e_j)(e_i) = \partial_{i,j} f(x_0).$$

Définition 11. *Si f est k fois différentiable en $x \in \Omega$, pour tous indices $1 \leqslant i_1, \dots, i_k \leqslant n$, f admet une dérivée partielle d'ordre k dans les directions $e_{i_1}, \dots, e_{i_k}$, définie par*

$$\partial_{i_1,i_2,\dots,i_k} f(x_0) = \partial_{i_1}(\partial_{i_2}(\dots \partial_{i_k} f(x_0)\dots)) = d^k f_{x_0}(e_{i_1}, \dots, e_{i_k}).$$

Proposition 10. *Pour $k \in \mathbb{N}^*$, on dit que f est de class $\mathcal{C}^k$ en $x_0 \in \Omega$ ssi f est k fois différentiable sur un voisinage de x_0 et $y \to d^k f_y$ est continue en x_0. C'est équivalent à: pour tous indices $1 \leqslant i_1, \dots, i_k \leqslant n$, f admet une dérivée partielle $x \mapsto \partial_{i_1,\dots,i_k} f(x)$ continue en x_0.*

Remarque 5. *Cette caractérisation ne fonctionnera pas en dimension infinie.*

Remarque 6. *Si f est deux fois différentiable en $x \in \Omega$, alors elle est de classe $\mathcal{C}^1$ en x_0. La réciproque est fausse en général.*

Théorème 4 (Théorème de Schwarz)**.** *Soit une application f 2 fois différentiable en $x_0 \in \Omega$. Alors pour $1 \leqslant i, j \leqslant n$,*

$$\partial_{i,j} f(x_0) = \partial_{j,i} f(x_0)$$

Il en découle que si f est k fois différentiable en x_0, pour tous indices $1 \leqslant i_1, \dots, i_k \leqslant n$, et toute bijection σ de $\mathfrak{S}_k$,

$$\partial_{i_1,\dots,i_k} f(x_0) = \partial_{i_{\sigma(1)},\dots,i_{\sigma(k)}} f(x_0).$$

Preuve dans le cas où f est $\mathcal{C}^2$ sur un voisinage de x_0:$\exists r > 0$ tel que f est de classe $\mathcal{C}^2$ sur $B(x_0, r)$.. L'idée est de montrer que pour $h, k \in \mathbb{R}$,

$$\lim_{(h,k)\to 0} \underbrace{\frac{f(x_0 + he_i + ke_j) - f(x_0 + he_i) - f(x_0 + ke_j) + f(x_0)}{hk}}_{A(h,k)} = \partial_i \partial_j f(x_0).$$

Par symétrie, la limite sera également $\partial_j \partial_i f(x_0)$, et par unicité les deux valeurs sont donc égales.

En dérivant (puis intégrant) la fonction $\varphi(t) = f(x_0 + he_i + te_j)$ on montre que

$$f(x_0 + he_i + ke_j) - f(x_0 + he_i) = \varphi(k) - \varphi(0) = \int_0^k \underbrace{\partial_j f(x_0 + he_i + te_j) dt}_{\varphi'(t) = \frac{d}{dt}(df_{x_0 + he_i + te_j})}(te_j)$$

$$\text{avec } h = 0 : f(x_0 + ke_j) - f(x_0) = \int_0^k \partial_j f(x_0 + te_j) dt$$

donc

$$\begin{aligned}
A(h,k) =& \frac{1}{hk}\left(\int_0^k [\partial_j f(x_0+he_i+te_j) - \partial_j f(x_0+te_j)]dt\right)\\
=& \frac{1}{hk}\left(\int_0^1 [\underbrace{\partial_j f(x_0+he_i+ske_j)}_{\psi(h)} - \partial_j \underbrace{f(x_0+ske_j)}_{\psi(0)}]kds\right) \quad \text{avec } sk=t\\
=& \frac{1}{hk}\left(\int_0^1\left(\int_0^h \underbrace{[\partial_i\partial_j f(x_0+te_i+ske_j)]}_{\psi'(t)} dt\right) kds\right)\\
=& \int_0^1\int_0^1 \frac{\partial_i\partial_j f(x_0+uhe_i+ske_j)hdukds}{hk} \quad \text{avec } hu=t\\
=& \int_0^1\int_0^1 \partial_i\partial_j f(x_0+uhe_i+ske_j)duds
\end{aligned}$$

$$\lim_{(h,k)\to(0,0)} A(h,k) = \int_0^1\int_0^1 \lim_{(h,k)\to(0,0)} \partial_i\partial_j f(x_0+uhe_i+ske_j)dsdu = \int_0^1\int_0^1 \partial_i\partial_j f(x_0)dsdu = \partial_i\partial_j f(x_0).$$

Il suffit ensuite d'appliquer la continuité sous le signe intégral, grâce au fait que $\partial_i\partial_j f$ est continue sur un voisinage compact de x_0, et elle y est donc bornée.

Cela implique que dans une dérivée partielle de n'importe quel ordre on peut intervertir deux indices proches, disons i_m et i_{m+1}

$$\begin{aligned}
\partial_{i_1,\dots,i_n} f(x_0) =& \partial_{i_1,\dots,i_{m-1}}(\partial_{i_m,i_{m+1}}(\partial_{i_{m+2},\dots,i_k} f(x_0)))\\
=& \partial_{i_1,\dots,i_{m-1}}(\partial_{i_{m+1},i_m}(\partial_{i_{m+2},\dots,i_k} f(x_0)))\\
=& \partial_{i_1,\dots,i_{m+1},i_m,\dots,i_k} f(x).
\end{aligned}$$

En intervertissant de cette manière plusieurs fois de suite des indices successifs de dérivation, on peut montrer que la valeur ne change pas si on met les indices dans un ordre quelconque $i_{\sigma(1)},\dots,i_{\sigma(k)}$. Par exemple

$$\partial_{3,2,1} f(x) = \partial_{3,1,2} f(x) = \partial_{1,3,2} f(x) = \partial_{1,2,3} f(x).$$

Pour le montrer dans un cadre général, on fait appel à un résultat de la théorie des permutations:

Lemme 1. *Toute permutation $\sigma \in \Sigma_k$ s'écrit comme un produit fini de composition de permutations de la forme $\tau_m, 1 \leqslant m < k$, où τ_m est la permutation qui intervertit m et $m+1$ et laisse inchangé les autres indices.*

Dans l'exemple précédent, la permutation

$$\sigma : 1 \to 3, 2 \to 1, 3 \to 2$$

s'écrit

$$\sigma = \tau_3 \circ \tau_1 \circ \tau_2.$$

Montrons que ce lemme permet de conclure dans le cas général: il existe donc $m_1,\dots,m_q$ tels que

$$\sigma = \tau_{m_1} \circ \tau_{m_2} \circ \dots \circ \tau_{m_q}.$$

On a montré que pout tout m, et indices $i_1,\dots,i_k$,

$$\partial_{i_1,\dots,i_k} f(x) = \partial_{i_{\tau_m(1)},\dots,i_{\tau_m(k)}} f(x).$$

On raisonne alors par récurrence sur q. On vient de montrer que c'est vrai pour $q = 1$. Pour q quelconque, supposons que c'est vrai pour $q-1$. On a alors, en posant $i'_j = \tau_{m_2} \circ \cdots \circ \tau_{m_q}(i_j)$

$$\partial_{i_{\sigma(1)},\dots,i_{\sigma(k)}} f(x) = \partial_{i_{\tau_1(i'_1)},\dots,i_{\tau'_1(i'_k)}} f(x) = \partial_{i'_1,\dots,i'_q} f(x).$$

Les $i'_j, 1 \leqslant j \leqslant k$ s'obtiennent à partir des $i_j, 1 \leqslant j \leqslant k$ par le produit de $q-1$ transpositions. On a donc d'après l'hypothèse de récurrence, l

$$\partial_{i'_1,\dots,i'_k} f(x) = \partial_{i_1,\dots,i_k} f(x),$$

ce qui prouve l'induction, et la propriété est vraie pour tout q.

La preuve du lemme repose sur de la combinatoire et est laissée en exercice.

□

Définition 12. *Soit f deux fois différentiable en $x \in \Omega$. On appelle Hessienne de f en x la matrice symétrique*

$$H_f(x) := (\partial_{i,j} f(x))_{1 \leqslant i,j \leqslant n}.$$

La symétrie vient du théorème de Schwarz. La Hessienne permet de calculer la différentielle d'ordre 2: Comme $d^2 f_x \in \mathcal{L}^2(\mathbb{R}^n, \mathbb{R}), \exists A = (a_{i,j})_{1 \leqslant i,j \leqslant n}$ tels que

$$d^2 f_x(h,k) = \sum_{i,j=1}^n a_{i,j} h_i k_j = h^T \cdot A \cdot k.$$

En particulier, en notant $(e_i)_i$ la base canonique de $\mathbb{R}^n$:

$$a_{i,j} = d^2 f_x(e_i, e_j) = \partial_{i,j} f(x) = \partial_{j,i} f(x).$$

On en déduit en particulier que $d^2 f$ est **symétrique**:

$$d^2 f(h,k) = d^2 f(k,h),$$

nous généraliserons cela en dimension infinie.

Proposition 11 (Développement de Taylor-Lagrange)**.** *Soit $x \in \Omega$ tel que f soit de classe $\mathcal{C}^{k+1}$ en x. Alors il existe $r > 0$ tel que pour $\|h\| \leqslant r$*

$$\left| f(x+h) - f(x) - \sum_{p=1}^k \frac{d^p f_x(h^p)}{p!} \right| \leqslant \sup_{z \in [x,x+h]} \frac{\|d^{k+1} f_z(h^{k+1})\|}{k!} \leqslant C \|h\|_E^{k+1}.$$

(Il n'y a pas besoin de la continuité de $d^{k+1} f$ pour la première inégalité.)

Donnons les deux premiers ordres pour $f : \mathbb{R}^n \to \mathbb{R}$:

$$\text{de classe } \mathcal{C}^2 \text{ sur } [x,x+h]: f(x+h) = f(x) + df_x(h) + O(\|h\|^2) = f(x) + \nabla f(x) \cdot h + O(\|h\|^2)$$

$$\text{de classe } \mathcal{C}^3 \text{ sur } [x,x+h]: f(x+h) = f(x) + \nabla f(x) \cdot h + \frac{1}{2} d^2 f_x(h^2) + O(h^3)$$

$$= f(x) + \sum_{i=1}^n \partial_i f(x) h_i + \frac{1}{2} \sum_{i,j=1}^n \partial_{i,j} f(x) h_i h_j + O(h^3).$$

Proof. Plaçons-nous dans le cas le plus simple: $n = 1, x = 0, h = 1$. On considère la fonction de $[0,1] \to \mathbb{R}$

$$\varphi : t \mapsto f(t) + \sum_{p=1}^{k} \frac{f^{(p)}(t)(1-t)^p}{p!}$$

La dérivée donne des sommes téléscopiques

$$\varphi'(t) = f'(t) + \sum_{p=1}^{k} \frac{-p(1-t)^{p-1} f^{(p)}(t)}{p!} + \sum_{p=1}^{k} \frac{f^{(p+1)}(t)(1-t)^p}{p!} = \frac{f^{(k+1)}(t)(1-t)^k}{k!}$$

Donc par l'IAF

$$|\varphi(1) - \varphi(0)| \leqslant \frac{\sup_{t\in[0,1]} |f^{(k+1)}(t)|}{k!}.$$

C'est-à-dire

$$|f(1) - f(0) - \sum_{p=1}^{k} \frac{f^{(p)}(0)1^p}{p!}| \leqslant \frac{\sup_{t\in[0,1]} |f^{(k+1)}(t)|}{k!}.$$

Dans un e.v.n. E quelconque, soit r tel que f soit k fois différentiables sur $B(x,r)$. On différentie en 0 la fonction $[0,1] \to F$

$$\varphi : t \mapsto f(x+th) + \sum_{p=1}^{k} \frac{(1-t)^p d^p f_{x+th}(h^p)}{p!}$$

La différentielle de $d^p f_{x+th}(h^p)$ en t est

$$u \mapsto ud(d^p f)_{x+th}(h)(h^p) = ud^{p+1} f_{x+th}(h^{p+1})$$

En effet, pour $u \to 0$:

$$d^p f_{x+(t+u)h}(h^p) = d^p f_{x+th}(h^p) + d(d^p f)_{x+th}(uh)(h^p) + o(u)$$

ce qui donne pour $u \in \mathbb{R}$

$$d\varphi_t(u) = \frac{(1-t)^k d^{k+1} f_{x+th}(h^{k+1})}{k!} u$$

et donc par l'IAF ($\|d^{k+1} f_{x+th}(h^{k+1})\| \leqslant |||d^{k+1} f_{x+th}||| \|h\|^{k+1}$)

$$\|f(x+h) - f(x) - \sum_{p=1}^{k} \frac{d^p f_x(h^p)}{p!}\|_F = \|\varphi(1) - \varphi(0)\|_F \leqslant \sup_{t\in[0,1]} \frac{\|d^{k+1} f_{x+th}\|_{\mathcal{L}_s^k(E,F)} \|h\|^{k+1}}{k!}$$

□

3.2 Différentielle d'ordre supérieur, cas général

Le cadre général est très similaire au cadre de la dimension finie, la différence majeure étant qu'on n'utilise pas les dérivées partielles, vu qu'il n'y a en général pas de base canonique. Pour $k \in \mathbb{N}^*$, on munit $\mathcal{L}^k(E,F)$ de la norme

$$|||\ell||| = \sup_{h_1,\ldots,h_k \neq 0} \frac{\|\ell(h_1,\ldots,h_k)\|_F}{\|h_1\|_E \ldots \|h_k\|_E}.$$

Notation 3. *On dit que* $\ell \in \mathcal{L}^k(E,F)$ *est symétrique si l'ordre de ses arguments n'influence pas sa valeur:* $\forall \sigma \in \mathfrak{S}_k, h_1, \dots, h_k \in E$,

$$\ell(h_1, \dots, h_k) = \ell(h_{\sigma(1)}, \dots, h_{\sigma(k)}).$$

On note $\mathcal{L}_s^k(E,F) \subset \mathcal{L}^k(E,F)$ *le sous-e.v.n. des applications* $k-$*linéaires symétriques continues, muni également de* $|||\cdot|||$.

Théorème 5 (Théorème de Schwarz, cas général)**.** *Soit* $x \in \Omega$ *et* $f : \Omega \to \mathbb{R}$ *2 fois différentiable en* x*. Alors* $\forall h, k \in E$,

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

autrement dit, $d^2 f_x \in \mathcal{L}_s^2(E,F)$*. De même, si* f *est* k *fois différentiable en* x*, alors* $d^k f_x \in \mathcal{L}_s^k(E,F)$.

On admet la preuve dans le cas infini. Comme on n'a pas d'intégrale à notre disposition, il faut remplacer l'outil intégral par l'IAF. On peut par contre en général définir une intégrale dans un espace de Banach (i.e. un e.v.n. complet)

Définition 13. *On dit que* f *est de classe* $\mathcal{C}^k$ *en un point* $x \in \Omega$ *si* f *est* k *fois différentiable sur un voisinage de* x *et* $y \to d^k f_y$ *est continue en* x*. On dit que* f *est de classe* $\mathcal{C}^\infty$ *en* x *si elle de classe* $\mathcal{C}^k$ *pour tout* $k \in \mathbb{N}$.

Proposition 12. *Soit* $k \in \mathbb{N}, f$ *une fonction de classe* $\mathcal{C}^k$ *sur un ouvert* $\Omega \subset E$*. On suppose qu'il existe* $\ell_p \in \mathcal{L}_s^p(E,F)$ *pour* $1 \leqslant p \leqslant k$ *telles que pour* $x \in \Omega, h \in E$,

$$f(x+h) = f(x) + \sum_{p=1}^{k} \ell_p(h^p) + o(h^k).$$

Alors pour $1 \leqslant p \leqslant k$, $\ell_p = \frac{1}{p!} d^p f$.

Voir Exo 1.

Proposition 13 (Développement de Taylor-Lagrange)**.** *Soit* $x \in \Omega, h \in E$ *tel que* f *soit de classe* $\mathcal{C}^{k+1}$ *en* x*. Alors il existe* $r > 0$ *tel que pour* $\|h\| \leqslant r$,

$$f(x+h) = f(x) + \sum_{p=1}^{k} \frac{d^p f_x(h^p)}{p!} + O_F(\|h\|^{k+1}).$$

La preuve en dimension finie peut être reprise verbatim.

Proposition 14 (Développement de Taylor-Young)**.** *Soit $x \in \Omega$ tel que $f : \Omega \to F$ soit k fois différentiable en x. Alors*

$$f(x+h) - f(x) - \sum_{p=1}^{k} \frac{d^p f_x(h^p)}{p!} = o_F(h^k).$$

Proof. On admet la preuve, mais si f est de classe $\mathcal{C}^{k+1}$ sur un voisinage de x, alors c'est une conséquence de Taylor-Lagrange, qui est plus fort ($o(\|h\|^k) \leqslant O(\|h\|^{k+1})$). En effet,

$$h \mapsto d^{k+1} f_{x+h} \mapsto |||d^{k+1} f_{x+h}|||$$

est continue en 0 et donc pour r suffisamment petit et $\|h\| \leqslant r$ $|||d^{k+1} f_{x+h}||| \leqslant |||d^{k+1} f_x||| + 1$. Donc

$$\|f(x+h) - f(x) - \sum_{p=1}^{k} \frac{d^p f_x(h^p)}{p!}\| \leqslant \frac{\sup_{\|h\| \leqslant r} |||d^{k+1} f_x|||}{(k+1)!} \|h\|^{k+1} = o(\|h\|^k).$$

□

3.3 Recherche d'extrema

Soit E un e.v.n. .

Définition 14. *Soit $A \subset E$, et $f : A \to \mathbb{R}$. Un point $x \in A$ est un minimum (resp. maximum) local de f s'il existe un voisinage V de x tel que pour tout $y \in V, f(x) \leqslant f(y)$ (resp. $f(x) \geqslant f(y)$).*

On dit que $x \in A$ est un minimum global (resp. maximum global) si pour tout $y \in A$,

$$f(y) \geqslant f(x)(\text{ resp. } f(y) \leqslant f(x)).$$

Le minimum (resp. maximum) est **strict** *si $\forall y \neq x, f(y) > f(x)$ (resp. $f(y) < f(x)$) (avec $y \in V \setminus \{x\}$ pour un extremum local et $y \in A \setminus \{x\}$ pour un extremum global). On dit que x est un extremum s'il est un minimum ou un maximum (global, local, ou strict).*

Définition 15. *On dit qu'une fonction f est minorée (majorée) par un nombre $a \in \mathbb{R}$ sur un ensemble $A \subset \mathbb{R}^d$ ssi pour tout $x \in A, f(x) \geqslant a$ (resp. $f(x) \leqslant a$). Une fonction est bornée si elle est minorée et majorée.*

Remarque 7. *Une application minorée n'admet pas forcément de minimum global, ou même local.*

L'existence et la recherche d'extrema sont des notions essentielles et souvent compliquées en mathématiques; beaucoup de problèmes s'écrivent sous cette forme. En dimension finie on a des résultats d'existence, en particulier le résultat suivant, conséquence directe du fait que l'image d'un compact par une application continue est un compact.

Théorème 6. *Soit K un compact de $\mathbb{R}^d$, et $f : K \to \mathbb{R}$ une fonction continue. Alors f admet au moins un minimum et un maximum global sur K.*

Exercice 3. Soit $K =]0,1[$. Trouvez un exemple de fonction continue bornée $f : K \to \mathbb{R}$ qui n'admet pas de maximum sur K. Donnez un exemple de fonction continue non-bornée sur K.

Exercice 4. Soit $A \subset \mathbb{R}^n$ fermé, $f : A \to \mathbb{R}$ continue et minorée sur A. Supposons que A est borné ou que

$$\lim_{\|x\| \to \infty, x \in A} f(x) = +\infty.$$

Montrez que f admet un minimum global sur A.

Soit $m = \inf\{f(x) : x \in A\}$. Comme f est minorée, $m > -\infty$. Soit

$$R = \sup\{\|x\| : f(x) \in [m, m+1]\}.$$

Comme $K = B(0,R) \cap A$ est un fermé borné, f y atteint son minimum en un point x_0. Comme $f(x) > m+1$ pour $x \notin K, f(x) > f(x_0)$. Comme de plus par définition $f(x) \geqslant f(x_0)$ pour $x \in K$, on a

$$f(x) \geqslant f(x_0) \text{ pour } x \in A.$$

Donc x_0 est un min global.

Théorème 7 (Points critiques)**.** *Soit $\Omega \subset E, f : \Omega \to \mathbb{R}$ différentiable en un point $x \in \Omega$. On dit que x est un point critique (PC), ou point singulier, si $df_x = 0$. Si x est un extremum local, alors x est un point critique.*

La recherche des PC est donc la première étape de la recherche d'extrema. La réciproque de ce théorème est fausse, comme on le verra, et l'existence d'un point critique ne dit rien sur sa nature (maximum, minimum, autre). Il faut pour cela faire un développement à un ordre supérieur. Le théorème est une conséquence du premier point de la proposition suivante:

Proposition 15. *Soit $x \in E$. Supposons qu'il existe $m \in \mathbb{N}$ et $\ell \in \mathcal{L}_s^m(E, \mathbb{R})$ tel que pour $h \in E$*

$$f(x+h) = f(x) + \ell(h^m) + o(\|h\|^m).$$

Alors

1. *Si x est un minimum local, $\forall h \in E, \ell(h^m) \geqslant 0$. Si m est impair (en paticulier $m = 1$) alors $\ell = 0$ (si $m = 1$ cela veut dire que x est un PC).*
2. *Si $\ell(h^m) > 0$ pour $h \neq 0$, alors x est un minimum local strict si $\dim(E) < \infty$.*
3. *Si $\exists c > 0$ tel que $\ell(h^m) \geqslant c\|h\|^m$ pour $h \in E$, alors x est un minimum local strict quelle que soit la dimension.*

Dans tous les autres cas, il faut faire une étude spécifique.

Corollaire 1. *On a avec les même notations*

- *Pour tout extrema local x, alors $df_x = 0$, i.e. $\nabla f(x) = 0$*
- *Si x est un maximum (resp. minimum) local, alors pour un certain m,*

$$df_x = 0\ , \dots , d^{2m-1} f_x = 0; d^{\,2m} f_x(h, \dots , h) \leqslant 0$$

  (resp. $\geqslant 0$) pour tout h

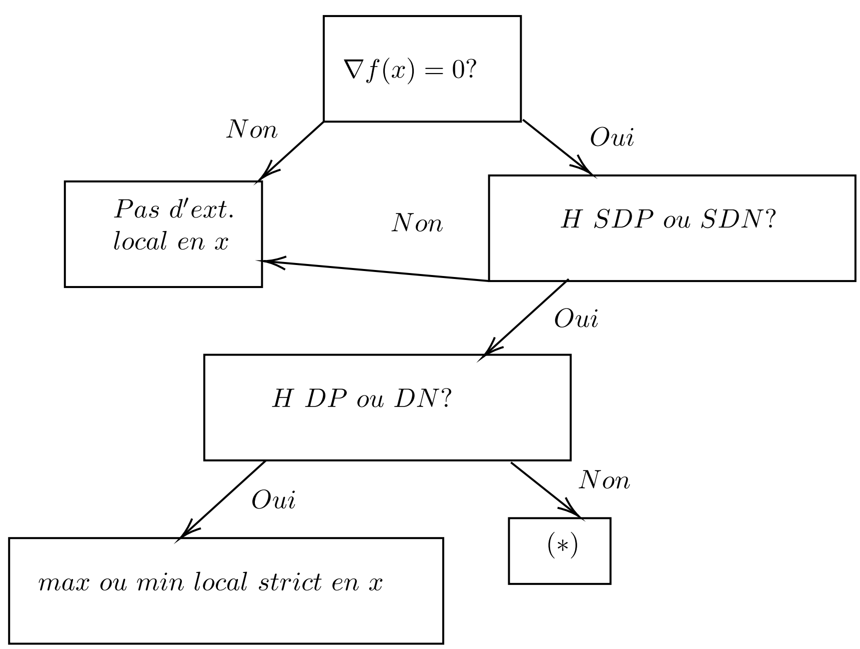
- *Si l'extrema est strict, les inégalités larges sont strictes*

Remarque 8. *Ce type de développement peut-être donné par Taylor-Young, mais on peut aussi le calculer directement sans passer par les différentielles. Le fait que f admette ce développement n'implique pas forcément qu'elle soit m fois différentiable (ou même 2 fois différentiable), voir exo 1 du TD.*

Le cas typique est $m = 2$ pour une fonction qui admet un développement du type

$$f(x + h) = f(x) + \nabla f(x).h + \frac{1}{2} h^T H h + o(h^2).$$

(H sera typiquement la Hessienne, mais ce type de développement ne requiert pas forcément que f soit deux fois différentiable, voir le TD). On a alors le diagramme suivant



Le cas (*) est le plus problématique, il faut étudier $d^k f_x$ pour $k > 2$, mais tous les cas possibles peuvent survenir, il est en particulier possible que $d^k f_x = 0$ pour tout k et x est un min. local strict (voir Exercice 5 ou Exercice 1 de la feuille de TD). Justifions les flèches du diagramme.

- On a $h^T H h + o(h^2) = o(h)$ donc
$$f(x+h) = f(x) + \nabla f(x) \cdot h + o(h),$$
on peut donc appliquer le cas 1) avec $m = 1$ si x est un ext local, et dans ce cas $\ell(h) = \nabla f(x) \cdot h$, qui nous dit que $\ell = 0$, donc $\nabla f(x) = 0$, donc x est un PC. Par contraposée, si $\nabla f(x) \neq 0$, alors x n'est pas un PC, et donc pas un extremum local.

- Si x est un PC on réécrit le développement
$$f(x+h) = f(x) + \ell(h) + o(h^2)$$
avec $\ell(h,k) = \frac{1}{2} h H k^T$, avec $m = 2$ le 1) nous dit que si x est un min local et $\ell(h,h) \geqslant 0$ pour tout h, et donc H est semi définie-positive:
$$\forall h \in E, h^T H h \geqslant 0.$$
Par contraposée, si H n'est pas SDP alors x n'est pas un min. local.

- Si l'on applique le point précédent à $-f$, comme
$$-f(x+h) = -f + h(-H)h^T + o(h)$$
si H n'est pas SDN ($\Leftrightarrow -H$ n'est pas SDP), alors x n'est pas un max. local.

- Appliquons la proposition avec $m = 2$ (i.e. x est un PC et $\ell(h) = \frac{1}{2} h H h^T$). Si H est définie positive (resp. définie négative) alors x est un min local strict (resp. max local strict) car nous sommes en dimension finie.

- La réciproque du point précédent n'est pas forcément vraie, prendre par exemple $f(x) = x^4$ sur $E = \mathbb{R}^1$. 0 est bien un min local/global strict, mais $df_x(h) = f'(0)h = 0$ et $d^2 f_x(h,h) = \frac{1}{2} f''(0)h^2 = 0$.

- Toujours grâce au point 2), si H n'est ni semi-définie positive, ni semi-définie négative, hHh^T n'a pas de signe constant (en h) et x n'est pas un extremum local. En dimension $n = 2$ c'est équivalent à $\det(H) < 0$.

Définition 16 (Rappel sur les matrices symétriques)**.** *Soit H une matrice symétrique $n \times n$.*

- *H est SDP (resp. DP) ssi pour tout $h \in \mathbb{R}^n \setminus \{0\}, hHh^T \geqslant 0$ (resp. > 0).*

- *H est SDP (resp. DP) ssi pour toute valeur propre λ de H, $\lambda \geqslant 0$ (resp. $\lambda > 0$).*

- *$H = (H_{i,j})_{1 \leqslant i,j \leqslant n}$ est DP ssi ses mineurs sont positifs, i.e. pour $1 \leqslant m \leqslant n$,*
$$\det \begin{pmatrix} H_{1,1} & H_{1,2} & ... & H_{1,m} \\ H_{2,1} & H_{2,2} & ... & H_{2,m} \\ ... & ... & ... & ... \\ H_{m,1} & ... & ... & H_{m,m} \end{pmatrix} > 0.$$

- *H est SDN (resp. DN) ssi $-H$ est SDP (resp. DP).*

- *Si $n = 2$, H est SDP ou SDN (resp. DP ou DN) ssi $\det(H) \geqslant 0$ (resp. > 0).*

Avertissement: Une matrice dont toutes les mineures sont négatives n'est pas forcément SDN, par exemple

$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$$

n'est ni SDP ni SDN. Dans l'autre sens, $-I_2$ est DN mais sa seconde mineure est positive. La seule bonne manière de déterminer si une matrice H est (S)DN est d'étudier si $-H$ est (S)DP.

Autre avertissement: si tous les mineurs sont $\geqslant 0$, cela n'implique pas forcément que la matrice est SDP.

Proof. On a donc pour $h \in E$,

$$f(x+h) = f(x) + \ell(h^m) + \|h\|^m \varepsilon(\|h\|)$$

donc pour $t \in \mathbb{R}$ suffisamment petit

$$f(x+th) = f(x) + t^m \ell(h^m) + |t|^m \|h\|^m \varepsilon(t\|h\|)$$

1. Si x est un minimum local ,

   $$0 \leqslant f(x+th) - f(x) = t^m(\ell(h^m) + o(1))$$

   pour $t \to 0$, donc si $\ell(h^m) \neq 0$, $t \mapsto t^m$ ne change pas de signe en 0, donc m est pair et $\ell(h^m) \geqslant 0$. On en déduit en particulier que si m est impair, $\ell(h^m) = 0$ pour tout $h \in E$.

2. On a en fait 2 implique 3: Comme $h \mapsto \ell(h^m)$ est continue et strictement positive sur le compact $K = \{h : \|h\| = 1\}$, elle y est strictement minorée par un $c > 0$. On a donc pour $h \neq 0$

   $$\ell(h^m) = \ell\left(\left(\|h\|\frac{h}{\|h\|}\right)^m\right) = \|h\|^m \ell\left(\left(\frac{h}{\|h\|}\right)^m\right) \geqslant \|h\|^m c.$$

   Donc on se ramène au cas suivant.

3. Soit $r > 0$ tel que $\varepsilon(\|h\|) < c/2$ pour $\|h\| < r$. Donc sur le voisinage $B(0, r)$,

   $$f(x+h) - f(x) \geqslant \ell(h^m) - \|h\|^m c/2 \geqslant c\|h\|^m - c\|h\|^m/2 > 0$$

   donc on a bien un minimum strict sur $B(x, r)$.

□

Exercice 5.

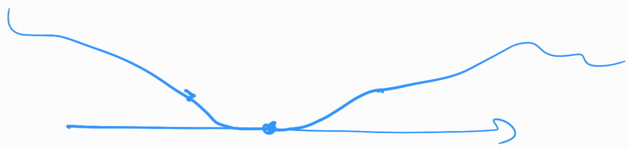
1. Trouver un exemple de fonction f de classe $\mathcal{C}^\infty$ telle que $f(x) = o(x^m) = o(\|x\|^m)$ pour tout m, et 0 est un minimum local strict. $f(x) = e^{-1/x^2}$. Montrer que toutes ses dérivées sont de la forme $P(1/x)f(x)$ où P est un polynome.

2. Trouver un exemple de fonction f différentiable et de point x tel que $d_x f = 0$ et $d^2 f_x(h) > 0$ pour $h \in E \setminus \{0\}$, mais x n'est pas un minimum local.

Voir TD3 exo 6.

Chap. 3 exo17

f C^∞ telle que $f(x)=o(x^n)$ $\forall n\in\mathbb{N}$ et 0 min. local strict.



Il faudrait que $f^{(n)}(0)=0$ pour $n\in\mathbb{N}$.

$$f(x)=\begin{cases} e^{-1/x^2} & \text{pour } x\neq 0\\ 0 & \text{si } x=0\end{cases}$$

f C^0 en 0

→ f C^∞ ?

→ $f^{(n)}(0)=0$?

$f(0)=0 \to$ $\boxed{0 \text{ min. local strict.}}$

$\forall x\neq 0,\ f(x)=e^{-x^{-2}}$

$$f'(x)=2x^{-3}e^{-x^{-2}}$$

$$f''(x)=\left(-6x^{-4}+(2x^{-3})^2\right)e^{-x^{-2}}$$

⋮

On peut montrer par récurrence que $\forall n\in\mathbb{N}$, $\exists P_n$ polynome tel que, $\forall x\neq 0$,

$$\boxed{f^{(n)}(x)=P_n\left(\frac{1}{x}\right)\times e^{-x^{-2}}}$$

Si on montre ça,

$f^{(n)}(x)\xrightarrow[x\to 0]{}0$ car

$\forall k\in\mathbb{N},\ \frac{1}{x^k}e^{-x^{-2}}\xrightarrow[x\to 0]{}0$

→ f est n fois dérivable en 0 pour tout n avec

$$\boxed{f^{(n)}(0)=\lim_{x\to 0} f^{(n)}(x)=0}$$

→ $\boxed{n=0}$ ok

Vrai pour $n\in\mathbb{N}$

$$f^{(n+1)}(x)=\left[\underbrace{\frac{-1}{x^2}P_n'\left(\frac{1}{x}\right)-2x^{-3}P_n\left(\frac{1}{x}\right)}_{P_{n+1}\left(\frac{1}{x}\right)}\right]\times e^{-x^{-2}}$$

où $P_{n+1}(x)=-x^2P_n'(x)-2x^3P_n(x)$ est bien un polynôme

→ Vrai pour n+1

→ Vrai $\forall n\in\mathbb{N}$ par récurrence

Chapter 4

Inversion locale

Définition 17 (Rappel)**.** *Pour E, F deux e.v.n. , on dit que $\ell : E \to F$ linéaire continue est inversible si $\exists \ell' : F \to E$linéaire continue telle que $\ell \circ \ell' = Id_F$ et $\ell' \circ \ell = Id_E$. On note $\ell' = \ell^{-1}$ et E et F sont alors déclarés isomorphes (en tant qu'e.v.n.). On note $\mathcal{L}(E, F)^* \subset \mathcal{L}(E, F)$ l'espace des applications linéaires continues inversibles d'inverse continue.*

Proposition 16 (rappel)**.** *Si E, F sont des espaces de Banach, $\mathcal{L}(E, F)^*$ est un ouvert de $\mathcal{L}(E, F)$ et l'application*

$$\begin{aligned} \mathcal{L}(E, F)^* &\to \mathcal{L}(F, E)^* \\ \ell &\to \ell^{-1} \end{aligned}$$

est continue.

Exercice 6. Soit $E = \mathcal{C}^1([0, 1]; \mathbb{R})$ muni de la norme

$$\|f\|_1 = \|f\|_\infty + \|f'\|_\infty, f \in E$$

et $F = \mathcal{C}^0([0, 1]; \mathbb{R})$ muni de la norme $\| \cdot \|_\infty$. On considère une fonction $f \in E$ et l'application linéaire

$$\ell_f(g) = f'g' + fg.$$

Montrer que si $f = 1$, ℓ_f est inversible. Sous quelles conditions sur f ℓ_f est-elle inversible?

On remarque que résoudre $h = \ell_f(g)$ revient à résoudre l'équation différentielle

$$f'(x)h'(x) + f(x)h(x) = g(x).$$

4.1 $\mathcal{C}^1$-difféomorphismes

E et F sont désormais des espaces de Banach.

Définition 18. *Soit $\Omega \subset E, f : \Omega \to F$ et $\Omega' = f(\Omega)$. On dit que f est un $\mathcal{C}^1$-difféomorphisme sur Ω (ou de Ω sur Ω') si*

1. *f est une bijection $\Omega \to \Omega'$*
2. *f est de classe $\mathcal{C}^1$ sur Ω*
3. *f^{-1} est de classe $\mathcal{C}^1$ sur Ω'.*

Proposition 17. *Si f est un $\mathcal{C}^1$-difféomorphisme sur Ω, alors pour $x \in \Omega$, $df_x \in \mathcal{L}(E,F)^*$ et*

$$(df_x)^{-1} = d(f^{-1})_{f(x)}, x \in \Omega.$$

Dans cette dernière écriture, -1 peut représenter l'inversion de la fonction, ou l'inversion d'une application linéaire. De plus, E et F sont alors isomorphes via l'application df_x (pour n'importe quel $x \in E$), ils ont en particulier nécessairement de la même dimension (finie ou infinie).

Proof. Soit $x \in \Omega$. On doit montrer que $d(f^{-1})_{f(x)} \circ df_x = Id_E$ et $df_x \circ d(f^{-1})_{f(x)} = Id_F$.

On a $f^{-1} \circ f(x) = x$ sur Ω. Donc par la formule de différentiation d'une application composée, pour $u \in E$,

$$d(f^{-1})_{f(x)} \circ df_x(u) = u$$

donc $d(f^{-1})_{f(x)} \circ df_x = Id_E$ est l'application identité sur E. De même, $f \circ f^{-1}(y) = y$ sur Ω', et donc $df_{f^{-1}(y)} \circ d(f^{-1})_y = Id_F$, ce qui prouve $df_x \circ d(f^{-1})_{f(x)}$ en posant $y = f(x)$. □

Exemple 3.

1. La fonction $x \mapsto \ln(x)$ est une bijection de classe $\mathcal{C}^1$ de $]0, +\infty[$ sur $\mathbb{R}$, avec $\ln^{-1} = \exp$. Comme $\exp$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}$, $\ln$ est bien un $\mathcal{C}^1$ difféomorphisme sur $]0, \infty[$. On a $d\ln_x(u) = \frac{1}{x}u$ pour $u \in \mathbb{R}$, qui est bien inversible d'inverse

$$(d\ln_x)^{-1}(u) = xu; u \in \mathbb{R}.$$

Donc pour $y \in \mathbb{R}$,

$$(d\exp_y^{-1})(u) = (d\ln_{\exp(y)})^{-1}(u) = \exp(y)u, u \in \mathbb{R}.$$

ce qui prouve bien que $\exp' = \exp$.

2. La fonction $f : x \mapsto x^3$ est bijective de $\mathbb{R}$ vers $\mathbb{R}$ et de classe $\mathcal{C}^1$ sur $\mathbb{R}$, mais ça n'est pas un $\mathcal{C}^1$-difféomorphisme car $f^{-1}(x) = x^{\frac{1}{3}}$ n'est pas de classe $\mathcal{C}^1$ sur $\mathbb{R}$. On peut le montrer avec

$$(f^{-1})'(x) = \frac{1}{3}x^{-2/3} \xrightarrow[x \to 0, x>0]{} +\infty,$$

alors qu'une application continue sur $[-1, 1]$ doit être bornée sur $[-1, 1]$. On peut par contre montrer que f est un $\mathcal{C}^1$-difféomorphisme de $\mathbb{R}^*$ vers $\mathbb{R}^*$.

Le théorème d'inversion locale fournit une réciproque locale au théorème précédent:

Théorème 8. *Soit $f : \Omega \to F$ une fonction de classe $\mathcal{C}^1$ en $x_0 \in \Omega$. Alors si df_{x_0} est inversible, il existe un voisinage U de x_0 dans Ω tel que f soit un $\mathcal{C}^1$-difféomorphisme de U dans $f(U)$ ($f(U)$ est un voisinage ouvert de $f(x_0)$ dans V).*

$f(U)$ est un voisinage car il est ouvert c'est l'image réciproque de l'ouvert U par l'application continue f^{-1}, i.e. $U = (f^{-1})^{-1}(U)$.

Exercice 7. On reprend l'exemple précédent et on pose $\varphi(f) = (f')^2 + f^2$. Il faut montrer que φ est différentiable avec

$$d\varphi_f(h) = \ell_f(h) = 2f'h' + 2fh.$$

Soit f telle que ℓ_f est inversible, et $g_0 = \varphi(f)$. Alors d'après le théorème précédent, il existe un voisinage V de g_0 et U de f tels que φ soit bijective de U vers V (avec $V = \varphi(U)$). L'inverse $\varphi^{-1}(g)$ nous donne donc pour tout $g \in V$ une solution de l'équation différentielle

$$(f')^2 + f^2 = g.$$

Heuristique

On dispose de la fonction f au voisinage de x_0, et la différentielle df_{x_0}, inversible. Soit y dans le voisinage de 0. On cherche $x = f^{-1}(y)$. La première tentative consiste à écrire

$$y = f(x) = f(x_0) + df_0(x) + o(x - x_0) \approx f(x_0) + df_0(x - x_0).$$

i.e.

$$x - x_0 \approx df_0^{-1}(y - f(x_0))$$

On pose donc

$$x_1 = x_0 + df_0^{-1}(y - f(x_0)).$$

Approximation Etant donné une approximation x_0 du point x tel que $f(x) = y$, on peut améliorer cette approximation via la fonction φ_y:

$$\varphi_y : x_0 \mapsto x_1 = x_0 + df_0^{-1}(y - f(x_0)).$$

Si cette heuristique est exacte, on peut donc sans cesse améliorer l'approximation en posant

$$\begin{aligned} x_2 =& \varphi_y(x_1) \\ & \dots \\ x_{n+1} =& \varphi_y(x_n); n \in \mathbb{N} \end{aligned}$$

et espérer que $x_n \to x$ quand $n \to \infty$.

Cette heuristique sera justifiée par le fait que φ_y est contractante si x_0 est suffisamment proche de x et converge donc par le théorème de Picard vers sont unique point fixe x: en effet, $\varphi_y(x) = x$.

On observe que

$$d(\varphi_y)_z(u) = u + df_0^{-1}(-df_z(u)) = df_0^{-1}(df_0(u) - df_z(u)) \approx 0.$$

Donc φ_y a des faibles variations tant qu'on est dans un voisinage de x_0. Ce type de système dynamique a tendance à converger rapidement vers sont point fixe, s'il en existe: $x_n \to x$ défini par $\varphi_y(x) = x$, et donc $f(y) - x = 0$.

Théorème 9 (Théorème du point fixe de Picard)**.** *Soit A un fermé non-vide d'un espace de Banach E et $\varphi : A \to A$ telle qu'il existe $c \in]0,1[$ tel que*

$$\|\varphi(x) - \varphi(y)\|_E \leqslant c\|x - y\|_E$$

pour $x, y \in E$ (on dit que φ est c-contractante). Alors il existe un unique $x \in A$ tel que $\varphi(x) = x$. Ce point fixe est la limite de n'importe quelle suite (x_n) définie par

$$\begin{cases} x_0 \in A \\ x_{n+1} = \varphi(x_n). \end{cases}$$

On montre les faits suivants:

1. Il Existe $r > 0$ tel que si $y \in B(0,r)$, $x_n \to x$. Ainsi par passage à la limite,

$$x = x + df_0^{-1}(y - f(x))$$

et donc $df_0^{-1}(y - f(x)) = 0$, donc $y = f(x)$ car df_0^{-1} est injective, on a solutionné le problème. On peut définir l'application réciproque

$$\begin{aligned} f^{-1} :& B(0,r) \to E \\ & y \to \lim_n x_n. \end{aligned}$$

2. Cette application est-elle $\mathcal{C}^1$? bijective? Fournit-elle l'unique solution?

Preuve du théorème d'inversion locale. Il suffit de traiter le cas où $x_0 = 0, f(x_0) = 0$. Le cas général s'en déduit facilement avec la fonction $\tilde{f}(x) = f(x - x_0) - f(x_0)$. Pour montrer que φ_y est contractante sur une boule $B(0,r)$, on utilise l'IAF: pour $x \in B(0,r)$,

$$\|\varphi_y(x) - \varphi_y(x')\| \leqslant \sup_{z \in B(0,r)} |||d(\varphi_y)_z||| \|x - x'\| \leqslant |||df_0^{-1}||| |||df_0 - df_z||| \|x - x'\|.$$

Il faut donc trouver $r > 0$ tel que $\|d\varphi\| \leqslant 1/2 < 1$ (par exemple). Comme $z \mapsto df_z$ est de classe $\mathcal{C}^1, df_z \to df_0$ quand $z \to 0$ et ilexiste $r_0 > 0$ tel que pour tout $r < r_0, z \in B(0, r_0)$, $|||df_0 - df_z||| < |||df_0^{-1}|||^{-1}/2$. On pose donc $A = B(0, r_0)$.

Il faut aussi s'assurer que $\varphi(A) \subset A$. On a pour $r_1 > 0, y \in B(0, r_1), z \in A = B(0, r_0)$,

$$\|\varphi_y(z)\| = \|\varphi_y(z) - \varphi_y(0) + \varphi_y(0)\| \leqslant \sup_{B(0,r_0)} |||d\varphi_y||| \|z\| + |||df_0^{-1}||| \|y\| \leqslant \frac{r_0}{2} + |||df_0^{-1}||| r_1.$$

On choisit donc $r_1 < |||df_0^{-1}|||^{-1} r_0/2$, ce qui assure $\varphi_y(z) \in B(0, r_0)$ pour $y \in B(0, r_1), z \in B(0, r_0)$.

On applique le théorème du point fixe avec $A = B(0, r_0)$, et on a $x \in A$ tel que $\varphi_y(x) = x$, c'est-à-dire $f(x) = y$. Comme c'est l'unique point fixe pour chaque $y \in B(0, r_1)$, cette opération réalise bien une bijection de $B(0, r_1)$ vers $f^{-1}(B(0, r_1))$. En posant $U = f^{-1}(B(0, r_1))$, on a bien une bijection de U vers $f(U)$.

Montrons que f^{-1} est de classe $\mathcal{C}^1$ sur $B(0, r_1)$, de différentielle $y \mapsto (df_{f^{-1}(y)})^{-1}$: soit $y \in B(0, r_1), x = f^{-1}(y), h \in F$ suffisamment petit. Soit $z = z_h$ tel que $f^{-1}(y + h) = x + z$, d'où

$h = f(x+z) - f(x)$. Alors

$$\begin{aligned}\|f^{-1}(y+h) - f^{-1}(y) - (df_x)^{-1}(h)\| =& \|x+z-x-(df_x)^{-1}(f(x+z)-f(x))\| \\ =& \|df_x^{-1}(df_x)(z) - df_x^{-1}(f(x+z)-f(x))\| \\ =& \|df_x^{-1}(df_x(z) - f(x+z) + f(x))\| \\ \leqslant& |||df_x^{-1}||| \|f(x+z) - f(x) - df_x(z)\| \\ \leqslant& \sup_{x \in f^{-1}(B(0,r_1))} |||df_x^{-1}||| o(z)\end{aligned}$$

Il faut donc contrôler $|||df_x^{-1}|||$. $x \mapsto |||df_x^{-1}|||$ est continue comme composée:

$$\begin{aligned}E \to \mathcal{L}(E,F)^* \to \mathcal{L}(F,E)^* \to \mathbb{R} \\ x \to df_x \to df_x^{-1} \to |||df_x^{-1}|||.\end{aligned}$$

On choisit $r_2 \in]0, r_1[$ tel que $\sup_{x \in B(0,r_2)} |||df_x^{-1}||| < \infty$. Il reste à montrer que $o(z) = o(h)$ pour conclure. On utilise $z = f^{-1}(y+h) - f^{-1}(y)$. Il faudrait donc montrer que f^{-1} est L-lipschitzienne, ce qui prouverait $o(\|z\|) = o(L\|h\|) = o(h)$.

$$\begin{aligned}\|z\| =& \|x+z-x\| = \|\varphi_{y+h}(x+z) - \varphi_y(x)\| \\ \leqslant& \|\varphi_{y+h}(x+z) - \varphi_y(x+z)\| + \|\varphi_y(x+z) - \varphi_y(x)\| \\ \leqslant& \|df_0^{-1}(y+h) - df_0^{-1}(y)\| + \sup_{B(0,r_2)} |||d\varphi_y||| \|z\| \\ \leqslant& |||df_0^{-1}||| \|h\| + \frac{1}{2}\|z\|.\end{aligned}$$

On en déduit $\|z\| \leqslant 2|||df_0^{-1}||| \|h\|$, ce qui conclut la preuve. On a en fait montré dans ce dernier calcul que f^{-1} est $2|||df_0^{-1}|||$-Lipschitzienne.

□

Théorème du point fixe. Il suffit de montrer que (x_n) est de Cauchy pour qu'elle converge. On a pour $n, p \in \mathbb{N}$

$$\|x_{n+p} - x_n\| \leqslant \sum_{i=1}^{p} \|x_{n+i} - x_{n+i-1}\|$$

On montre par récurrence que pour $m \in \mathbb{N}$,

$$\|x_{m+1} - x_m\| \leqslant \|x_1 - x_0\| c^m.$$

Donc

$$\|x_{n+p} - x_n\| \leqslant \sum_{i=1}^{p} c^{n+i} \|x_1 - x_0\| \leqslant C c^n$$

où

$$C = \|x_1 - x_0\| \sum_{i=0}^{\infty} c^i < \infty.$$

Donc pour $\varepsilon > 0$, $\exists n_0 \in \mathbb{N}$ tel que pour $n \geqslant n_0, Cc^n < \varepsilon$, et donc $\forall p \in \mathbb{N}$,

$$\|x_{n+p} - x_n\| \leqslant \varepsilon,$$

ce qui prouve que (x_n) est bien de Cauchy, donc convergente comme E est un Banach. La limite satisfait

$$x = \lim_n x_{n+1} = \lim_n \varphi(x_n) = \varphi(x)$$

car φ est continue (car c-lipschitzienne). □

Théorème 10 (Théorème d'inversion globale)**.** *Soit $f : \Omega \to F$ de classe $\mathcal{C}^1$ sur Ω. Alors f est un $\mathcal{C}^1$-difféomorphisme de Ω vers $f(\Omega)$ ssi f est injective et $\forall x \in \Omega, df_x \in \mathcal{L}(E,F)^*$.*

Proof. Il n'y a que la réciproque à prouver. Soit donc f injective sur Ω, avec df_x inversible pour $x \in \Omega$, et soit $f^{-1} : f(\Omega) \to \Omega$ son application réciproque. Il suffit donc de montrer que f^{-1} est de classe $\mathcal{C}^1$.

Soit $y \in f(\Omega), x = f^{-1}(y)$. On sait d'après le résultat précédent qu'il existe un voisinage U de x tel que f soit un $\mathcal{C}^1$-difféomorphisme $U \mapsto f(U)$. Par unicité, la réciproque de f sur U est la restriction de f^{-1} à $f(U)$. En particulier, f^{-1} est de classe $\mathcal{C}^1$ sur $f(U)$, avec $d(f^{-1})_y = (df_x)^{-1}$. Donc f^{-1} est de classe $\mathcal{C}^1$ en tout $y \in f(\Omega)$. □

4.2 Théorème des fonctions implicites

Rappel Une fonction $f : E \times F \to G$ est de classe $\mathcal{C}^1$ en $x = (x_1, x_2)$ ssi les applications partielles

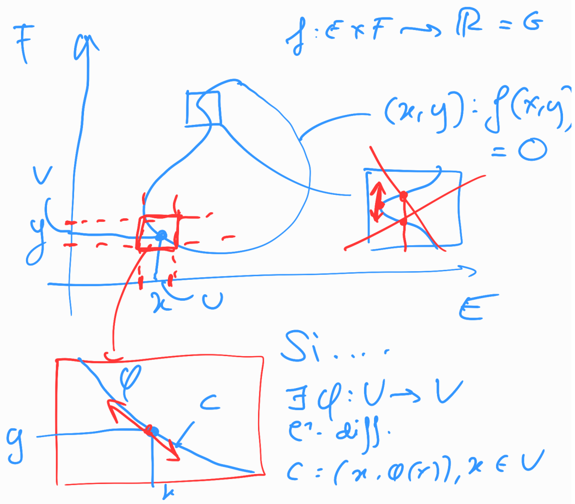
$$y \mapsto f(x_1, y) \text{ et } y \mapsto f(y, x_2)$$

sont de classe $\mathcal{C}^1$ en resp. x_2, x_1, on note alors resp. $d_2 f_x$ et $d_1 f_x$ leurs différentielles, et on a pour $h = (h_1, h_2) \in E \times F$

$$df_x(h_1, h_2) = d_1 f_x(h_1) + d_2 f_x(h_2).$$

Si réciproquement f est différentiable (resp. $\mathcal{C}^1$) en x, alors les applications partielles sont différentiables (resp. $\mathcal{C}^1$) en x de différentielles

$$d_1 f_x(h_1) = df_x(h_1, 0) \text{ et } d_2 f_x(h_2) = df_x(0, h_2).$$



Théorème 11 (Théorème des fonctions implicites)**.** *Soit E, F, G tois espaces de Banach, $\Omega \subset E \times F$ ouvert, et $f : \Omega \to G$ de classe $\mathcal{C}^1$. Soit $(x_0, y_0) \in \Omega$ tel que $f(x_0, y_0) = 0$; et*

$$d_2 f_{(x_0,y_0)} \in \mathcal{L}(F, G)^*.$$

Alors il existe U_{x_0}, U_{y_0} des voisinages de x_0, y_0 tels que $U_{x_0} \times U_{y_0} \subset \Omega$ et une fonction $\varphi : U_{x_0} \to U_{y_0}$ telle que

$$\forall (x, y) \in U_{x_0} \times U_{y_0}, (f(x, y) = 0 \Leftrightarrow y = \varphi(x)).$$

On a de plus: φ est de classe $\mathcal{C}^1$ en $x \in U_{x_0}$, et pour $h \in E$,

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

Proof. On pose

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

qui vérifie $g(x_0, y_0) = (x_0, 0)$. L'idée est de montrer que g est un $\mathcal{C}^1$-difféomorphisme d'un voisinage $U_{x_0} \times U_{y_0}$ vers un voisinage $V_{x_0} \times V_0$, car on aura ensuite $\forall x \in V_{x_0}, \forall y \in U_{y_0}$,

$$f(x, y) = 0 \Leftrightarrow (x, y) = g^{-1}(x, 0) \Leftrightarrow x \in V_{x_0}, y = \varphi(x) := g_2^{-1}(x, 0)$$

autrement dit $\varphi(x)$ est la seconde composante de $g^{-1}(x,0)$, bien définie et de classe $\mathcal{C}^1$ sur V_{x_0} car g^{-1} est de classe $\mathcal{C}^1$. On a $V_{x_0} = U_{x_0}$ car $V_{x_0} \times V_0 = g(U_{x_0} \times U_{y_0}) = U_{x_0} \times f(U_{x_0} \times U_{y_0})$.

D'après le théorème d'inversion locale, il suffit donc de montrer que $dg_{(x_0,y_0)} \in \mathcal{L}(E \times F, E \times G)^*$ (g est bien de classe $\mathcal{C}^1$ car ses composantes sont de classe $\mathcal{C}^1$). On a pour $h = (h_1, h_2) \in E \times F$,

$$dg_{(x_0,y_0)}(h_1, h_2) = (h_1, df_{(x_0,y_0)}(h)) = (h_1, d_1 f_{(x_0,y_0)}(h_1) + d_2 f_{(x_0,y_0)}(h_2))$$

Pour $h' = (h'_1, h'_2) \in E \times G$, on a bien

$$dg_{(x_0,y_0)}(h) = h' \Leftrightarrow h'_1 = h_1 \text{ et } d_2 f_{(x_0,y_0)}(h_2) = h'_2 - d_1 f_{(x_0,y_0)}(h_1) \Leftrightarrow h'_1 = h_1 \text{ et } h_2 = d_2 f^{-1}_{(x_0,y_0)}(h'_2 - d_1 f_{(x_0,y_0)}(h'_1)).$$

Donc $dg_{(x_0,y_0)}$ est linéaire continue inversible, donc c'est un isomorphisme.

Pour prouver le dernier point, on différencie $f(x, \varphi(x)) = f(u(x)) = 0$ avec $u(x) = (x, \varphi(x))$: pour $h \in E$,

$$0 = df_{u(x)}(du_x(h)) = df_{(x,\varphi(x))}(h, d\varphi_x(h)) = d_1 f_{(x,\varphi(x))}(h) + d_2 f_{(x,\varphi(x))}(d\varphi_x(h)).$$

On en déduit

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

$\square$

Exemple 4. TD4 exo8

Part II

Equations différentielles

Chapter 5

Equations différentielles: existence et unicité

Soit E un espace de Banach (en général, $\dim(E)$ sera finie).

5.1 Vocabulaire et définition

Soit $I \subset \mathbb{R}$ un intervalle ouvert, $\Omega_1, \ldots, \Omega_{n-1}$ des ouverts de E. Soit $f : I \times \Omega_1 \times \cdots \times \Omega_{n-1} \to E$. On note E_f l'équation différentielle

$$y^{(n)}(t) = f(t, y(t), y'(t), \ldots, y^{(n-1)}(t)). \tag{E_f}$$

Définition 19. *On appelle solution la donnée d'un couple (J, y) où $J \subset I$ et $y : J \to E$ est n fois différentiable et satisfait (E_f) pour $t \in J$.*

On dit que la solution est maximale si pour toute solution (J', z) telle que $J \subset J'$ et $z = y$ sur J, alors $J = J'$ (et donc $z = y$).

Attention! Il peut y avoir plusieurs solutions maximales distinctes. Exemple: $y = 0$ et $y = 1$ sont solutions maximales sur $\mathbb{R}$ de l'équation $y' = 0$... Des conditions de type Cauchy-Lipschitz peuvent imposer l'unicité.

- Si $E = \mathbb{R}$, l'équation est dite **scalaire** (ex: $y'(t) = ty(t) + e^{-t}$ sur $I = \mathbb{R}$).
- Si f ne dépend pas de t, E_f est dite **autonome (ex: $y'(t) = y(t)$ sur $I = \mathbb{R}$).**
- Si $f(t, y_0, \ldots, y_{n-1})$ est linéaire en $Y = (y_0, \ldots, y_{n-1})$, E_f est dite **linéaire sans second membre** (ex: $y'' = ty'(t) - y(t)$).
  - Si de plus f ne dépend pas de t, E_f est dite **linéaire à coeficients constants** (ex: $y''(t) = 3y(t)$).
- Si $f(t, y_0, \ldots, y_{n-1}) = g(t, y_0, \ldots, y_{n-1}) + x(t)$ pour une fonctions g linéaire en Y et une fonction $x : I \to E$, f est dite **linéaire avec second membre** (à coefficients constants si g ne dépend pas de t) (ex: $y''(t) = t + y'(t) - 2ty(t)$).

Remarque 9. *Si l'ED est autonome, alors pour toute solution y et $a \in \mathbb{R}$, la fonction $y_a(t) = y(t-a)$ est une solution (pas nécessairement distincte, cf. cas où $y =$Cte).*

Exemple 5. On considère l'ED $y' = -y^2$ sur $I = \mathbb{R}$, d'ordre $n = 1$ sur $E = \mathbb{R}$, correspondant à $f(t, y) = -y^2$. C'est donc une équation autonome non-linéaire. On a $y(t) = 1/t$ qui est solution sur $J_+ =]0, \infty[$ ou sur $J_- =]-\infty, 0[$. De même, $y_a(t) = y(t-a) = \frac{1}{t-a}$ est solution sur $J_{a,+} =]a, \infty[$ et sur $J_{a,-} =]-\infty, a[$. Toutes ces solutions sont maximales car on ne peut pas les prolonger par continuité en a. La solution $y \equiv 0$ est solution (évidemment maximale) sur $J = I$.

On appelle problème de Cauchy, étant donné des conditions initiales $t_0 \in \mathbb{R}$ et $x_0 \in \mathbb{R}$, la détermination d'une solution y de l'ED qui vérifie

$$y(t_0) = x_0.$$

On voit que si $x_0 \neq 0$, alors il y a une unique solution en résolvant $\frac{1}{t_0-a} = x_0$, i.e. $a = t_0 - \frac{1}{x_0}$. Si $x_0 = 0$, la solution est $y \equiv 0$. On verra grâce au Théorème de Cauchy-Lipschitz que ce sont les uniques solutions.

On dit que E_f est une ED d'ordre n sur E. Remarquons tout d'abord que quitte à changer d'espace, toute ED se ramène à une ED d'ordre 1: soit

$$Y(t) = (y(t), y'(t), \dots, y^{(n-1)}(t))$$

à valeurs dans E^{n-1}, alors y est solution de E_f ssi Y est solution de E_F

$$Y'(t) = F(t, Y) \tag{E_F}$$

avec

$$F(t, x_0, \dots, x_{n-1}) = (x_0, \dots, x_{n-1}, f(t, x_0, \dots, x_{n-1})).$$

En effet, $Y'(t) = F(t, Y)$ s'écrit

$$\begin{cases} y^{(n)}(t) = & F_n(t, Y) = f(t, y(t), \dots, y^{(n-1)}(t)) \\ y^{(n-1)}(t) = & F_{n-1}(t, Y) = y^{(n-1)}(t) \\ \dots \end{cases}$$

donc seule la n-ème composante est non-triviale.

Exemple 6. L'équation $y''(t) = 3y'(t) + 2y(t) + x(t)$ s'écrit aussi

$$Y'(t) = (y''(t), y'(t)) = F(t, Y(t))$$

avec $Y(t) = (y'(t), y(t))$ et $F(t, (y_1, y_2)) = (3y_1 + 2y_2 + x(t), y_1)$. Remarquons que F se met sous la forme

$$F(t, Y) = (AY^t)^t + X(t)$$

où

$$A = \begin{pmatrix} 3 & 2 \\ 1 & 0 \end{pmatrix}, \ X(t) = (x(t), 0).$$

L'équation (E_f) est donc une ED linéaire d'ordre 2 à coefficients constants avec second membre.

De la même manière, toute ED est équivalente à une ED autonome si l'on ajoute une dimension:

$$y' = f(t, y, \dots, y^{(n-1)}) \Leftrightarrow Y' = F(Y)$$

avec $Y(t) = (y^{(n-1)}(t), \dots, y'(t), y(t))$ et $F(y_1, \dots, y_n) = (f(y_1, \dots, y_n), y_{n-1}, \dots, y_1)$.

5.2 Théorème de Cauchy-Lipschitz

On appelle *Problème de Cauchy* de paramètres $t_0 \in I, y_0, \dots, y_{n-1} \in E$ la résolution de E_f sous la restriction que

$$\begin{aligned} y(t_0) &= y_0 \\ y'(t_0) &= y_1 \\ &\dots \\ y^{(n-1)}(t_0) &= y_{n-1}. \end{aligned}$$

Si certaines hypothèses sont vérifiées, ce problème a une unique solution.

Théorème 12 (Cauchy-Lipschitz local)**.** *Soit $(t_0, y_0) \in I \times \Omega$ et $f : I \times \Omega \to E$ continue. On suppose que f est "localement lipschitzienne par rapport à son second argument", c'est-à-dire qu'il existe $L \in \mathbb{R}$, J voisinage de t_0 et U voisinage de y_0 tels que*

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

pour $t \in J, x, y \in U$. Alors

- *il existe $\tau > 0$ tel que le problème de Cauchy*

$$\begin{aligned} y' &= f(t,y) \\ y(t_0) &= y_0 \end{aligned}$$

  admette une solution $y : [t_0 - \tau, t_0 + \tau] \to U$.

- *Il y a unicité dans le sens où si il y a une autre solution z sur un voisinage de t_0, alors il existe $0 < \tau' \leqslant \tau$ tel que y et z coincident sur $[t_0 - \tau', t_0 + \tau']$.*

Si de plus f est de classe $\mathcal{C}^k$, alors y est de classe $\mathcal{C}^{k+1}$.

La conséquence de ce théorème et que deux solutions ne peuvent pas se "croiser", car si elles se croisent en t_0, $y_0 := y(t_0) = z(t_0)$ ce sont donc toutes les deux l'unique solution du problème de Cauchy $y(t_0) = y_0$.

Remarque 10. *Si f est de classe $\mathcal{C}^1$ sur $I \times \Omega$ et F est de dimension finie, alors f vérifie les hypothèses de CL en tout $(t_0, y_0) \in I \times \Omega$.*

Proof. Soit $J \subset I, K \subset \Omega$ des compacts tels que (t_0, y_0) soit dans l'intérieur de $J \times K$. Comme $\partial_2 f$ est continue sur le compact $J \times K$, elle y est bornée par une constante L, et par l'IAF on a pour $t \in J$

$$|f(t,x) - f(t,y)| \leqslant \sup_{z \in K} |\partial_2 f(t,z)||x - y| \leqslant L|x - y|.$$

□

Exemple 7. On reprend l'exemple précédent $y' = f(t,y)$ avec $f(t,y) = -y^2$. f est polynomiale, donc $\mathcal{C}^\infty$ en y, donc pour tout $(t_0, y_0) \in \mathbb{R} \times \mathbb{R}$ elle vérifie les hypothèses du TCL. Par exemple la solution y_a vérifie pour $a \in \mathbb{R}$

$$y_a(t_0) = \frac{1}{t_0 - a} = y_0$$

ssi $a = t_0 - \frac{1}{y_0}$, si $y_0 \neq 0$. Pour que t_0 soit dans l'intervalle de définition, il faut choisir $J_{a,+}$ si $y_0 > 0$ et $J_{a,-}$ si $y_0 < 0$.

On a donc la solution maximale :

- Si $y_0 > 0$, $y(t) = \frac{1}{t-a}$ sur $J_{a,+}$ avec $a = t_0 - \frac{1}{y_0}$
- Si $y_0 < 0$, $y(t) = \frac{1}{t-a}$ sur $J_{a,-}$ avec $a = t_0 - \frac{1}{y_0}$
- Si $y_0 = 0$, $y \equiv 0$ est la solution.

Ce sont les uniques solutions locales. Toute solution maximale y tombe nécessairement dans un de ces trois cas de figure.

Exemple 8. Soit l'ED $y' = 3|y|^{2/3}$. On remarque que $y(t) := t^3$ vérifie

$$y'(t) = 3t^2 = 3|t^3|^{2/3} = 3|y(t)|^{2/3}$$

est solution, ainsi que la solution $z \equiv 0$. Comme c'est une solution autonome, les fonctions

$$y_a(t) = (t-a)^3$$

sont également solution. On voit en particulier que le problème de Cauchy $y(0) = 0$ admet deux solutions, y et z. La raison pour laquelle le TCL ne s'applique pas est que $f(t, y) = |y|^{2/3}$ n'est pas Lipschitzienne en 0 car

$$|\frac{d}{dy} y^{2/3}| = |y^{-1/3}| \xrightarrow[t \to 0]{} \infty.$$

Par contre, pour $y_0 \neq 0$ et $t_0 \in \mathbb{R}$, on voit qu'on peut trouver a tel que y_a vérifie ces conditions, il faut que

$$(t-a)^3 = y_0.$$

Cette solution sera localement unique d'après le TCL car $(t, y) \mapsto |y|^{2/3}$ est $\mathcal{C}^\infty$ sur $\mathbb{R} \times \mathbb{R}^*$, mais pas globalement unique comme on le verra au TD4.

Proposition 18 (TCL pour une ED d'ordre quelconque)**.** *Soit $\Omega = \Omega_1 \times \cdots \times \Omega_n$ un produit d' ouverts de E, et $f : I \times \Omega \to E$. Alors si f est localement lipschitzienne par rapport à Y au voisinage d'un point $(t_0, Y_0 = (y_0, \ldots, y_{n-1}))$, le problème de Cauchy*

$$y^{(n)} = f(t, y(t), \ldots, y^{(n-1)}(t)), \quad y^{(i)}(t_0) = y_i, 0 \leqslant i \leqslant n-1$$

admet localement une unique solution dans le sens précédent.

Il suffit de transformer l'ED $y^{(n)} = f(t, y, \ldots, y^{(n-1)})$ en l'ED d'ordre 1 $Y' = F(t, Y)$ selon la méthode précédente et d'appliquer le TCL d'ordre 1.

Autrement dit, on a l'unicité de la solution pour une ED linéaire d'ordre n qui satisfait les hypothèses du TCL ssi on impose n conditions initiales, une pour chaque dérivée de y.

Théorème 13 (Unicité globale)**.** *Avec les notations précédentes: supposons que (I, y_1) et (I, y_2) soient deux solutions du problème de Cauchy $y(t_0) = y_0$, et que $f(t, y(t))$ soit localement lipschitzienne en $y(t)$ pour tout $t \in I$. Alors $y_1 = y_2$.*

Proof. L'ensemble

$$E = \{t \in I : y_1(t) = y_2(t)\} = (y_1 - y_2)^{-1}(\{0\})$$

est un fermé de I car $y_1 - y_2$ est continue. Par unicité locale des solutions, pour tout $t \in E, \exists \tau > 0$ tel que $]t-\tau, t+\tau[\subset E$, autrement dit E est un ouvert de I. Comme I est connexe, $E = I$.

Démo équivalente: Soit $a = \inf I; b = \inf A$ où $A = \{t \leqslant t_0 : y_1(s) = y_2(s), s \in [t, t_0]\}$. Par continuité de $y_1 - y_2, A$ est un fermé donc $b \in A$. Si $b > a, b \in I$, comme les hypothèses de CL sont vérifiées en b, il existe $\tau > 0$ tel qu'il y ait une solution sur $[b-\tau, b+\tau]$, et par unicité cette solution coïncide avec y_1 et y_2 sur $[b-\tau', b+\tau']$ avec $\tau' > 0$. Donc $b - \tau' \in A$, ce qui contredit $b = \inf A$. Donc $b = a$. De même, $\forall t > t_0, t \in I, y_1(t) = y_2(t)$.

□

Preuve du Théorème de Cauchy-Lipschitz. Remarquons que y est solution équivaut à

$$y(t) = \underbrace{y_0 + \int_{t_0}^{t} f(s, y(s))ds}_{g_y(t)}$$

pour t dans un voisinage ouvert de t_0 (l'intégrale est bien défini dans un espace de Banach). Donc on va chercher y comme un point fixe de l'application $\varphi : y \to g_y$, pour cela il faut se placer dans un espace où φ est contractante.

Soit $\tau > 0$ tel que $K = [t_0 - \tau, t_0 + \tau] \subset J$ et $\tau L < 1/2$. Soit F l'espace des fonctions continues $K \to E$ muni de la norme

$$\|f\| = \sup_{t \in K} \|f(t)\|.$$

Alors F est un espace de Banach (cf cours de Topologie). On voudrait définir $\varphi : F \to F$ qui à y associe g_y, comme défini au-dessus. On aurait pour $y, z \in F$

$$\|\varphi(y) - \varphi(z)\|_F = \sup_{t \in K} |\int_{t_0}^{t} [f(s, y(s)) - f(s, z(s))]ds| \leqslant \sup_{t \in K} \int_{t_0}^{t} L|y(s) - z(s)|ds \leqslant \sup_{t \in K} L\|y - z\||t_0 - t| \leqslant L\tau\|y - z\|.$$

Alors φ serait 1/2-contractante, et il existerait un unique point fixe y de φ, qui vérifie bien $y(t) = g_y(t)$, et donc en dérivant $y'(t) = f(t, y(t))$, avec $y(t_0) = y_0$.

Le problème est que $f(s, y(s))$ n'est bien définie que si y est à valeur dans Ω, et de toute façon f n'est lipschitzienne que dans un voisinage de (t_0, y_0). Ici on choisit une boule fermée $B(y_0, R) \subset \Omega$ avec $R > 0$ et $\tau > 0$ tels que f est bornée et L-lipschitzienne $/y$ sur $[t_0 - \tau, t_0 + \tau] \times B(y_0, R)$ et $\tau L < 1/2$ (au pire on diminue τ). On définit

$$A = \{y \in F : y(t_0) = y_0 \text{ et } \forall t \in [t_0 - \tau, t_0 + \tau], y(t) \in B(y_0, R)\}.$$

On a bien que A est un fermé: Si $y_n \to y$ pour la norme uniforme sur $[t_0 \pm \tau]$, $\forall t, y_n(t) \to y(t)$, et donc $y(t) \in B(y_0, R)$ (et $y(t_0) = \lim_n y_n(t_0) = y$). Il ne reste donc plus qu'à vérifier que A est stable par φ. On a

$$\|\varphi(y)(t) - y_0\| = \|g_y(t) - y_0\| \leqslant |t - t_0| \underbrace{\sup_{(t,y) \in [t_0 \pm \tau] \times B(0,R)} \|f(t, y)\|}_{M} \leqslant \tau M.$$

Quitte à diminuer encore τ, on suppose $\tau M \leqslant R$ (M n'est pas modifiée).

$M < \infty$ car on a supposé f bornée au moment de choisir R, τ.

Remarque pour plus tard: on a bien une solution sur $[t - \tau, t + \tau]$ dès lors que $M < \infty, \tau M \leqslant R, \tau L \leqslant 1/2$.

Si z est une autre solution définie sur $[t_0 - \delta, t_0 + \delta]$, alors en posant $\tau' = \min(\tau, \delta)$, y et z sont toutes deux solutions sur $[t_0 - \tau', t_0 + \tau']$. En refaisant le raisonnement ci-dessous avec le TPF, par unicité du point fixe, y et z coincident sur $[t_0 - \tau', t_0 + \tau']$.

□

Remarquons que le TPF nous fournit une manière de construire la solution: On part d'une fonction y_1 quelconque, et on définit $y_{n+1} = g_{y_n} = \varphi(y_n)$. Le TPF assure que y_n converge vers une solution dans F.

5.3 Solutions maximales

Etant donné une solution, le résultat suivant nous permet de dire si elle est maximale ou pas dans les conditions du théorème de CL: si une solution maximale n'est pas défnie sur tout l'intervalle I où f vérifie CL, alors la solution "explose" quelque part sur I.

Théorème 14. *[Théorème des bouts]On pose $I =]a, b[$ avec $a, b \in \mathbb{R} \cup \{-\infty, \infty\}$. Soit f continue et localement lipschitzienne par rapport à y sur $I \times \Omega$.*

- *Pour $t_0 \in I, y_0 \in \Omega$, il existe une unique solution maximale (J, y) au problème de Cauchy $y(t_0) = y_0, y' = f(t, y(t))$, avec $J =]\alpha, \beta[\subseteq I$.*
- *Soit $\beta = b$, soit: pour tout compact $K \subset \Omega, \exists s < \beta$ tel que pour $s \leqslant t \leqslant \beta$, $y(t) \in \Omega \setminus K$. On a le même comportement avec les bornes inférieures.*

Autrement dit, si y est bornée sur un intervalle strictement inclus dans I, alors elle ne peut être solution maximale. On dit aussi qu'"y n'explose pas en temps fini", et on peut la prolonger en une solution sur un plus grand intervalle.

Dans un espace de dimension infinie, c'est moins évident car on ne sait pas forcément quels sont les compacts.

Remarque 11. *L'équation $y' = |y|^{2/3}$ (cf. TD4) nous donne un contre-exemple: $f(t, y) = |y|^{2/3}$ n'est pas localement y-Lipschitzienne sur $\Omega = \mathbb{R}$ et on n'a en effet pas de solution maximale unique au problème de Cauchy $y(t_0) = y_0$, même si f est lipschitzienne au voisinage de (t_0, y_0). Par exemple les fonctions $y_1(t) = (t/3)^3$ et*

$$y_2(t) = \begin{cases} (t/3)^3 \ \textit{si}\ t > 0 \\ 0 \ \textit{sinon} \end{cases}$$

sont toutes deux solutions maximales du problème de Cauchy avec $y(3) = 1$. Il faut vérifier que y_2 est dérivable en 0: pour $t > 0$

$$\frac{y_2(t) - y_2(0)}{t} = \frac{t^2}{3^3} \to 0$$

et la limite est également 0 à gauche, donc $y_2'(0) = 0$ (on peut montrer que y_2 est de classe $\mathcal{C}^2$).

Proof. On va se concentrer sur la borne supérieure car la borne inférieure est traitée exactement de la même manière. On part d'une solution locale définie sur $[\alpha_0, \beta_0[\subset I$, qui existe grâce au théorème de CL local. Soit

$$\beta = \sup(E) \text{ où } E := \{t \geqslant \alpha_0 : y \text{ peut être étendue comme solution sur } [\alpha_0, t[.\}$$

Montrons qu'il y a une solution y définie sur $[\alpha_0, \beta[$. On considère des bornes $\beta_n \to \beta, \beta_n < \beta$, et des solutions y_n qui étendent y sur $]\alpha_0, \beta_n[$. Par la Proposition 13, y_{n+1} coïncide avec y_n sur $[\alpha_0, \beta_n[$ (et avec y sur $[\alpha_0, \beta_0[$). On définit alors pour $t \in [\alpha_0, \beta[$

$$y(t) = y_n(t) \text{ pour } n \text{ tel que } \beta_n > t \ (y(t) \text{ ne dépend pas de } n \text{ vu que les solutions coïncident}).$$

- **1er cas:** $\beta < b$ et y sort de tout compact, alors elle ne peut pas être prolongée par continuité en β, sinon on aurait $y([\alpha_0, \beta]) \subset K$ pour $K compact$. Dans ce cas y est bien maximale "à droite".

- **2ème cas:** y est contenue dans un compact K (i.e. $y([\alpha_0, \beta[) \subset K)$. Alors on va montrer par l'absurde que $\beta = b$ et que donc la solution est bien maximale "à droite".

  Supposons donc $\beta < b$ et montrons qu'on peut prolonger y sur un intervalle $[\beta, \beta + \tau]$, ce qui contredit la définition de β. Par compacité, il y a $z_0 \in K$ et une suite $t_n \to \beta$ tels que $y(t_n) \to z_0$. Soit $\tau > 0$ et $R > 0$ telle que f soit bornée par M et L-lipschitzienne sur $[\beta \pm 2\tau] \times B(y_0, 2R)$.

  On souhaite appliquer le TCL en un couple $(t_n, z_n := y(t_n))$, qui nous donne une solution sur un intervalle $[t_n \pm \tau_n]$. Par unicité la solution sera forcément y, donc si l'on trouve $t_n + \tau_n > \beta$, on pourra prolonger la solution au-delà de β, contradiction. Pour n suffisament grand, $z_n \in B(y_0, R)$ et donc $B(z_n, R) \subset B(y_0, 2R)$. Donc f est bien bornée par M et L-Lipschitzienne sur $[t_n \pm \tau_n] \times B(z_n, R)$. Si l'on choisit $\tau_n = \min(M/2R, 1/3L)$, on a une solution. Il suffit donc de pendre t_n pour que $t_n + \tau_n > \beta$.

$\square$

Remarque 12. *Si $E = \mathbb{R}^d$ ce résultat prend la forme suivante: Soit $\beta = b$, soit*

$$\limsup_{s \to \beta, s < \beta} \|y(t)\| = \infty, \tag{*}$$

i.e. il existe des temps $s_n \to \beta, s_n < \beta$ tels que $\|y(s_n)\| \to \infty$. Pour le montrer il faut prendre le compact $K_n = [-n, n]^d$ et choisir s_n tel que $\|y(t)\| \notin K_n$ pour $s_n \leqslant t \leqslant \beta$.

En dimension 1, $(*)$ équivaut à

$$\limsup_{s \to \beta} y(t) = +\infty \text{ or } \liminf_{s \to \beta} y(t) = -\infty.$$

Pour résumer, il peut se passer trois choses:

- Soit l'extrémité de la solution maximale est $\pm\infty$, par exemple
  - équation $y' = y$,
  - solution $y(x) = \exp(x)$, définie sur $\mathbb{R}$
- Soit l'équation "explose" à l'extrémité,exemple:
  - équation $y' = -y^2$
  - solution $y(x) = \frac{1}{x-a}$ sur $]a, +\infty[$
  - Explose à l'extrémité:

$$y(x) \xrightarrow[x \to a_+]{} +\infty$$

- Soit $f(y(x))$ n'est plus localement lipshitzienne à l'extrémité:
  - équation $y' = 3|y|^{2/3}$
  - Solution $y(x) = x^3$ sur $]0, \infty[$,
  - Pas d'unicité au voisinage de $y(x) = 0$ car $f(y) = y^{2/3}$ n'est pas lipschitzienne en 0.

5.4 Quelques techniques

Pour les équations linéaires, on va faire tout un chapitre dessus. Voyons d'autres techniques pour des équations non linéaires:

- **Equation séparée:**
$$\varphi(y(x))y'(x) = a(x)$$
avec a, φ continues. On fait une analyse-synthèse:
  - Analyse: Supposons que (J, y) soit une solution telle que $\varphi(y(x))$ ne s'annule pas et $a(x)$ ne s'annule pas pour $x \in J$. Alors si on connait une primitive ψ de φ on a pour $x_0, x \in J$,
$$(\psi(y(x)))' = a(x)$$
$$\psi(y(x)) - \psi(y(x_0)) = \int_{x_0}^{x} a(t)dt$$
ce qui donne une forme explicite si ψ est injective (sinon cela fait parfois plusieurs solutions, voir ci-dessous)
  - Synthèse: On sait désormais que si une solution existe, elle doit avoir la forme trouvée lors de l'analyse. On pose donc y sous cette forme, et on calcule $y'\varphi(y)$ pour voir si elle est effectivement solution. On pourra ensuite faire une résolution complète.
  - En pratique les notations sont plus légères, voyons le sur l'exemple
$$\begin{aligned} y'(1+y) =& x \\ \frac{dy}{dx}(1+y) =& x \\ dy(1+y) =& xdx \\ \int(1+y)dy =& \int xdx + C \\ y + \frac{y^2}{2} =& x^2/2 + C' \\ 2y + y^2 =& x^2 + C' \\ (1+y)^2 =& xx^2 + C' - 1 = x^2 + c \\ y_{c,\varepsilon}(x) =& \varepsilon\sqrt{x^2 + c} - 1. \end{aligned}$$
avec $\varepsilon \in \{-1, +1\}$. Remarquons que si $c < 0$ cette équation n'est bien définie que sur $J_c = [\sqrt{-c}, +\infty]$, autrement $J_c = \mathbb{R}$. **synthèse** On vérifie réciproquement facilement que pour chaque ε et constante c, la fonction $y_{c,\varepsilon}$ est solution sur J_c. D'après l'analyse ce sont les seules solutions possibles, donc on a toutes les solutions. Pour savoir si elles sont maximales il faut étudier la possibilité d'un prolongement en $\sqrt{-c}$ si $c < 0$ (cf. TD5, exo 3).

- Equations qui font intervenir une puissance de y, de la forme
$$y'(x) + \alpha(x)y(x) + \beta(x)y^p(x) = 0$$
C'est ce qu'on appelle une **équation de Bernoulli**. L'astuce consiste à poser $z(x) = y(x)^{1-p}$, cette fois en supposant que y ne s'annule pas. On a alors en divisant l'équation par y^p
$$y'y^{-p} + \alpha y^{1-p} + \beta = 0$$
$$\frac{z'}{1-p} + \alpha z = -\beta.$$
Il "suffit" donc de résoudre l'ED linéaire avec second membre ci-dessus, ce qui donne toutes les solutions u. On en déduit les solutions en y en posant $y = z^{\frac{1}{1-p}}$.

- **Equations de Riccati:** Ce sont les équations de Bernoulli avec $p = 2$ et avec second membre:

$$y' + \alpha y + \beta y^2 = \gamma.$$

Il faut trouver une solution particulière y_P (souvent de la forme $y_P(x) = x^k$), et ensuite poser $y = y_P + z$. On a

$$\begin{aligned} y \text{ solution } &\Rightarrow (y_P + z)' + \alpha(y_P + z) + \beta(y_P^2 + 2y_P + z^2) = 0 \\ &\Rightarrow z' + (\alpha + 2\beta y_P)z + \beta z^2 = 0 \text{ car } y_P' + \alpha y_P + \beta y_P^2 = \gamma. \end{aligned}$$

On a donc une équation de Bernoulli **homogène** pour z, qu'on résout en posant $u = z^{1-2} = 1/z$, avec $z' = -u'/u^2$, voir ci-dessus. On en déduit $y = y_P + z$ (cf. TD6 exo 8)

Chapter 6

Equations linéaires et sous-linéaires

6.1 Equations linéaires scalaires d'ordre 1

A connaître ou savoir retrouver.

Sans second membre

Soit I un intervalle de $\mathbb{R}, E = \mathbb{R}$. On considère l'équation linéaire d'ordre 1

$$y'(t) = a(t)y(t), t \in I.$$

où a est continue sur I.

Remarquons tout d'abord que $f(t, y) := a(t)y$ est localement lipschitzienne sur tout $J := [t - \tau, t + \tau] \times E$ car $|f(t, y) - f(t, x)| \leqslant \|a\|_J |x - y|$. On voit facilement que pout tout $t_0 \in I$ et constante $\lambda \in \mathbb{R}$,

$$y(t) = \lambda \exp(\int_{t_0}^{t} a(s)ds), t \in I,$$

est solution:

$$y'(t) = a(t)\lambda \exp(\int^{t} a) = a(t)y(t).$$

(On peut aussi retrouver ces solutions avec la méthode pour variables séparées). L'unique solution du problème de Cauchy $y(t_0) = y_0$ est donc

$$y(t) = y_0 \exp(\int_{t_0}^{t} a(s)ds).$$

En particulier, en dimension finie, cette solution est globale et définie sur $\mathbb{R}$ car pour $t \in \mathbb{R}$, $\int_{t_0}^{t} a$ n'explose pas car a est continue et donc bornée sur le compact $[t_0, t]$. On verra avec le lemme de Gronwall que pour ce type d'équations (linéaire) les solutions sont en général définies sur $\mathbb{R}$ tout entier

On voit que changer de point de départ revient à changer la constante: pour $t_0' \in \mathbb{R}$

$$\lambda \exp(\int_{t_0'}^{t} a(s)ds) = \lambda \exp(\int_{t_0'}^{t_0} a + \int_{t_0}^{t} a) = \lambda' \exp(\int_{t_0}^{t} a)$$

avec $\lambda' = \lambda \exp(\int_{t_0'}^{t_0} a(s)ds)$. Donc on a montré que l'espace vectoriel de dimension 1

$$E = \{\lambda y_1, \lambda \in \mathbb{R}\} = \mathbb{R}y_1$$

où

$$y_1(t) = \exp(\int_{t_0}^t a)$$

contient toutes les solutions, et elles sont donc globales en dimensions finie.

Avec second membre

Soit maintenant une équation avec second membre

$$y'(t) = a(t)y(t) + b(t) \tag{E}$$

où b est continue $I \to E$. Pour la résoudre, on cherche une solution particulière avec la méthode dite de “variation de la constante”: soit $\lambda : I \to \mathbb{R}$ de classe $\mathcal{C}^1$ et

$$y(t) = \lambda(t)\exp(\int_{t_0}^t a(s)ds) = \lambda(t)y_1(t).$$

Alors on a $y' = \lambda' \exp(...) + \lambda a \exp(...) = ay + \lambda' \exp(...)$. Donc y est solution si $\lambda' \exp(..) = b$, c'est-à-dire si pour une constante $y_0 \in \mathbb{R}$,

$$\lambda(t) = y_0 + \int_{t_0}^t b(s)\exp(-\int_{t_0}^s a(u)du)ds.$$

On a alors

$$y(t) = y_0 y_1(t) + y_p(t)$$

où

$$y_p(t) = y_1(t)\int^t b(s)\exp(-\int^s a)ds \tag{(SP)}$$

On remarque que $y_P(t_0) = 0$. Comme a et b sont continues et donc bornées sur tout compact, y n'explose pas en temps fini et elle est donc globale. On en déduit la proposition suivante.

Proposition 19. *Les solutions de l'équation E sur un intervalle $I \subset \mathbb{R}$ forment l'espace affine de solutions globales (définies sur I):*

$$\mathcal{E} = \{y_0 y_1 + y_p, \lambda \in \mathbb{R}\} = E + y_p = \mathbb{R}y_1 + y_p$$

où y_1, y_p sont définies ci-dessus, et $t_0 \in I$ est un point de référence fixé.

Etant donné le problème de Cauchy (t_0, y_0); l'unique solution est définie sur $\mathbb{R}$ par

$$y(t) = y_0 y_1(t) + y_p(t)$$

car $y(t_0) = y_0$.

6.2 Equations sous-linéaires

Le prochain résultat permet d'estimer si effectivement une solution peut exploser en un temps fini.

Remarque: y est à valeurs dans $\mathbb{R}$, pas dans E, voir après la preuve pour la méthode pour traiter des fonctions vectorielles.

Théorème 15 (Lemme de Gronwall)**.** *Soit $y(t), t \in I$ une fonction $I \to \mathbb{R}$ telle qu'il existe des fonctions continues a et c sur I telles que $a(t) \geqslant 0$ et pour $t, t_0 \in I$*

$$y(t) \leqslant c(t) + \int_{t_0}^{t} a(s)y(s)ds.$$

Alors

$$y(t) \leqslant c(t) + \int_{t_0}^{t} c(s)a(s)\exp(\int^{s} a)ds.$$

Proof. On pose

$$z(t) = \int_{t_0}^{t} a(s)y(s)ds.$$

On a $z'(t) \leqslant a(t)(z(t)+c(t)) = az+b$ si l'on pose $b = ac$. Soit u la solution de l'ED linéaire $u' = au+b$ telle que $u(t_0) = 0$. La section précédente nous donne explicitement

$$u(t) = \exp(\int_{t_0}^{t} a)\int_{t_0}^{t} \exp(-\int_{t_0}^{s} a)b(s)ds,$$

On a donc

$$\begin{aligned} u' &= au+b \\ z' &\leqslant au+b \\ z(t_0) &= u(t_0). \end{aligned}$$

on se doute que "de proche en proche" $z \leqslant u$ sur tout I, montrons-le: On a

$$\begin{aligned} (z\exp(-\int^{t} a))' &= (z'-az)\exp(-\int_{t_0}^{t} a) \leqslant ac\exp(-\int_{t_0}^{t} a) \text{ car } \exp(-\int a) \geqslant 0 \\ \text{on intègre: } z\exp(-\int_{t_0}^{t} a) &\leqslant \int_{t_0}^{t} a(s)c(s)\exp(-\int_{t_0}^{s} a)ds = u\exp(-\int_{t_0}^{t} a). \end{aligned}$$

Alors $z \leqslant u$. et donc

$$y \leqslant c+z \leqslant c+u$$

qui est l'inégalité voulue. □

On peut appliquer ce résultat à des ED en dimension > 1: on a parfois un contrôle du type

$$\|Y(t)\| \leqslant b(t) + \int a(s)\|Y(s)\|ds.$$

C'est le cas par exemple pour une équation du type $Y'(t) = f(t, Y(t)) + b(t)$ où f est linéaire $/Y$.

Théorème 16. *Soit E un espace de Banach, $\Omega \subset E$. Soit une ED de la forme $Y' = f(t, Y)$ pour $f : I \times \Omega \to E$ telle qu'il existe une fonction continue $a : I \to \mathbb{R}_+$ satisfaisant*

$$\forall t, y, z \in I \times \Omega \times \Omega, \|f(t, y) - f(t, z)\| \leqslant a(t)\|y - z\|.$$

Alors pour tout problème de Cauchy $Y(t_0) = Y_0$, il existe une unique solution globale $Y(t), t \in I$.

Une telle équation est dite “sous-linéaire”.

Proof. On voit déjà que les hypothèses du TCL sont vérifiées car

$$\|f(t, y) - f(t, z)\| \leqslant a(t)\|y - z\|,$$

et comme a est continue, elle est bornée au voisinage de t_0, donc f est localement lipschitzienne au voisinage de (t_0, y_0). Soit donc Y une solution maximale (il en existe au moins une par le théorème des bouts **??**).

En intégrant on a

$$\begin{aligned} Y(t) =& Y(t_0) + \int_{t_0}^t f(s, Y(s))ds \\ Y(t) - Y(t_0) =& \int_{t_0}^t (f(s, Y(s)) - f(s, Y(t_0)))\, ds + \underbrace{\int_{t_0}^t f(s, Y(t_0))ds}_{c(t)} \\ \|Y(t) - Y(t_0)\| \leqslant& \int_{t_0}^t a(t)\|Y(s) - Y(t_0)\|ds + c(t). \end{aligned}$$

Donc en appliquant le lemme de Gronwall à $y(t) := \|Y(t) - Y(t_0)\|$, on a

$$\|Y(t) - Y(t_0)\| \leqslant \underbrace{c(t) + \int_{t_0}^t c(s)a(s) \exp(\int^s a)dsdt}_{\varphi(t)}.$$

Comme c et a sont continues, φ est continue. Comme $[t_0, t]$ est un compact, $K_t := \varphi([t_0, t])$ est un compact de E. On en déduit que $Y(t) - Y(t_0)$ ne peut pas sortir du compact K_t en un temps fini t. Donc d'après le théorème des bouts, cela implique que Y est globale. □

6.3 Application exponentielle

Sur $\mathbb{R}$, l'ED $y' = ay$ se résout par $y(t) = \exp(a(t - t_0))y_0$. Comment résoudre l'équation linéaire du 1er ordre à coefficients constants

$$Y' = AY$$

en dimension quelconque (finie), quand A est une matrice? Comme en dimension 1, la solution est de la forme

$$Y(t) = \exp((t - t_0)A)Y_0$$

où $\exp(\cdot)$ est l'exponentielle de matrices et Y_0 est un vecteur. Définissions l'exponentielle de matrice. On va le faire par analogie avec la formule

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}, x \in \mathbb{R}.$$

Théorème 17. *On se place dans un espace de Banach E. L'application*

$$\ell \in \mathcal{L}(E) \to \exp(\ell) \in \mathcal{L}(E)$$

est bien définie via

$$\exp(\ell) = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{l^k}{k!}$$

où ℓ^k est $\ell \circ \ell \cdots \circ \ell$.

Proof. Il faut donc montrer que la suite

$$\sum_{k=0}^{n} \frac{f^k}{k!}$$

converge dans $\mathcal{L}(E)$ quand $n \to \infty$. Comme E est complet, il suffit de montrer que cette suite est de Cauchy. On a pour $p, n \in \mathbb{N}$,

$$\| \sum_{k=n}^{n+p} \frac{\ell^k}{k!} \| \leqslant \sum_{k=n}^{n+p} \frac{\|\ell^k\|}{k!}$$

On rappelle que la norme triple est sous-miltiplicative: pour $\ell, \ell' \in \mathcal{L}(E)$,

$$|||\ell \circ \ell'||| = \sup_{\|h\|=1} \|\ell \circ \ell'(h)\| \leqslant \sup_{\|h\|=1} |||\ell||| \|\ell'(h)\| \leqslant |||\ell||| |||\ell'|||.$$

On dit que $||| \cdot |||$ est une norme matricielle.

Par récurrence, $|||\ell^k||| \leqslant |||\ell|||^k$. Et donc

$$\| \sum_{k=n}^{n+p} \frac{\ell^k}{k!} \| \leqslant \sum_{k=n}^{n+p} \frac{\|\ell^k\|}{k!} \leqslant \sum_{k=n}^{n+p} \frac{|||\ell|||^k}{k!}.$$

Cette suite est donc bien de Cauchy car on a la convergence

$$\lim_{n\to\infty} \sum_{k=0}^{n} \frac{|||\ell|||}{k!} \to \exp(|||\ell|||)$$

En remplaçant ℓ par A, la preuve est le même pour une matrice. Il faut simplement adopter une norme d'algèbre sur l'espace des matrices, par exemple

$$\|A\| = \sup_{\|h\|=1} \|Ah\|.$$

□

Proposition 20. *Pour $\ell, \ell' \in \mathcal{L}(E,E)$ qui commutent, c'est-à-dire $\ell \circ \ell' = \ell' \circ \ell$, on a*

$$\exp(\ell + \ell') = \exp(\ell)\exp(\ell').$$

En particulier, pour $t, s \in \mathbb{R}$, $\exp(t\ell)$ et $\exp(s\ell)$ commutent et pour $t \in \mathbb{R}$, $\exp(t\ell)$ est inversible d'inverse $\exp(-t\ell)$.

Proof. Les deux séries $\exp(\ell) = \sum_{k=0}^{\infty} \frac{(\ell)^k}{k!}$ et $\exp(\ell') = \sum_{m=0}^{\infty} \frac{(\ell')^m}{m!}$ convergent absolument, donc on peut intervertir les sommes et sommer dans l'ordre que l'on veut:

$$\begin{aligned}
\exp(\ell) \circ \exp(\ell') &= \sum_{k=0}^{\infty} \sum_{m=0}^{\infty} \frac{(\ell)^k (\ell')^m}{k!m!} \\
&= \sum_{n=0}^{\infty} \sum_{l=0}^{n} \frac{(\ell)^l (\ell')^{n-k}}{l!(n-k)!} \text{ (changement de variable } n = k+m, l = k) \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{l=0}^{n} (\ell)^l (\ell')^{n-l} \frac{n!}{l!(n-l)!} \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} (\ell + \ell')^n \text{ en utilisant la formule du binome} \\
&= \exp(\ell + \ell').
\end{aligned}$$

On a donc

$$\exp(\ell) \circ \exp(-\ell) = \exp(-\ell) \circ \exp(\ell) = I$$

donc $\exp(\ell)$ et $\exp(-\ell)$ sont bijectives linéaires continues réciproques.

□

Proposition 21. *Soit I un intervalle, et $\ell \in \mathcal{L}(E,E)$.*
Alors la dérivée de $t \mapsto \exp(t\ell)$ dans $\mathcal{L}(E,E)$ est $\ell \circ \exp(t\ell)$.

Proof. L'idée est assez simple: Pour $t \in \mathbb{R}$

$$\frac{d}{dt} \sum_{k=0}^{\infty} \frac{(t\ell)^k}{k!} = \sum_{k=1}^{\infty} \frac{kt^{k-1}\ell^k}{k!} = \ell \circ \left(\sum_{k=1}^{\infty} \frac{(t\ell)^{k-1}}{(k-1)!} \right) = \ell \exp(t\ell).$$

Pour le prouver rigoureusement: stoit $k \in \mathbb{N}^*, s \in \mathbb{R}$. On a

$$\frac{((t+s)\ell)^k - (t\ell)^k}{s} = \frac{(skt^{k-1} + s^2 k \gamma_t)\ell^k}{s}$$

où $|\gamma_t| \leqslant |t| + |t|^{k-2}$. en ne gardant que les termes d'ordre 1 en s quand on développe. Donc

$$\| \frac{\exp((t+s)\ell) - \exp(t\ell)}{s} - \sum_k kt^{k-1}\ell^k \| \leqslant s \sum_k k(|t| + |t|^{k-2}) \|\ell^k\| \xrightarrow[s \to 0]{} 0$$

Donc

$$\frac{d}{dt}\sum_{k=0}^{\infty}\frac{(t\ell)^k}{k!}=\sum_{k=1}^{\infty}\frac{kt^{k-1}\ell}{k!}=\lim_{n\to\infty}\ell\circ\sum_{k=1}^{n}\frac{(t\ell)^{k-1}}{(k-1)!}=\ell\exp(t\ell).$$

On a utilisé la continuité de ℓ:

$$\ell(\lim_n\sum_{k=0}^{n}\frac{(t\ell)^k}{k!})=\lim_n\ell(\sum_{k=0}^{n}\frac{(t\ell)^k}{k!})$$

□

On peut résoudre les équations linéaires à coefficients constants:

6.4 Equations linéaires à coefficients constants en dimension quelconque

Rappel: Sur un espace de Banach E, l'intégrale d'une fonction $f:\mathbb{R}\to E$ de norme intégrable est bien définie, et pour $\ell\in\mathcal{L}(E,E)$,

$$\ell(\int_a^b f(t)dt)=\int_a^b\ell(f(t))dt.$$

On ne fait pas la démonstration générale ici, mais en dimension finie c'est facile: si $E=\mathbb{R}^n, f(t)=(f_1(t),\dots,f_n(t))$, avec $e=(e_1,\dots,e_n)$ la base canonique,

$$\int_a^b f(t)dt=\left(\int_a^b f_i(t)dt\right)_{1\leqslant i\leqslant n}=\sum_{i=1}^n\left(\int_a^b f_i(t)dt\right)e_i,$$

et

$$\ell\left(\int_a^b f(t)dt\right)=\ell\left(\sum_{i=1}^n\left(\int_a^b f_i(t)\right)e_i\right)=\sum_{i=1}^n\int_a^b f_i(t)dt\ell(e_i)=\int_a^b\sum_{i=1}^n f_i(t)\ell(e_i)=\int_a^b\ell(f(t))dt.$$

La forme des solutions d'une ED linéaire à coefficients constants s'écrit de manière très similaire au cas d'une équation d'ordre 1 à coefficients constants (il suffit de remplacer $\ell(y)$ par une multiplication $a\times y$ pour passer de la dimension quelconque à la dimension 1):

Remarque: on l'utilisera de manière théorique pour les preuves, mais en pratique on ne fera que de la dimension finie.

Théorème 18. *Soit E un espace de Banach, $I\subset\mathbb{R}$ un intervalle, et $\ell\in\mathcal{L}(E,E)$ et $t\mapsto B(t)\in E$ continue et l'ED*

$$Y'(t)=\ell(Y(t))+B(t).$$

Alors les solutions homogènes globales sont les fonctions de la forme

$$Y_{y_0}(t)=\exp((t-t_0)\ell)(y_0),t\in\mathbb{R},y_0\in E,$$

et la solution particulière qui s'annule en t_0, obtenue par variation de la constante, est

$$Y_p(t)=\int_{t_0}^t\exp((t-s)\ell)(B(s))ds,t\in I.$$

La solution au problème de Cauchy $Y(t_0)=y_0$ est donc

$$Y(t)=Y_{y_0}(t)+Y_p(t)$$

Proof. La méthode et les calculs sont sensiblement les mêmes que pour la dimension 1.

Sans second membre, on a pour $y_0 \in E$

$$Y_{y_0}(t) = \exp((t-t_0)\ell)(y_0)$$

qui vérifie bien $Y'_{y_0}(t) = \ell(Y(t)), Y_{y_0}(t_0) = y_0$, et cette solution est globale. Par unicité de la solution au problème de Cauchy, toutes les solutions sont de cette forme là: soit une solution Y et $y_0 = Y(t_0)$. Alors par unicité $Y = Y_{y_0}$ sur l'ensemble de définition de Y.

Dans la suite on écrit $R(t) := \exp((t-t_0)\ell) \in \mathcal{L}(E,E)$ pour simplifier les notations. Avec second mmbre, dans la méthode de la variation de la constante, il faut juste remplacer la multiplication par une constante $\lambda \in \mathbb{R}$ par lapplication de l'argument $\lambda \in E$ à la fonction linéaire $R(t)$. La méthode de variation de la constante devient alors

$$Y_p(t) = \underbrace{\exp((t-t_0)\ell)}_{R(t)\in\mathcal{L}(E,E)}(\lambda(t))$$

Pour dériver, on peut l'écrire comme une fonction de deux arguments $Y_p(t) = R(t)(\lambda(t)) = \varphi(t, \lambda(t))$. On a donc pour $s \in \mathbb{R}$, en utilisant le fait que à t fixé, $R(t)$ est linéaire continue,

$$\begin{aligned} dY_{p,t}(s) =& d_1\varphi_{(t,\lambda(t))}(s) + d_2\varphi_{(t,\lambda(t))}(s) = sR'(t)(\lambda(t)) + R(t)(d\lambda(t)(s)) \\ Y'_p(t) =& R'(t)(\lambda(t)) + R(t)(\lambda'(t)) = \ell(R(t)(\lambda(t))) + R(t)(\lambda'(t)) = \ell(Y_p(t)) + R(t)(\lambda'(t)). \end{aligned}$$

λ doit donc vérifier

$$\begin{aligned} R(t)(\lambda'(t)) = B(t) \\ \lambda'(t) = R(t)^{-1}(B(t)) = \exp(-(t-t_0)\ell)(B(t)) \end{aligned}$$

ce qui donne la solution en intégrant:

$$\begin{aligned} Y_P(t) &= \exp((t-t_0)\ell)(\lambda(t)) \\ &= \exp((t-t_0)\ell) \circ \int_{t_0}^t \exp(-(s-t_0)\ell)(B(s))ds \\ &= \int_{t_0}^t \exp((t-t_0)) \circ \exp(-(s-t_0)\ell)(B(s))ds \\ &= \int_{t_0}^t \exp((t-s)\ell)(B(s))ds \end{aligned}$$

On a bien $Y_{y_0}(t_0) + Y_P(t_0) = y_{y_0}(t_0) = y_0$, donc par unicité du problème de Cauchy on a toutes les solutions. □

En pratique il peut être compliqué de calculer l'application exponentielle, et à un ordre pas trop grand on fait les calculs à la main.

6.5 Vision "algèbre linéaire"

Soit $\mathcal{E}_0$ l'espace des solutions homogènes. L'application

$$\begin{aligned} S : E &\to \mathcal{E}_0 \\ y_0 &\mapsto Y_{y_0} = R_{t_0}(\cdot)(y_0). \end{aligned}$$

est linéaire car $R_{t_0}(t) \in \mathcal{L}(E, E)$. Donc $\mathcal{E}_0 = S(E)$ est un espace vectoriel. Montrons que S est une bijection. Il suffit de montrer que S est injective. Soit $y_0 \in E$ telle que Y_{y_0} est la solution nulle. Alors $y_0 = Y_{y_0}(t_0) = 0$, donc S est injective.

Cela implique que l'espace des solutions est de même dimension que E.

Dimension finie Soit $y_1, \dots, y_n$ une base de E. On note $Y_i = S(y_i) = Y_i$ la solution qui vaut y_i en t_0. Comme S est injective, (Y_i) forme une base de $\mathcal{E}_0$. Toute solution homogène $Y = Y_{y_0}$ est de la forme

$$Y = \sum_{i=1}^{n} \alpha_i Y_i.$$

On peut trouver les coefficients α_i comme ceux de la décomposition de y_0 dans la base (y_i):

$$y_0 = \sum_{i=1}^{n} \alpha_i y_i$$

dans la base y_i, alors Y_{y_0} se décompose $\sum_{\alpha_i} Y_i$, car

$$Y = S(y_0) = R_{t_0}(\cdot)(\sum_i \alpha_i y_i) = \sum_i \alpha_i Y_i$$

Variation de la constante Il s'agit de faire varier y_0 dans E, ou de manière équivalente de faire varier les coefficients α_i dans la base $(y_1, \dots, y_n)$. Si l'on dispose déjà d'une base $(Y_i)_{1 \leqslant i \leqslant n}$ de solutions, une solution particulière s'obtient donc via

$$Y_P(t) = \sum_{i=1}^{n} \alpha_i(t) Y_i(t),$$

en trouvant les bonnes valeurs pour $\alpha_i(t)$ en résolvant $Y_P'(t) = \ell(Y_P(t)) + b(t)$.

$$\begin{aligned} Y_P' &= \sum_i \alpha_i' Y_i + \sum_i \alpha_i Y_i' \\ &= \sum_i \alpha_i' Y_i + \sum_i \alpha_i \ell(Y_i) \\ &= \sum_i \alpha_i' Y_i + \ell(\sum_i \alpha_i Y_i) \\ &= \sum_i \alpha_i' Y_i + \ell(y_P). \end{aligned}$$

Donc Y_P est solution équivaut à $Y_P' = \ell(Y_P) + B$ et donc

$$\begin{aligned} \sum_i \alpha_i' Y_i &= B \\ R_{t_0}(t)(\sum_i \alpha_i' y_i) &= B(t), t \in I, \\ \alpha'(t) = R_{t_0}(t)^{-1}(B(t))\alpha(t) &= \int R_{t_0}(s)^{-1} B(s) ds \end{aligned}$$

en notant $\alpha'(t)$ le vecteur de coordonnées les $\alpha_i'(t)$. On retrouve l'expression de la solution particulière si $R_{t_0}(t) = \exp((t - t_0)\ell)$:

$$Y_P(t) = R_{t_0}(t) \left(\int R_{t_0}(s)^{-1}(B(s)) ds \right) = \exp((t - t_0)\ell) \left(\int \exp(-(s - s_0))(B(s)) ds \right) = \int_{t_0}^{t} \exp((t - s)\ell)(B(s)) ds.$$

Equation scalaire d'ordre n

On rappelle que toute équation scalaire (E) d'ordre n peut se mettre sous une forme (équivalente) d'une ED (E') d'ordre 1 dans un espace de dimension n. Soit $\mathcal{E}_0$ l'espace vectoriel des solutions homogènes de (E) et $\mathcal{E}_0'$ l'espace vectoriel des solutions homogènes de (E'). Le passage d'une solution $y \in \mathcal{E}_0$ à une solution $Y \in \mathcal{E}_0'$ se fait via l'application

$$\begin{aligned}\psi : \mathcal{E}_0 &\to \mathcal{E}_0' \\ y &\mapsto Y : (x \mapsto y(x), y'(x), \ldots, y^{(n-1)}(x)).\end{aligned}$$

En particulier, cette application est linéaire injective, et donc bijective car $\mathcal{E}_0$ et $\mathcal{E}_0'$ ont la même dimension. Si (y_1, y_2) est une base de $\mathcal{E}_0$, alors $(Y_1 = \psi(y_1), Y_2 = \psi(y_2))$ est une base de $\mathcal{E}_0'$. Réciproquement, d'une base (Y_1, Y_2) on peut obtenir une base $(y_1 = \psi^{-1}(Y_1), y_2 = \psi^{-1}(Y_2))$ de $\mathcal{E}_0$.

En particulier, si (E) est solution avec second membre, toute solution de (E) se met sous la forme

$$y = \sum_{i=1}^{n} \alpha_i f_i + y_p$$

, où les α_i sont des scalaires, et y_P une solution particulière.

Baisse de l'ordre d'une ED linéaire par la variation de la constante: Si on a une solution homogène y d'une ED linéaire de la forme

$$\sum_{i=0}^{n} \alpha_i(t) y^{(i)}(t) = s(t)$$

alors si l'on pose $z(t) = y(t)u(t)$, z est solution si u vérifie

$$u(t) \sum_{i=0}^{n} \alpha_i y^{(i)}(t) + \sum_{i=1}^{n} \beta_i(t) u(t) = s(t)$$

où les β_i dépendent des α_i et des $y^{(i)}$, donc u' est solution de l'ED d'ordre $n-1$: $\sum_{i=0}^{n-1} (u')^{(i)} \beta_{i+1} = s(t)$.

Exemple d'utilisation de la variation de la constante pour une ED scalaire d'ordre 2 Soit l'ED linéaire avec second membre et coefficients constants:

$$y''(t) - y(t) = e^t$$

C'est équivalent à l'équation du 1er ordre

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

où

$$\begin{aligned}\ell(Y) =& AY, A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ B(t) =& \begin{pmatrix} e^t \\ 0 \end{pmatrix}, \\ Y(t) =& \begin{pmatrix} y'(t) \\ y(t) \end{pmatrix}\end{aligned}$$

On remarque que e^t et e^{-t} sont solution homogènes.

On a donc des solutions homogènes de la forme $y(t) = \alpha e^t + \beta e^{-t}$, où $\alpha, \beta \in \mathbb{R}$, ou en termes de Y:

$$Y(t) = \alpha \begin{pmatrix} e^t \\ e^t \end{pmatrix} + \beta \begin{pmatrix} -e^{-t} \\ e^{-t} \end{pmatrix} = M_t \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \text{ où } M_t = \begin{pmatrix} e^t & -e^{-t} \\ e^t & e^{-t} \end{pmatrix}$$

. **Méthode 2** Le théorème précédent nous dit que les solutions homogènes de l'ED en dimension 2 sont de la forme

$$Y(t) = \exp(tA) \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

pour $\alpha, \beta \in \mathbb{R}$. Prenons par exemple $t_0 = 0, y_0 = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$. L'unique solution homogène telle que $Y(t_0) = y_0$ est $\exp((t - t_0))(y_0)$, c'est-à-dire $\exp(tA) \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$. Il faut donc calculer $\exp(tA)$.

Exercice 8. Montrer par le calcul que pour $t \in \mathbb{R}$,

$$\exp(tA) = \frac{1}{2} \begin{pmatrix} e^t + e^{-t} & e^t - e^{-t} \\ e^t - e^{-t} & e^t + e^{-t} \end{pmatrix}$$

Pour $t = 1$. Comme $A^2 = I$, on a

$$\begin{aligned} A^{2k} &= I, k \in \mathbb{N}, \\ A^{2k+1} &= A^{2k} A = A, k \in \mathbb{N}. \end{aligned}$$

Donc

$$\begin{aligned} \exp(A) &= \sum_{k=0}^{\infty} \frac{A^k}{k!} = \sum_{k \text{ pair}} \frac{I}{k!} + \sum_{k \text{ impair}} \frac{A}{k!} \\ &= I \sum_{k \text{ pair}} \frac{1}{k!} + A \sum_{k \text{ pair}} \frac{1}{k!} \\ &= \begin{pmatrix} \frac{e^1 + e^{-1}}{2} & \frac{e^1 - e^{-1}}{2} \\ \frac{e^1 - e^{-1}}{2} & \frac{e^1 + e^{-1}}{2} \end{pmatrix} \end{aligned}$$

Pour t quelconque, on peut faire les mêmes calculs en partant de $A^{2k} = t^{2k} I, A^{2k+1} = t^{2k+1} A$.

Solution particulière La méthode de la variation de la constante est donc

$$y_P(t) = \alpha(t) e^t + \beta(t) e^{-t}$$

Vu la forme du second membre, on cherche sous la forme $0 = \beta$, c'est-à-dire

$$y_P(t) = \alpha(t) e^t$$

Donc

$$\begin{aligned} y_P' &= (\alpha' + \alpha) e^t \\ y_P'' &= (\alpha'' + 2\alpha' + \alpha) e^t \\ y_P'' - y_P &= (\alpha'' + 2\alpha') e^t \end{aligned}$$

Donc y_P est solution si $\alpha'' + 2\alpha' = 1$. On choisit par exemple $\alpha(t) = t/2$, et on a la solution

$$y_P(t) = \frac{t}{2} e^t.$$

On en déduit la forme générale des solutions:

$$y(t) = \alpha e^t + \beta e^{-t} + \frac{t e^t}{2}$$

Méthode 2 On peut donc trouver une SP avec l'exponentielle via

$$Y_P(t) = \int_0^t \exp((t-s)A)(B(s))ds$$

et ensuite y_P est la deuxième composante de cette égalité:

$$y_P(t) = \int_0^t \frac{e^{t-s} - e^{-(t-s)}}{2} e^s ds = \frac{1}{2}(te^t - e^{-t}\int_0^t e^{2s}ds) = \frac{te^t}{2} - e^{-t}[\frac{1}{2}e^{2s}]_0^t = \frac{te^t}{2} - \frac{1}{2}(e^t - e^{-t}).$$

Cette méthode nous donne une seconde solution particulière, mais ça ne contredit pas le premier calcul, car il n'y a pas unicité de la solution particulière. On observe bien que cette seconde solution est de la forme

$$\alpha e^t + \beta e^{-t} + \frac{te^t}{2}$$

comme on l'avait trouvé précédemment.

6.6 Coefficients non constants

Lorsqu'on a une équation du type

$$Y' = \ell_t(Y_t) + B(t)$$

où $\forall t \in I, \ell_t \in \mathcal{L}(E,E)$ on ne peut pas donner la solution explicite en général. Il y a quand même des choses qu'on peut dire.

Définition 20. *Pour $t_0, t \in \mathbb{R}$, on note $R_{t_0}(t)(y_0)$ la valeur au temps t de l'unique solution maximale qui vaut y_0 au temps t_0 (elle est donnée par le problème de Cauchy et elle est globale).*

Par exemple dans le cas où $\ell_t = \ell$,

$$R_{t_0}(t)(y_0) = \exp((t-t_0)\ell)(y_0)$$

La vision algèbre linéaire marche toujours: si $y_0 = \sum_i \alpha_i y_i$, la solution homogène qui vaut y_0 en t_0 est

$$Y_{y_0}(t) = \sum_i \alpha_i R_{t_0}(t)(y_i).$$

L'application $y_0 \mapsto Y_{y_0}$ est linéaire injective car $Y = 0$ implique $y_0 = Y(t_0) = 0$. Si on trouve une base. de solutions y_i, on a alors

$$Y = \sum_i \alpha_i Y_i.$$

Proposition 22. *Soit $t \mapsto \ell_t$ continue. Supposons que pour $t, s \in \mathbb{R}, \ell_t$ et ℓ_s commutent: $\ell_t \circ \ell_s = \ell_s \circ \ell_t$. On peut alors résoudre l'ED exactement comme dans le cas à coefficients constants, en remplaçant $\exp((t-t_0)\ell)$ par $\exp(\int_{t_0}^t \ell_s ds)$:*

$$R_{t_0}(t)(y_0) = \exp(\int_{t_0}^t \ell_s ds)(y_0),$$

les solutions homogènes sont les

$$Y_{y_0}(t) = R_{t_0}(t)(y_0).$$

La solution particulière s'écrit d'une manière analogue au cas de la dimension 1*:*

$$Y_P(t) = \exp(\int_{t_0}^{t} \ell_s ds) \circ \int_{t_0}^{t} \exp(-\int_{t_0}^{s} \ell_s)(B(s))ds$$

et toutes les solutions sont les

$$Y_{y_0}(t) + Y_P(t)$$

pour y_0 *qui décrit* E*.*

Proof. sous forme d'exercice. On ne fait que la dimension finie.

1. On suppose que $t \mapsto \ell_t$ est différentiable. Montrez que pour $t \in \mathbb{R}, h \in E, \ell_t'(\ell_t(h)) = \ell_t(\ell_t'(h))$.
   Soit $h \in E$
   $$\begin{aligned} \ell_t'(\ell_t)(h) &= \lim_{s\to 0} \left(\frac{(\ell_{t+s}(\ell_t(h)) - \ell_t(\ell_t(h)))}{s} \right) \\ &= \lim_{s\to 0} \left(\frac{(\ell_t(\ell_{t+s}(h)) - \ell_t(\ell_t(h)))}{s} \right) \text{ par commutativité} \\ &= \ell_t \left(\lim_{s\to 0} \frac{\ell_{t+s}(h) - \ell_t(h)}{s} \right) \text{ par continuité de } \ell_t \\ &= \ell_t(\ell_t'(h)). \end{aligned}$$

2. Montrer que
   $$t \mapsto \exp(\ell_t)$$
   est différentiable, et a pour dérivée $\ell_t' \circ \exp(\ell_t)$. On pourra s'inspirer de la preuve de la proposition 21.
   Soit $s \in \mathbb{R}$. On a $\ell_{t+s} = \ell_t + s\ell_t' + s\varepsilon_t(s)$ où $\varepsilon_t(s) \to 0$ quand $s \to 0$.
   $$\begin{aligned} \lim_n \sum_{k=0}^{n} \frac{\ell_{t+s}^k}{k!} &= \lim_n \sum_k \frac{(\ell_t + s\ell_t' + s\varepsilon_t(s))^k}{k!} \\ &= \lim_n \sum_k \frac{(\ell_t + ks(\ell_t' + \varepsilon_t(s))^{k-1} + s^2(...))}{} \end{aligned}$$

3. Déduisez-en la dérivée de
   $$A_t := \exp(\int_{t_0}^{t} \ell_s ds).$$
   Etablissons débord que $A_t \circ A_s = A_s \circ A_t$: il faut utiliser Fubini dans un Banach (ok pour dimension finie)
   $$A_t \circ A_s = \int_{t_0}^{t} \ell_u du \circ \int_{t_0}^{s} \ell_v dv = \int_{[t_0,t]\times[t_0,s]} \ell_u \circ \ell_v dudv = \int_{[t_0,s]\times[t_0,t]} \ell_v \circ \ell_u dvdu = A_s \circ A_t.$$

4. Calculez la dérivée de $t \mapsto A_t$.

   On a

$$\frac{\int_{t_0}^{t+s} \ell_u du - \int_{t_0}^{t} \ell_u du}{s} - \ell_t = \frac{\int_t^{t+s} (\ell_u - \ell_t) du}{s}$$

   Comme ℓ_t est continue, elle est uniformément continue, et $\|\ell_u - \ell_t\| \leqslant \varepsilon$ pour s suffisamment petit.

5. Montrez que pour $y_0 \in E$,

$$Y_{y_0}(t) = \exp(\int_{t_0}^{t} \ell_s ds)(y_0)$$

   est la solution au problème de Cauchy

$$\begin{cases} Y'(t) = \ell_t(Y(t)) \\ Y(t_0) = Y_0. \end{cases}$$

   Découle directement des questions précédentes.

6. On suppose que E est de dimension finie, et que $y_1, \ldots, y_n$ est une base de E. Déduisez-en une base de l'espace des solutions homogènes.

   L'application $S : y_0 \mapsto Y_{y_0}$ est linéaire injective, donc $\mathcal{E}_0 = S(E)$ et (Y_{y_i}) forme une base.

   Variation de la constante

7. Déduisez-en l'espace des solutions.

$\square$