Calcul différentiel

Partie 3 sur 3 du cours de topologie et calcul différentiel.

Patrick Bernard

December 16, 2013

Nous exposons ici le calcul différentiels dans le cadre des espaces de Banach. Presque tous les résultats sont déjà intéressants et utiles en dimension finie, et il ne faut pas hésiter à lire ce cours en supposant que tous les espaces sont de dimension finie.

1 Différentiabilité

Soient E et F des espaces de Banach. L'application $f : E \longrightarrow F$ est dite *différentiable* (ou Fréchet-différentiable) en x_0 si il existe une forme linéaire continue $L \in \mathcal{L}(E, F)$ telle que

$$f(x_0 + x) = f(x_0) + L \cdot x + \|x\|\epsilon(x)$$

où $\epsilon : E \longrightarrow F$ est une fonction qui tend vers 0 en 0. On dit alors que L est la différentielle de f en x_0, on notera par la suite $df(x_0)$ cette différentielle. On rappelle que la courbe $\gamma : \mathbb{R} \longrightarrow F$ est dite dérivable en t_0 si la limite

$$\gamma'(t_0) = \lim_{t \longrightarrow 0} \frac{\gamma(t_0 + t) - \gamma(t_0)}{t}$$

existe. On dit alors que $\gamma'(t_0)$ est la dérivée de γ en t_0. La courbe γ est différentiable si et seulement si elle est dérivable, et sa différentielle en t_0 est $d\gamma(t_0) \cdot t = t\gamma'(t_0)$. On remarque que $d\gamma(t_0)$ est, par définition, un élément de $\mathcal{L}(\mathbb{R}, F)$ alors que $\gamma'(t_0)$ est un élément de F.

Exercice 1. *L'application $l \longmapsto l(1)$ est un isomorphisme isométrique entre les espaces de Banach $\mathcal{L}(\mathbb{R}, F)$ et F.*

Si f est différentiable en x_0, alors chacune des courbes $t \longmapsto \gamma_v(t) = f(x_0 + tv), v \in E$ est dérivable en 0. On dit que f est dérivable dans la direction v en x_0. On note

$$\delta f(x_0, v) := \gamma_v'(0) = df(x_0) \cdot v$$

la dérivée directionnelle de f. L'existence de dérivées directionnelles de f dans toutes les directions en x_0 n'implique pas la différentiabilité de f, sauf si E est de dimension 1.

Exemple 2. *Considérons la fonction $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ telle que*

- $f(x, y) = 1$ *si* $y = x^2$ *et* $x \neq 0$*,*
- $f = 0$ *si* $x = 0$ *ou* $y \neq x^2$*.*

On vérifie facilement que toutes les dérivées directionnelles existent et que $\delta f(0, v) = 0$ pour tout v.

La fonction de l'exemple n'est pas continue en 0, donc elle n'est pas différentiable en 0, au vu de la propriété suivante, qui découle immédiatement de la définition.

Propriété 3. *Si f est différentiable en x_0, alors elle est continue en x_0.*

Si $\delta f(x,v)$ existe, alors $\delta f(x,tv)$ existe pour tout $t \in \mathbb{R}^*$ et $\delta f(x,tv) = t\delta f(x,v)$. Cependant, même si $\delta f(x,v), \delta f(x,w)$ et $\delta f(x,v+w)$, on a pas forcément $\delta f(x,v+w) = \delta f(x,w) + \delta f(x,v)$.

Definition 4. *On dit que la fonction f est Gâteau-différentiable en x_0 si toutes les dérivées directionnelles de f en x_0 existent, et si l'application $v \longmapsto \delta f(x_0,v)$ est linéaire et continue.*

Toute fonction différentiable au sens de Frechet est différentiable au sens de Gâteau, avec $\delta f(x_0,v) = df(x_0)\cdot v$. L'exemple montre que la différentiabilité au sens de Gâteau n'implique pas la différentiabilité au sens de Fréchet, même en dimension finie. Elle n'implique même pas la continuité de f.

Pour mieux détailler la différence entre différentiabilité au sens de Gâteau est de Fréchet, on peut considérer la fonction

$$\Delta_v(t) = \frac{f(x+tv) - f(x) - tL(v)}{t\|v\|}.$$

L'application linéaire $L \in \mathcal{L}(E,F)$ est la différentielle au sens de Gâteau de f en x si et seulement si chacune des fonctions Δ_v tend vers 0 en 0. L'application linéaire $L \in \mathcal{L}(E,F)$ est la différentielle au sens de Frechet de f en x si et seulement les fonctions $\Delta_v, v \in S_E$ tendent uniformément vers 0 en 0 (où S_E est la sphère unité de E).

Propriété 5.

- *La courbe $\gamma(t) : \mathbb{R} \longrightarrow F$ est dérivable en t_0 si et seulement si elle est Frechet différentiable en t_0, ce qui est aussi équivalent à la différentiabilité au sens de Gâteau en t_0. On a alors la relation $d\gamma(t_0)\cdot s = s\gamma'(t_0)$.*
- *Toute application linéaire continue $L \in \mathcal{L}(E,F)$ est différentiable au sens de Fréchet en chaque point, avec $dL(x_0) = L$.*

La propriété suivant est valable aussi bien au sens de Fréchet qu'au sens de Gâteau:

Proposition 6. *Si g est différentiable en x_0 et si f est différentiable en $g(x_0)$, alors $f\circ g$ est différentable en x_0 et $d(f\circ g)(x_0) = df(g(x_0))\circ dg(x_0)$.*

DÉMONSTRATION.

$$\begin{aligned}
f\circ g(x_0+x) &= f\Big(g(x_0) + dg(x_0)\cdot x + \|x\|\epsilon(x)\Big)\\
&= f(g(x_0)) + df(g(x_0))\cdot\Big(dg(x_0)\cdot x + \|x\|\epsilon(x)\Big)\\
&+ \Big\|dg(x_0)\cdot x + \|x\|\epsilon(x)\Big\|\epsilon\Big(dg(x_0)\cdot x + \|x\|\epsilon(x)\Big)\\
&= f\circ g(x_0) + df(g(x_0))\circ dg(x_0)\cdot x + r(x)
\end{aligned}$$

avec un reste $r(x) = \|x\|\epsilon(x)$ (on note par $\epsilon(x)$ des fonctions toutes différentes qui tendent vers zéro quand x tend vers zéro). □

Par exemple si $\gamma : \mathbb{R} \longrightarrow E$ est une courbe dérivable et si $f : E \longrightarrow F$ est une application différentiable, alors la courbe $f\circ\gamma$ est différentiable, et

$$(f\circ\gamma)'(t_0) = df(\gamma(t_0))\cdot\gamma'(t_0).$$

Lorsque $E = E_1\times E_2\times\cdots\times E_n$ est un produit, on note $\partial_i f(x) = \partial_i f(x_1,x_2,\ldots x_n)$ la différentielle partielle par rapport à la i-ème variable, lorsqu'lle existe, qui est la différentielle en zéro de l'application

$$E_i \ni y \longmapsto f(x_1,x_2,\ldots,x_i+y,\ldots,x_n).$$

C'est donc un élément de $\mathcal{L}(E_i, F)$. Si f est différentiable en $x = (x_1, \dots x_n)$, alors ses différentielles partielles existent et on a

$$df(x) \cdot y = df(x) \cdot (y_1, \dots, y_n) = \partial_1 f(x) \cdot y_1 + \cdots + \partial_n f(x) \cdot y_n.$$

Dans le cas, particulièrement important, où $E = \mathbb{R}^n$ on parlera plus souvent de dérivée partielle que de différentielles partielles, et on notera

$$\partial_i f(x) \quad \text{ou} \quad \partial_{x_i} f(x) \quad \text{ou} \quad \frac{\partial f}{\partial x_i}(x)$$

la dérivée directionnelle de f dans la direction e_i (où $(e_1, \dots, e_n)$ est la base canonique de $\mathbb{R}^n$). C'est donc la dérivée en zéro de la courbe $t \longmapsto f(x + te_i)$. Le fait d'utiliser la même notation $\partial_i f(x)$ pour la différentielle partielle (qui est une application linéaire de $\mathbb{R}$ dans F) et pour la dérivée (qui est un vecteur de f) ne crée pas trop de confusion en général. Par exemple, la formule

$$df(x) \cdot (y_1, \dots, y_n) = \partial_1 f(x) \cdot y_1 + \cdots + \partial_n f(x) \cdot y_n.$$

s'écrie de la même façon dans les deux cas, mais avec une interprétation différente de l'opération représentée par $\cdot$. Dans un cas, c'est l'opération d'évaluation d'une application linéaire ($L \cdot v = L(v)$), et dans l'autre, c'est le produit par un scalaire $f \cdot t = tf$.

Un cas particulièrement important est celui ou E et F sont de dimension finie, $E = \mathbb{R}^n$ et $F = \mathbb{R}^m$. Dans ce cas, la différentielle $df(x) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, lorsqu'elle existe, se représente par une matrice réelle $m \times n$ que l'on appelle *matrice Jacobienne* de f en x. On la note parfois $df(x)$ ou $Jf(x)$ si on veut être plus précis.

Propriété 7. *Les coéfficients de la matrice Jacobienne $Jf(x)$ d'une fonction $f = (f_1, \dots, f_n)$ différentiable en x sont*

$$[Jf(x)]_{i,j} = \frac{\partial f_j}{\partial x_i}.$$

Démonstration. Pour tout $l \in F'$, l'application $t \longmapsto l \circ f(a + t(b-a))$ est C^1 □

Rappelons encore une fois que l'existence de toutes les dérivées partielles $\partial_i f^j(x)$ n'implique pas la différentiabilité de f au sens de Frechet en x.

Démonstration. Notons $\pi_j = \mathbb{R}^m \longrightarrow \mathbb{R}$ la projection sur le j-ème facteur, et $\gamma_i(t)$ la courbe $x + te_i, 1 \leqslant i \leqslant n$. On a alors $f_j(x + te_i) = \pi_j \circ f \circ \gamma_i$. En appliquant la formule de composition et en rappelent que $d\pi_j(x) = \pi_j$ pour tout x, on obtient:

$$\partial_i f_j(x) = (\pi_j \circ f \circ \gamma_i)'(0) = \pi_j \circ df(x) \circ e_i = [Jf(x)]_{i,j}.$$

□

Introduisons une dernière notation classique: Si H est un espace de Hilbert, et si $f : H \longrightarrow \mathbb{R}$ est différentiable en x_0, alors sa différentielle $df(x_0)$ est un point du dual H'. En identifiant canoniquement H et son dual H', on peut toutefois considérer cette différentielle comme un élément de H, que l'on appelle *gradient* de f en x_0, et que l'on note $\nabla f(x_0)$. On a, par définition

$$df(x_0) \cdot v = \langle \nabla f(x_0), v \rangle$$

pour tout $v \in H$. Dans le cas où H est l'espace Euclidien $\mathbb{R}^n$, la différentielle de f est la forme linéaire

$$y \longmapsto df(x) \cdot y = \partial_1 f(x) y_1 + \cdots + \partial_n f(x) y_n,$$

et le gradient de f s'écrit

$$\nabla f(x) = (\partial_1 f(x), \dots, \partial_n f(x)).$$

Théorème 1 (Inégalité des accroissements finis)**.** *Soit U un ouvert de E et $f : U \longrightarrow F$ une application. Soit $[a,b] \subset U$ un segment tel que f est Gâteau différentiable en chaque point de $[a,b]$, alors*

$$\|f(b)-f(a)\| \leqslant \big(\sup_{x\in[a,b]} \|Gf(x)\|\big)\|b-a\|.$$

Démonstration. On peut supposer que $E=\mathbb{R}$ et que $b>a$ sinon on remplace la fonction f par la fonction $\mathbb{R} \ni t \longmapsto f(a+t(b-a))$. On fixe une constante $C > \sup_{x\in]a,b[} \|Gf(x)\|$. La courbe f est dérivable sur $[a,b]$, donc continue. L'ensemble $A \subset [a,b]$ des points x tels que $\|f(x)-f(a)\| \leqslant C\|x-a\|$ est donc fermé. Soit $M \in [a,b]$ le suprémum des réels $x \in [a,b]$ tels que $[a,x] \subset A$. On remarque que $a \in A$, donc $M \in [a,b]$. Comme A est fermé, on a $[a,M] \subset A$. Comme la fonction f est dérivable en M, et que $\|f'(M)\| < C$, il existe $\epsilon > 0$ tel que $\|f(x)-f(M)\| \leqslant C\|x-M\|$ pour tout $x \in [M, M+\epsilon]$. Mais alors

$$\|f(x)-f(a)\| \leqslant \|f(x)-f(M)\| + \|f(M)-f(a)\| \leqslant C(x-M)+C(M-a) = C(x-a)$$

pour tout $x \in [M, M+\epsilon]$. Ceci est en contradiction avec la définition de M, sauf si $M=b$. On a montré que $\|f(b)-f(a)\| \leqslant C(b-a)$ pour tout $C > \sup_{x\in[a,b]} \|Gf(x)\|$, et donc que $\|f(b)-f(a)\| \leqslant \big(\sup_{x\in[a,b]} \|Gf(x)\|\big)(b-a)$. □

Dans la preuve ci-dessus, on aurait pu se ramener au cas $F=\mathbb{R}$ en considérant la fonction $g = l\circ f$, où l est une forme linéaire $l \in F'$ telle que $l(f(b)-f(a)) = \|b-a\|$ et $\|l\| = 1$. En appliquant l'inégalité des accroissements finis à l'application $f(x)-L$, où $L \in \mathcal{L}(E,F)$, on obtient:

Corollaire 8. *Sous les hypothèses de l'inégalité des accroissements finis, pour tout $L \in \mathcal{L}(E,F)$ on a*

$$\|f(b)-f(a)-L\cdot(b-a)\| \leqslant \Big(\sup_{x\in[a,b]} \|df(x)-L\|\Big)\|b-a\|.$$

En particulier,

$$\|f(b)-f(a)-df(a)\cdot(b-a)\| \leqslant \Big(\sup_{x\in[a,b]} \|df(x)-df(a)\|\Big)\|b-a\|.$$

Ce corollaire permet d'obtenir des estimations du reste du développement limité d'ordre 1 de f en a en fonction d'informations sur sa différentielle. Les expressions suivantes impliquent formellement l'inégalité des accroissements finis, mais nécessitent une hypothèse un peu plus forte.

Propriété 9. *Si f est Gâteau différentiable en chaque point de l'intervalle $[a,b] \subset E$, et si l'application $t \longmapsto df(a+t(b-a)$ est continue (de $[0,1]$ dans $\mathcal{L}(E,F)$, alors*

$$f(b)-f(a) = \int_0^1 df(a+t(b-a))\cdot(b-a)dt = \int_0^1 df(a+t(b-a))dt\cdot(b-a)$$

Démonstration. Pour tout $l \in F'$, on considère la fonction réelle $g : t \longmapsto l\circ f(a+t(b-a))$. C'est une fonction C^1, donc

$$\begin{aligned} l(f(b)-f(a)) = g(1)-g(0) = \int_0^1 g'(t)dt &= \int_0^1 [l\circ df(a+t(b-a))]\cdot(b-a)dt \\ &= l\cdot \int_0^1 df(a+t(b-a))\cdot(b-a)dt. \end{aligned}$$

Comme ceci est vrai pour tout $l \in F'$, on a la première égalité. La second égalité consiste juste à faire sortir de l'intégrale l'application linéaire continue de $\mathcal{L}(E,F)$ dans F consistant à évaluer en $(b-a)$. □

2 Intégration des fonctions à valeurs vectorielles

Soit X un espace mesurable, μ une mesure finie sur X, et B un espace de Banach. Étant donnée une application $f : X \longrightarrow B$, on dit que

$$I = \int f(x)d\mu(x)$$

si I est un point de B tel que $l(I) = \int l \circ f(x)d\mu(x)$ pour tout $l \in B'$. On demande donc en particulier que $l \circ f$ soit intégrable pour tout $l \in B'$. On dit que f est intégrable si elle admet une intégrale I, cette intégrale est alors unique, et on a

$$\|I\| \leqslant \int \|f(x)\|d\mu(x).$$

En effet, on peut prendre un forme linéaire $l \in B'$ telle que $l(I) = \|I\|$ et $\|l\| = 1$, on a alors $l(I) = \int l \circ f \leqslant \int \|f\|$. De plus, pour $L \in \mathcal{L}(B, F)$, l'application $L \circ f$ est intégrable si f l'est, et

$$\int L \circ f d\mu = L \cdot \int f d\mu.$$

En effet, pour tout $l \in F'$, on a $\int l \circ (L \circ f) = \int (l \circ L) \circ f = (l \circ L) \cdot \int f = l\Big(L \cdot \int f\Big)$.

Propriété 10. *Les applications intégrables constituent un espace vectoriel, et l'intégrale est une application linéaire.*

Démonstration. Si f et g sont intégrables, alors pour tout $l \in B'$ on a

$$l\left(a \int f d\mu + b \int g d\mu\right) = a \int l \circ f d\mu + b \int l \circ g d\mu = \int l \circ (af + bg) d\mu.$$

□

La difficulté consiste à caractériser les fonctions intégrable. Nous nous contenterons ici d'un énoncé très partiel :

Propriété 11. *Si f est mesurable et d'image précompacte, alors elle est intégrable pour toute mesure μ finie.*

En particulier, si X est un espace topologique compact muni de sa tribu Borélienne, et si f est continue, alors elle est intégrable pour toute mesure finie μ sur X. Bien sur, il suffit de supposer qu'il existe une partie $Y \subset X$ de mesure totale et telle que $f(Y)$ est précompacte.

Démonstration. Soit F l'image de f. En utilisant la précompacité, on construit, pour tout $\epsilon > 0$, une partition mesurable finie A_i de F dont les éléments ont un diamètre inférieur à ϵ. On choisit alors un point a_i de A_i, et on considère la fonction g qui prend la valeur a_i sur $f^{-1}(A_i)$. Il est facile de vérifier que la fonction g est intégrable et que $\int g d\mu = \sum_i a_i \mu(f^{-1}(A_i))$. On fait cette construction pour chaque $\epsilon_n = 2^{-n}$ et on appelle g_n les applications ainsi construites. On voit que $g_n \longrightarrow f$ dans l'espace $b(X, B)$ des fonctions bornées de X dans B (muni de la norme uniforme). La suite $I_n := \int g_n d\mu$ est alors de Cauchy dans B, puisque

$$\|I_m - I_n\| \leqslant \int \|g_m - g_n\| d\mu \leqslant \mu(X) \|g_m - g_n\|.$$

Elle a donc une limite I. Pour tout $l \in B'$, on a

$$\int l \circ g_n d\mu = l(I_n) \longrightarrow l(I)$$

et $\int l \circ g_n d\mu \longrightarrow \int l \circ g d\mu$ puisque $l \circ g_n$ converge uniformément vers $l \circ g$. On en conclut que $\int l \circ f d\mu = l(I)$ pour tout $l \in B'$, et donc que I est l'intégrale de f. □

3 Applications C^1

Definition 12. *Soit U un ouvert de E. L'application $f : U \longrightarrow F$ est dite C^1 si elle est différentiable en chaque point de U et si la différentielle dépend continument du point, c'est à dire si l'application $df : U \longrightarrow \mathcal{L}(E, F)$ est continue.*

Il n'est pas utile de spécifier si l'on parle de différentielle au sens de Fréchet ou au sens de Gâteau au vu du Théorème ci-dessous.

Théorème 2. *Toute application C^1 au sens de Gâteau sur un ouvert U de E est aussi C^1 au sens de Fréchet sur U.*

Démonstration. Notons $G(x)$ la différentielle de Gâteau de f en x. En utilisant le Corollaire 8, on a

$$\|f(x+y) - f(x) - G(x) \cdot y\| \leqslant \|y\| \sup_{z \in [x,x+y]} \|G(z) - g(x)\|.$$

Comme $z \longmapsto G(z)$ est continue en x, le terme $\sup_{z \in [x,x+y]} \|G(z) - G(x)\|$ tend vers zéro lorsque y tend vers 0. On conclue que $G(x)$ est une différentielle au sens de Fréchet. □

Propriété 13. *Soit U un ouvert de E et $f : U \longrightarrow F$ une application C^1. Si $[a, b] \subset U$, alors*

$$f(b) - f(a) = \int_0^1 df(a + t(b-a)) \cdot (b-a) dt = \int_0^1 df(a + t(b-a)) dt \cdot (b-a)$$

Dans le cas où $E = E_1 \times \cdots \times E_n$ est un produit fini, on a

Théorème 3. *Soit U un ouvert de $E = E_1 \times \cdots \times E_n$. La fonction f est C^1 sur U si et seulement si ses différentielles partielles $\partial_i f, 1 \leqslant i \leqslant n$ existent et sont C^1.*

En particulier, si E est de dimension finie, l'existence de dérivées partielles n'implique pas la différentiabilité, mais l'existence de dérivées partielle continues implique la différentiabilité!

Démonstration.

On fait la preuve dans le cas $n = 2$, le cas général s'en déduit par récurrence sur n. Au vu du théorème précédent, il suffit de démontrer que f est Gâteau différentiable avec

$$G(x) \cdot (y_1, y_2) = \partial_1 f(x) \cdot y_1 + \partial_2 f(x) \cdot y_2.$$

On a en effet

$$\begin{aligned}
&f(x_1 + ty_1, x_2 + ty_2) - f(x) \\
=&f(x_1 + ty_1, x_2 + ty_2) - f(x_1 + ty_1, x_2) + f(x_1 + ty_1, x_2) - f(x_1, x_2) \\
=&\left(\int_0^1 \partial_2 f(x_1 + ty_1, x_2 + sy_2) ds\right) \cdot (ty_2) + \left(\int_0^1 \partial_1 f(x_1 + sy_1, x_2) ds\right) \cdot (ty_1) \\
=&(\partial_2 f(x) + \epsilon(t)) \cdot (ty_2) + (\partial_1 f(x) + \epsilon(t)) \cdot ty_1 \\
=&\Big(\partial_1 f(x) \cdot y_1 + \partial_2 f(x) \cdot y_2\Big) t + t\epsilon(t).
\end{aligned}$$

□

4 Différentielle seconde

On vient de s'intéresser à la continuité de l'application $df : E \longrightarrow \mathcal{L}(E,F)$ qui au point x associe la différentielle $df(x)$. On va maintenant étudier la différentiabilité de cette application.

Definition 14. *L'application $f : E \longrightarrow F$ est dite deux fois différentiable en x_0 si il existe un voisinage U de x_0 dans E tel que f est différentiable en chaque point de x, et si l'application $df : E \longrightarrow \mathcal{L}(E,F)$ est différentiable en x_0. La différentielle seconde $d^2f(x_0) := d(df)(x_0)$ est alors un élément de $\mathcal{L}(E,\mathcal{L}(E,F))$.*

Pour décrire l'espace $\mathcal{L}(E,\mathcal{L}(E,F))$ qui intervient naturellement dans la définition ci-dessus, faisons un petit détour algébrique.

Lemme 15. *Soient E_1, E_2 et F des espaces vectoriels normés. Pour une application bilinéaire $b : E_1 \times E_2 \longrightarrow F$, les propriétés ci-dessous sont équivalentes:*

- *b est continue en $(0,0)$.*
- *Il existe une constant $C > 0$ telle que $\|b(x_1,x_2)\| \leqslant C\|x_1\|\|x_2\|$.*
- *b est différentiable et*

$$db(x_1,x_2)\cdot(y_1,y_2) = b(x_1,y_2) + b(y_1,x_2).$$

- *b est contiue.*

Démonstration. Si b est continue en $(0,0)$, alors il existe $\epsilon > 0$ tel que $\|b(x_1,x_2)\| \leqslant 1$ pour $\|x_1\| \leqslant \epsilon$ et $\|x_2\| \leqslant \epsilon$. En utilisant la bilinéarité, on conclut que $\|b(x_1,x_2)\| \leqslant \epsilon^{-2}\|x_1\|\|x_2\|$.

Si $\|b(x_1,x_2)\| \leqslant C\|x_1\|\|x_2\|$, on développe

$$b(x_1+y_1, x_2+y_2) = b(x_1,x_2) + b(x_1,y_2) + b(y_1,x_2) + b(y_1,y_2).$$

Comme $\|b(y_1,y_2)\| \leqslant C\|y_1\|\|y_2\| \leqslant C\|y\|^2 = \|y\|\epsilon(y)$, on en conclut que b admet un développement limité au premier ordre, et est donc différentiable.

La différentiabilité implique la continuité, qui implique la continuité en 0. □

On note $\mathcal{L}^2(E_1,E_2;F)$ l'ensemble des applications bilinéaires continues de $E_1 \times E_2$ dans F. On le muni de la norme

$$\|b\| := \sup_{x_1 \in B_{E_1}, x_2 \in B_{E_2}} \|b(x_1,x_2)\|.$$

Propriété 16. *L'espace $\mathcal{L}^2(E_1,E_2;F)$ est un espace de Banach (si F l'est) pour la norme*

$$\|b\| := \sup_{x_1 \in B_1, x_2 \in B_2} \|b(x_1,x_2)\|.$$

Il est canoniquement isométrique à $\mathcal{L}(E_1,\mathcal{L}(E_2,F))$ par l'ismomorphisme J tel que

$$J(L)(x_1,x_2) = L(x_1)(x_2).$$

Démonstration. On vérifie immédiatement que J est un isomorphisme isométrique. On en conclut que $\mathcal{L}^2(E_1, E_2; F)$ est un Banach. $\square$

Propriété 17. *La partie $\mathcal{L}^2_s(E; F) \subset \mathcal{L}^2(E, F; F)$ constituée des application bilinéaires symétriques est un sous-espace fermé (donc un Banach).*

Ayant identifié $\mathcal{L}(E, \mathcal{L}(E, F))$ à $\mathcal{L}^2(E, F; F)$, on considèrera la différentielle seconde $d^2 f$ comme une forme bilinéaire.

Proposition 18. *Si f est deux fois différentiable en x_0, on a la formule*

$$d^2 f(x_0) \cdot (y, z) = dg(x_0) \cdot y,$$

avec $g(x) := df(x) \cdot z$.

Démonstration. Considérons l'application linéaire continue $Z : \mathcal{L}(E, F) \longrightarrow F$ donnée par $Z(L) = L(z)$. L'application $g : x \longmapsto Z \circ df(x)$ est différentiable en x_0, et

$$dg(x_0) \cdot y = Z \circ d^2 f(x) \cdot y = d^2 f(x) \cdot (y, z).$$

$\square$

Théorème 4 (Schwartz)**.** *Si f est deux fois différentiable en x_0, alors*

$$d^2 f(x_0) \in \mathcal{L}^2_s(E, F).$$

Démonstration. Supposons pour simplifier que $x_0 = 0$, et posons $B(x, y) := d^2 f(0) \cdot (x, y)$. On pose

$$\Delta(x, y) = f(x + y) - f(x) - f(y) + f(0) - B(x, y).$$

On va montrer que $\Delta(x, y) = (\|x\|^2 + \|y\|^2)\epsilon(x, y)$. Par symétrie, on en conclut que $B(x, y) - B(y, x) = (\|x\|^2 + \|y\|^2)\epsilon(x, y)$ et donc que la forme quadratique $B(x, y) - B(y, x)$ est nulle.

L'application Δ est différentiable au voisinage de 0, et elle vérifie $\Delta(x, 0) = 0$ pour tout x. Par l'inégalité des accroissements finis, on a donc

$$\|\Delta(x, y)\| \leqslant \|y\| \sup_{z \in [0, y]} \|\partial_2 \Delta(x, z)\|,$$

avec

$$\partial_2 \Delta(x, z) = df(x + z) - df(z) - B[x],$$

où l'on note $B[x]$ l'application linéaire $B(x, .)$, qui n'est autre que $d^2 f(0) \cdot x$. On a donc

$$\begin{aligned}\partial_2 \Delta(x, z) &= df(x + z) - df(0) - B[x + z] - \big(df(z) - df(0) - B[z]\big) \\ &= \|x + z\|\epsilon(x + z) + \|z\|\epsilon(z) = (\|x\| + z\|)\epsilon(\|x\| + \|z\|)\end{aligned}$$

puisque $B[x] = d^2 f(0) \cdot x$. Comme $\|z\| \leqslant \|y\|$, on en conclut que

$$\Delta(x, y) = \|y\|(\|x\| + \|y\|)\epsilon(\|x\| + \|y\|) = (\|x\|^2 + \|y\|^2)\epsilon(x, y).$$

$\square$

Dans le cas où $E = E_1 \times \cdots \times E_n$ est un produit, on définit les différentielles partielles secondes

$$\partial^2_{ij} f = \partial_i(\partial_j f) : U \longrightarrow \mathcal{L}(E_i, \mathcal{L}(E_j, F)) \approx \mathcal{L}^2(E_i, E_j; F).$$

Si f est deux fois différentiable en x_0, alors toutes ses différentielles partielles secondes existent et

$$d^2 f(x_0) \cdot (y, z) = \sum_{i,j} \partial^2_{ij} f(x_0) \cdot (y_i, z_j).$$

En particulier, dans le cas d'une application $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, la différentielle seconde s'identifie à une forme bilinéaire symétrique sur $\mathbb{R}^n$, dont la matrice $Hf(x_0)$ (dit matrice Hessienne de f) vérifie

$$[Hf(x_0)]_{i,j} = \partial^2_{ij} f(x_0).$$

Proposition 19. *Soit $f : E \longrightarrow \mathbb{R}$ une fonction deux fois différentiable en x_0 qui admet un minimum local en ce point. Alors $df(x_0) = 0$ et $d^2 f(x_0)$ est positive.*

Démonstration. Si $df(x_0)$ n'est pas nulle, alors il existe y tel que $df(x_0) \cdot y \neq 0$. Mais alors la fonction réelle $t \longmapsto f(x_0 + ty)$ n'a pas de minimum en $t = 0$, ce qui est une contradiction. On a donc $d^2 f(x_0) = 0$.

Si $d^2 f(x_0)$ n'est pas positive, alors il existe y tel que $d^2 f(x_0) \cdot (y, y) < 0$. Le développement limité à l'ordre 2 de la fonction $t \longmapsto f(x_0 + ty)$ en 0 étant

$$f(x_0 + ty) = f(x_0) + (t^2/2) d^2 f(x_0) \cdot (y, y) + t^2 \epsilon(t)$$

on conclut que cette fonction n'a pas de minimum local en $t = 0$, ce qui est une contradiction. □

5 Différentielles d'ordre supérieur

On note $\mathcal{L}^k(E, F)$ l'espace des applications k-linéaires continues de E^k dan F. On le munit de la norme

$$\|F\| := \sup_{x_i \in B_{E_i}} \|F(x_1, x_2, \ldots, x_k)\|.$$

L'espace $\mathcal{L}(E, \mathcal{L}^{k-1}(E, F))$ s'identifie à $\mathcal{L}^k(E, F)$. On dit que f est k fois différentiable en x_0 si elle est $k - 1$ fois différentiable au voisinage de x_0 et si l'application

$$E \ni x \longmapsto d^{k-1} f(x) \in \mathcal{L}^k(E, F)$$

est différentiable en x_0. On pose alors $d^k f(x_0) = d(d^{k-1} f)(x_0)$, c'est un élément de $\mathcal{L}(E, \mathcal{L}^{k-1}(E, F))$, que l'on identifie à un élément de $\mathcal{L}^k(E, F)$.

On note $\mathcal{L}^k_s(E, F)$ le sous-espace des applications multilinéaires symétriques, c'est à dire telles que

$$F(x_1, \ldots x_n) = F(x_{\sigma(1)}, \ldots, x_{\sigma(n)})$$

pour tout permutation σ.

Théorème 5. *Si f est k fois différentiable en x_0, alors sa différentielle $d^k f(x_0)$ est symétrique.*

Démonstration. On le démontre par récurrence, le cas $k = 2$ ayant déjà été démontré. Comme $d^{k-1} f(x)$ est symétrique pour tout x, on déduit de l'expression

$$d^k f(x_0) \cdot v = d\Big(d^{k-1} f(x) \cdot (v_2, \ldots, v_k)\Big) \cdot v_1$$

que $d^k f(x_0)$ est symétrique par rapport aux variables $(v_2, \ldots, v_n)$. La relation

$$d^k f(x_0) \cdot v = d^2\Big(d^{k-2} f(x) \cdot (v_3, \ldots v_k)\Big) \cdot (v_1, v_2)$$

montre, en appliquant le cas $k = 2$ à la fonction $x \longmapsto d^{k-2} f(x) \cdot (v_3, \ldots v_k)$, que $d^k f(x_0)$ est symétrique en les variables (v_1, v_2). $\square$

Pour $a_k \in \mathcal{L}_s^k(E, F)$, on note $a_k x^k := a_k(x, x \ldots, x)$. Un polynome de E dans F est une expression de la forme

$$P(x) = a_0 + a_1 x + \ldots + a_k x^k,$$

avec $a_i \in \mathcal{L}_s^i(E, F)$, en notant $\mathcal{L}_s^0(E, F) := F$, $\mathcal{L}_s^1(E, F) := \mathcal{L}(E, F)$.

Propriété 20. *Les polynomes sont différentiables à tous ordres. De plus, on a*

$$d^l(a_k x^k) = \frac{k!}{(k-l)!} a_k x^{k-l}$$

pour tout $a_k \in \mathcal{L}_s^k(E, F)$. Dans cette expression, $a_k x^{k-l}$ est l'élément de $\mathcal{L}_s^{k-l}(E, F)$ défini par

$$a_k x^{k-l} \cdot (v_1, \ldots, v_l) = a_k(v_1, \ldots, v_l, x, \ldots, x).$$

Démonstration. Nous allons montrer que l'application $a_k x^k$ est différentiable de différentielle $k a_k x^{k-1}$. Le résultat général s'en déduit par récurrence en considérant l'application a_k comme un élément de $\mathcal{L}_s^{k-l}(E, \mathcal{L}_s^l(E, F))$. On a

$$a_k(x + y)^k = \sum_{i=0}^{k} C_k^i a_k x^{k-i} y^i = a_k x^k + k a_k x^{k-1} y + \|y\| \epsilon(y).$$

$\square$

Proposition 21. *Si $f : E \longrightarrow F$ est k fois différentiable en $g(x_0)$ et si $g : G \longrightarrow E$ est k fois différentiable en x_0, alors $f \circ g$ est k fois différentiable en x_0.*

Démonstration. On fait la preuve par récurrence, le cas $k = 1$ a déjà été démontré. Nous ferons l'hypothèse que la conclusion est vraie pour toutes les fonctions f et g au rang $k - 1$. On a

$$d(f \circ g)(x) = df(g(x)) \circ dg(x).$$

L'application dg est $k-1$ fois différentiable en x_0 par hypothése, ainsi que l'applications $df \circ g : x \longmapsto df(g(x))$, par l'hypothèse de récurrence. Comme la composition est une application bilinéaire, et donc C^∞, de $\mathcal{L}(E, F) \times \mathcal{L}(F, G)$ dans $\mathcal{L}(E, G)$, on peut encore appliquer l'hypothèse de récurrence et on conclut que $d(f \circ g)$ est $k-1$ fois différentiable en x_0, et donc que $f \circ g$ est k fois différentiable en x_0. $\square$

Propriété 22. *L'application $L \longmapsto L^{-1}$, de $\mathcal{GL}(E, F)$ dans $\mathcal{L}(F, E)$ est C^∞.*

Démonstration. Fixons $L \in \mathcal{GL}(E, F)$ et notons σ l'application $l \longmapsto l^{-1}$. On a vu que

$$\sigma(L - G) = \sigma(L) \circ \sigma(I - G\sigma(L)) = \sigma(L) \sum_{i=0}^{\infty} (G\sigma(L))^i = \sigma(L)(I + G\sigma(L)) + \|G\| \epsilon(G)$$

donc σ est différentiable en L, et $d\sigma(L)\cdot G = -\sigma(L)\circ G\circ\sigma(L)$. On conclut de cette expression que $d\sigma$ est C^k si σ est C^k, et donc, par récurence, que σ est C^∞. $\square$

Définission le polynome de taylor

$$T_y^k f(x) := \sum_{i=0}^{k} \frac{1}{i!} d^i f(y)\cdot(x-y)^i.$$

Nous allons étudier dans quelle mesure c'est une bonne approximation de $f(x)$.

Théorème 6 (Inégalité de Taylor-Lagrange)**.** *Si f est $(k+1)$ fois différentiable sur $[y,x]$, alors*

$$\|f(x)-T_y^k f(x)\| \leqslant \frac{1}{(k+1)!}\sup_{z\in[y,x]}\|d^{k+1}f(z)\|\|x-y\|^{k+1}$$

Démonstration. On fait la preuve par récurrence sur k. Le cas $k=1$ est l'inégalité des accroissements finis. On remarque que $d(T_y^k f)(x) = T_y^{k-1}(df)(x)$. En posant $g(x)=f(x)-T_y^k f(x)$, on a donc

$$dg(x) = df(x) - T_y^{k-1}(df)(x).$$

On applique l'hypothothèse de récurrence à df à l'ordre $k-1$, et on obtient (en remarquant que $dg(0)=0$):

$$\|dg(x)\| \leqslant \frac{1}{k!}\sup_{z\in[y,x]}\|d^{k+1}f(z)\|\|x-y\|^k.$$

Le lemme ci-dessous implique l'inégalité cherchée puisque $g(0)=0$. $\square$

Lemme 23. *Soit $g: E\longrightarrow F$ une fonction différentiable telle que $\|dg(x)\|\leqslant C\|x\|^k$ et $g(0)=0$. Alors $\|g(x)\|\leqslant C\|x\|^{k+1}/(k+1)$.*

Démonstration. Soit $l\in F'$ de norme au plus 1 et $x\neq 0$. On considère la fonction $h(t) = l\circ g(tx/\|x\|)$. On voit que $|h'(t)|\leqslant Ct^k$ pour tout $t\geqslant 0$. La fonctoon $h(t)-Ct^{k+1}/(k+1)$ a une dérivée négative, elle est donc décroissante, donc négative. On a donc $h(t)\leqslant Ct^{k+1}/(k+1)$. De le même façon, $h(t)\geqslant Ct^{k+1}/(k+1)$, et finalement

$$|l\circ g(x)| = |h(\|x\|)| \leqslant C\|x\|^{k+1}/(k+1).$$

$\square$

Théorème 7 (Inégalité de Taylor-Young)**.** *Si f est k fois différentiable sur $[y,x]$, alors*

$$\|f(x)-T_y^k f(x)\| \leqslant \frac{1}{k!}\sup_{z\in[y,x]}\|d^k f(z)-d^k f(y)\|\|x-y\|^k.$$

En particulier, si $d^k f$ est continue en y, alors f admet en y le développement limité

$$f(x) = T_y^k f(x) + \|x-y\|^k\epsilon(x-y).$$

Démonstration. On applique l'inégalité de Taylor-Lagrange à l'ordre $k-1$ à la fonction $g(x) = f(x)-T_y^k f(x)$, dont le développement de Taylor à l'odre $k-1$ en y est nul, et tel que $d^k g(z) = d^k f(z) - d^k f(x)$ pour tout z. $\square$

Corollaire 24. *Soit $f = E \longrightarrow \mathbb{R}$ une application C^2 au voisinage de 0. Si $df(0) = 0$ et si $d^2f(0)$ est définie positive, c'est à dire si il existe $a > 0$ tel que $d^2f(0) \cdot (x, x) \geqslant a\|x\|^2$ pour tout $x \in E$, alors f admet un minimum local strict en 0.*

Démonstration. On a

$$f(x) = f(0) + \frac{1}{2}d^2f(0) \cdot x^2 + \|x\|^2\epsilon(x) \geqslant f(0) + \left(\frac{a}{2} + \epsilon(x)\right)\|x\|^2.$$

□

Théorème 8 (Formule de Taylor). *Si f est C^{k+1} au voisinage du segment $[y, x]$, alors*

$$\begin{aligned} f(x) - T_y^k f(x) &= \frac{1}{k!}\int_0^1 (1-t)^k d^{k+1}f(y + t(x-y)) \cdot (y-x)^{k+1} dt \\ &= \frac{1}{k!}\int_0^1 (1-t)^k d^{k+1}f(y + t(x-y)) dt \cdot (y-x)^{k+1}. \end{aligned}$$

Démonstration. On fait la preuve par récurrence sur k en intégrant par parties. Posons

$$g(t) = \frac{1}{k!}(1-t)^k d^k f(y + t(x-y)) \cdot (x-y)^k$$

On a

$$\begin{aligned} g'(t) =& \frac{1}{k!}(1-t)^k d^{k+1}f(y + t(x-y)) \cdot (x-y)^{k+1} \\ & - \frac{1}{(k-1)!}(1-t)^{k-1} d^k f(y + t(x-y)) \cdot (x-y)^k \end{aligned}$$

donc

$$\begin{aligned} g(0) - g(1) =& \frac{1}{k!} d^k f(y + t(x-y)) \cdot (x-y)^k \\ =& \frac{1}{k!}\int_0^1 (1-t)^k d^{k+1}f(y + t(x-y)) \cdot (y-x)^{k+1} dt \\ & - \frac{1}{(k-1)!}\int_0^1 (1-t)^{k-1} d^k f(y + t(x-y)) \cdot (x-y)^k dt, \end{aligned}$$

ce qui permet de déduire la formule au rang k de la formule au rang $k-1$.

6 Fonctions convexes

La fonction $f : E \longrightarrow]-\infty, +\infty]$ est dite convexe si $f(\alpha x + \beta y) \leqslant \alpha f(x) + \beta f(y)$ pour touts $x \neq y$ dans E, et tous α et β dans $]0, 1[$ tels que $\alpha + \beta = 1$. On dit que f est *strictement convexe* si l'inégalité est stricte. L'ensemble des points $x \in E$ en lesquels la fonction convexe f prend une valeure réelle est une partie convexe de E, appelée domaine de la fonction f.

Propriété 25. *Une fonction strictement convexe admet au plus un minimum.*

Démonstration. Si x et y sont deux minimas distincts, alors la stricte convexité implique que f prend des valeurs strictement inférieures à $f(x) = f(y)$ sur l'intervalle $]x, y[$, ce qui est une contradiction. □

Rappelons quelques caractérisations des fonctions convexes dérivables d'une variable réelle:

Propriété 26. *Soit $f : I \longrightarrow \mathbb{R}$ une fonction convexe dérivable, où I est un intervalle ouvert de $\mathbb{R}$. Alors les points suivants sont équivalents:*

- *Pour tous $x < y$ dans I, on a*

$$f'(x) \leqslant \frac{f(y) - f(x)}{y - x} \leqslant f'(y).$$

- *f' est croissante.*
- *f est convexe.*

Démonstration. Le premier point implique clairement le second.

Si f' est croissante, alors la fonction $y \longmapsto f(\alpha x + \beta y) - \alpha f(x) - \beta f(y)$ a une dérivée positive pour $y \geqslant x$, donc elle est positive pour $y \geqslant x$, ce qui implique la convexité de f.

Si f est convexe, alors le taux d'accroissement $f(y) - f(x)/(y - x)$ est une fonction croissante de y et une fonction décroissante de x pour $y > x$. En effet, pour $z = ty + (1 - t)x$, $t \in]0, 1[$, on a

$$\frac{f(z) - f(x)}{z - x} \leqslant \frac{tf(y) + (1 - t)f(x) - f(x)}{t(y - x)} = \frac{f(y) - f(x)}{y - x}.$$

La preuve est identique pour la décroissance en x. En prenant la limite $z \longrightarrow x$ par valeurs supérieures, on obtient que

$$f'(x) \leqslant \frac{f(y) - f(x)}{y - x},$$

l'autre inégalité s'obtient en prenant la limite $z \longrightarrow y$ par valeurs inférieures des taux $(f(y) - f(z))/(y - z)$. □

Proposition 27. *Soit U un ouvert convexe de l'espace de Banach E, et soit $f : U \longrightarrow \mathbb{R}$ une fonction C^1. Les propriétés suivantes sont équivalentes:*

- *f est convexe.*
- *$f(x) - f(y) \geqslant df(y) \cdot (x - y)$ pour tous x et y dans U.*

Démonstration. On remarque qu'une fonction est convexe si et seulement si sa restriction à chaque droite affine de E est convexe. Ceci permet de se ramener au cas où $E = \mathbb{R}$, et à la propriété 26. □

On dit que $l \in E'$ est une sous-différentielle de f en x_0 si

$$f(x) \geqslant f(x_0) + l \cdot (x - x_0)$$

pour tout x.

Proposition 28. *Soit $f : E \longrightarrow] - \infty, +\infty]$ une fonction convexe, et soit x_0 un point tel que $f(x_0) \in \mathbb{R}$. Si de plus, l'une des hypothèses suivantes est satisfaite:*

- *E est de dimension finie,*
- *f est localement majorée au voisinage de x_0,*

alors f admet une sous-différentielle en x_0.

Démonstration. On applique le théorème de Hahn Banach à la fonction $p(x) = f(x_0+x)-f(x_0)$. On trouve une forme linéaire $l \leqslant p$. Si E est de dimension finie, cette forme linéaire est continue. Si f est localement majorée, aliors p est majorée au voisinage de 0, donc l est continue. □

On verra que f est alors continue en x_0.

Proposition 29. *Soit U un ouvert convexe de E et $f : U \longrightarrow \mathbb{R}$ une fonction convexe à valeurs réelles. Si f est localement majorée, alors elle est continue, et même localement Lipschitz. C'est toujours le cas lorsque E est de dimension finie.*

Démonstration. Supposons d'abord que E est de dimension finie, d. Alors, la norme de E étant équivalente à la norme l^∞, les boules fermés de la norme l^∞ consituent une base de voisinage de chaque point de E. Ces boules fermées sont les enveloppes convexes de leurs 2^{d+1} sommets. Comme f est convexe, son maximum sur un boule l^∞ est majoré par son maximum sur les 2^{d+1} sommets. Ce maximum est donc fini, on conclut que f est localement majorée.

Revenons au cas général d'un espace normé quelconque, avec f localement majorée. Comme f admet une sous-différentielle en chaque point, elle est aussi localement minorée, et donc localement bornée. Fixons maintenant une boule $B(x_0, r)$ sur laquelle $|f| \leqslant M$. Soient maintenant deux points y et x dans la boule $B(x_0, r/2)$, avec $f(y) \geqslant f(x)$. Soit z le point en lequel la demi-droite issue de x et passant par y intersecte la sphère $S(x_0, r)$. En utilisant la convexité de f, sur cette demi-droite, on voit que

$$\frac{f(y)-f(x)}{\|y-x\|} \leqslant \frac{f(z)-f(x)}{\|z-x\|} \leqslant \frac{4M}{r}$$

ce qui monrtre que f est Lipschitzienne sur $B(x_0, r/2)$. □

Proposition 30. *Soit U un ouvert convexe de E et $f : U \longrightarrow \mathbb{R}$ une fonction deux fois différentiable. Alors les propriétés suivantes sont équivalentes:*

- *La fonction f est convexe.*
- *La différentielle seconde $d^2f(x)$ est positive pour chaque $x_0 \in U$.*

Démonstration. On se ramène au cas où $E = \mathbb{R}$ en se restreignant à des droites affines en rappelant que $d^2f(x_0) \cdot v^2$ est la dérivée seconde de la fonction $t \longmapsto f(x_0 + tv)$ en 0. Dans ce cas, on a vu que la convexité est équivalente à la croissance de la dérivée, et donc à la positivité de la dérivée seconde. □

7 Théorème d'inversion locale

Soient E et F des espaces de Banach, U et V des ouverts de E et F. On dit que l'application $f : E \supset U \longrightarrow V \subset F$ est un *difféomorphisme* si elle est bijective, différentiable, et d'inverse différentiable. C'est un difféomorphisme C^k si f et son inverse sont C^k. Si f est un difféomorphisme, alors $df(x)$ est un isomorphisme de Banach pour tout $x \in U$, et

$$d(f^{-1})(f(x)) = (df(x))^{-1}.$$

Ceci découle du calcul de la différentielle des compositions $I = f \circ f^{-1} = f^{-1} \circ f$.

Propriété 31. *Pour que l'application $f : E \supset U \longrightarrow V \subset F$ soit un difféomorphisme (C^k), il suffit que f soit un difféomorphisme et qu'elle soit C^k.*

Démonstration. En notant à nouveau σ l'application $L \longmapsto L^{-1}$, on a $d(f^{-1}) = \sigma \circ df \circ f^{-1}$. Si f est C^k et si f^{-1} est $C^l, 0 \leqslant l \leqslant k-1$, alors on en conclut que $d(f^{-1})$ est C^l et donc que f est C^{l+1}. Par récurrence, f est donc C^k. □

On commence par un théorème global:

Théorème 9. *Soit E un espace de Banach. Soit $f = E \longrightarrow E$ une application telle que $(I - f)$ est contractante, c'est à dire qu'elle est Lipschitz de constante $L < 1$. Alors, l'application f est bijective, et son inverse est Lipschitz de constante $(1 - L)^{-1}$. Si de plus f est différentiable en x, alors f^{-1} est différentiable en $y = f(x)$ et $d(f^{-1})(y) = (df(x))^{-1}$.*

Démonstration. L'équation $f(x) = y$ se réécrit $x - f(x) - y = x$. Pour chaque y, le terme de gauche est une fonction contractante de x, qui admet donc un seul point fixe. Autrement dit, pour chaque y, cette équation a une solution et une seule, ce qui montre que f est une bijection. On a de plus

$$\|x - x'\| \leqslant \|f(x) - f(x')\| + \|(I - f)(x) - (I - f)(x')\| \leqslant \|f(x) - f(x')\| + L\|x - x'\|$$

et donc

$$(1 - L)\|x - x'\| \leqslant \|f(x) - f(x')\|$$

ce qui montre que f^{-1} est Lipschitz de constante $(1 - L)^{-1}$. Dans le cas ou f est différentiable en x, on a

$$\|(I - df(x)) \cdot v\| = \|(I - f)f(x + v) - (I - f)(x) + \|v\|\epsilon(v)\| \leqslant (L + \epsilon(v))\|v\|.$$

On en déduit que $\|I - df(x)\| \leqslant L < 1$, et donc que $df(x)$ est inversible. Il reste à montrer que $G := (df(x))^{-1}$ est bien la différentielle de f^{-1} en $f(x)$. On estime pour ceci

$$g(f(x) + w) - g(f(x)) - G \cdot w = g(f(x + G \cdot w) + \|w\|\epsilon(w)) - x - G \cdot w = \|w\|\epsilon(w)$$

où on a utilisé le caractère Lipschitz de g. □

Théorème 10. *Soit $f : E \longrightarrow F$ une application C^1. Si $df(x)$ est un isomorphisme de Banach, alors f est un difféomorphisme local au voisinage de x. C'est à dire qu'il existe des voisinages ouverts U et V de x et $f(x)$ tels que $f : U \longrightarrow V$ est un difféomorphisme (qui est alors un difféomorphisme C^1).*

Démonstration. En considérant l'application $df(x)^{-1} \circ f$, on se ramène au cas où $E = F$ et $df(x) = Id$. L'application $I - f$ est alors C^1, et sa différentielle en x est nulle. Par l'inégalité des accroissements finis, on en conlut qu'elle est Lipschitz de constante $1/4$ sur la boule $\bar{B}(x, r)$ si r est choisi assez petit. Nous allons montrer qu'il existe alors une application $g : E \longrightarrow E$ Lipschitz de constante $1/2$ (sur tout E) et telle que $f = I - g$ sur $B(x, r)$. Par le théorème précédent, $I - g$ est donc globalement inversible, et c'est localement un difféomorhisme. Pour montrer l'existence de la fonction g, on pose par exemple $g = (I - f) \circ P$, où P est la projection sur une petite boule $\bar{B}(x, r)$ décrite dans le lemme ci-dessous. Comme P est 2-Lipschitz sur E et prend ses valeurs dans $B(x, r)$, et comme $(I - f)$ est $(1/4)$-Lipschitz sur cette boule, on conclut que g est $(1/2)$-Lipschitz sur $\mathbb{R}$. □

Lemme 32. *Soit $\bar{B}(r)$ la boule fermée de rayon r et de centre 0 dans l'espace de Banach E. Soit P la "projection" $P : E \longrightarrow \bar{B}(r)$ sur $\bar{B}(r)$ qui fixe la boule $\bar{B}(r)$ et qui, a tout point $x \notin B(r)$ associe le point $P(x) = rx/\|x\|$. L'application P est 2-Lipschitz (elle est même 1-Lipschitz dans le cas où E est un espace de Hilbert).*

Démonstration. On considère deux points x et x', et on suppose que $\|x\| \geqslant \|x'\|$.

Si x et x' sont dans $\bar{B}(r)$, alors on a $P(x') - P(x) = x' - x$.

Si $x \notin \bar{B}(x)$, $x \in \bar{B}(x)$, on a

$$P(x) - P(x') = \frac{r}{\|x\|}x - x' = \left(\frac{r}{\|x\|} - 1\right)x + (x - x')$$

donc $\|P(x) - P(x')\| \leqslant (\|x\| - r) + \|x - x'\| \leqslant 2\|x - x'\|$.

Finalement, si aucun des deux points x et x' n'est dans $\bar{B}(r)$, alors,

$$P(x) - P(x') = \frac{r}{\|x\|}x - \frac{r}{\|x'\|}x' = \frac{r}{\|x\|}(x - x') + \left(\frac{r}{\|x\|} - \frac{r}{\|x'\|}\right)x'$$

donc $\|P(x) - P(x')\| \leqslant \|x - x'\| + (\|x\| - \|x'\|)r/\|x\| \leqslant 2(\|x\| - \|x'\|)$. □

7.1 Formes Normales des applications C^1

Théorème 11. *Soit $f : E \longrightarrow F$ une application qui est C^1 au voisinage de x_0. Supposons de plus que le noyau N et l'image R de df sont fermés et facteurs directs, et choisissons des supplémentaires fermés E_1 et F_1, de sorte que*

$$E = E_1 \oplus N \quad , \quad F = R \oplus F_1.$$

Il existe alors un difféomorphisme local $\varphi : R \times N \longrightarrow E$ tel que $\varphi(0) = x_0$ et tel que

$$f \circ \varphi(r, n) = f(x_0) + r + g(r, n)$$

où $g : R \times N \longrightarrow F_1$ est une application C^1 telle que $dg(0) = 0$.

De plus, il existe un difféomorphisme local $\psi : F \longrightarrow R \times F_1$ tel que $\psi(f(x_0)) = 0$ et tel que

$$\psi \circ f \circ \varphi(r, n) = (r, h(r, n))$$

où $h : R \times N \longrightarrow F_1$ est une application C^1 définie au voisinage de 0 et telle que $dh(0,0) = 0$ et $h(n, 0) = 0$ pour tout n dans un voisinage de 0 (de N).

Démonstration. En notant π_N la projection sur le noyau N parallèlement à E_1 et π_R la projection sur R parallèlement à F_1, on considère l'application $F : E \longrightarrow R \times N$ donnée par

$$F : x \longmapsto (\pi_R \circ (f(x - x_0) - f(x_0)), \pi_N(x - x_0)).$$

Cette application est différentiable, et $dF(0) \cdot v = (\pi_R \circ df(x_0) \cdot v, \pi_N \cdot v)$ Comme la restriction de $df(x_0)$ à E_1 est un isomorphisme, cette application linéaire est un isomorphisme. On peut donc appliquer le théorème d'inversion local à F, on note φ le difféomorphisme local inverse, défini sur un voisinage de 0 dans $R \times N$. On a alors $\pi_R \circ (f - f(x_0)) \circ \varphi(r, n) = r$, et donc $f \circ \varphi$ est de la forme

$$(r, n) \longmapsto f(x_0) + (r, g(r, n)).$$

Comme R est l'image de $df(x_0)$, on a $\pi_{F_1} \circ df(x_0) = 0$, et donc $dg(0) = \pi_{F_1} \circ d(f \circ \varphi)(0) = 0$.

Pour le second point, il suffit de poser

$$\psi : y \longmapsto (\pi_R(y), \pi_{F_1}(y) - g(\pi_R(y), 0)) - f(x_0).$$

On voit alors que $\psi \circ (f \circ \varphi)(r, n) = (r, g(r, n) - g(r, 0))$. □

7.2 Submersions

Une application C^1 est une *submersion* en x_0 si sa différentielle en x_0 est surjective et de noyau facteur direct. On peut alors appliquer le théorème de formes normales:

Si on note N le noyau et E_1 un supplémentaire, il existe un difféomorphisme local $\varphi : N \times F \longrightarrow E$ tel que

$$f \circ \varphi(n, x) = f(x_0) + x.$$

Autrement dit, une submersion est localement équivalente à une projection linéaire. Réciproquement, il est clair que toute projection linéaire est une submersion.

7.3 Immersions

Une application C^1 est une *immersion* en x_0 si sa différentielle en x_0 est injective et d'image fermée et facteur direct.

Le théorème de forme normale donne des difféomorphismes locaux ψ et φ telles que $\psi \circ f \circ \varphi(x) = (x, h(x))$ avec $h = 0$. Autrement dit, on a

$$\psi \circ f \circ \varphi(x) = (x, 0).$$

Une immersion est localement équivalente à l'inclusion linéaire d'un sous-espace facteur direct.

7.4 Applications de rang ou de corang constant

On rappelle que le rang d'une application linéaire est la dimension de son image, et le corang est la codimension de son image.

Corollaire 33. *Si f est C^1 au voisinage de x_0, de rang ou de corang fini et constant, et de noyau facteur direct, alors il existe des difféomorphismes locaux ψ et φ tels que*

$$\psi \circ f \circ \varphi(r, n) = (r, 0).$$

Si $L \in \mathcal{L}(E, F)$ est de rang fini, et si son noyau est facteur direct, alors tout $G \in \mathcal{L}(E, F)$ assez proche de L a un rang au moins égal à celui de L et a un noyau facteur direct.

Démonstration. On applique la seconde version du théorème de forme normale, qui donne des difféomorphismes locaux ψ et φ tels que $\psi \circ f \circ \varphi(r, n) = (r, h(r, n))$. Comme le rang ou le corang de df sont finis et localement constants, on a nécessairement $\partial_n h(r, n) = 0$ au voisinage de 0. Comme de plus $h(r, 0) = 0$, ceci implique que h est nulle au voisinage de 0. □

Contrairement aux cas des immersions et submersions évoqués ci-dessus, on a besoin d'ici d'une hypothèse sur $df(x)$ dans un voisinage de x_0, et pas seulement en x_0. En fait, les submersions sont les applications de corang localement nul. La raison pour laquelle il suffit de supposer la surjectivité en x_0 de $df(x_0)$ pour savoir que $df(x)$ est localement surjective (c'est à dire de corang nul) est la semi-continuité supérieure du corang. Plus précisément, on a:

Propriété 34. *Soit $L \in \mathcal{L}(E, F)$ une application d'image fermée, de corang fini et de noyau facteur direct. Alors, toute $G \in \mathcal{L}(E, F)$ assez proche de L a un corang fini, inférieur à celui de L, et un noyau facteur direct.*

Démonstration. Soit E_1 un supplémentaire du noyau de L et π une projection sur l'image de L. Alors $\pi \circ L_{E_1}$ est un isomorphisme. Il en est donc de même de $\pi \circ G_{E_1}$. Ceci implique que l'image de G_{E_1} est de codimension inférieure à celle de L. De plus, ceci implique (voir le cours sur les opérateurs de Fredholm) que l'image de G et le noyau de G sont fermés et facteurs directs. □

7.5 Sous Variétés

Soit E un Banach et N un sous-espace vectoriel fermé facteur direct de E. Soit M une partie de E et x_0 un point de M. Les propriétés suivantes sont équivalentes:

1. Il existe un supplémentaire fermé E_1 de N dans E et une fonction $g : N \longrightarrow E_1$ telle que $dg_0 = 0$ et telle que M est localement le graphe de g. Plus précisément, il existe un voisinage ouvert U et 0 dans N et un voisinage ouvert V de 0 dans E_1 tels que
$$M \cap (x_0 + U + V) = \{x_0 + (x, g(x)), x \in U\}.$$

2. Il existe une submersion $P : E \longrightarrow E_1$ telle que le noyau de $dP(x_0)$ est N et telle que, localement, $M = P^{-1}(0)$. Plus précisément, il existe un voisinage ouvert U de x_0 et une submersion $P : U \longrightarrow E_1$ telle que $M \cap U = \{x \in U : P(x) = 0\}$.

3. Il existe un difféomorphisme local ψ vérifiant $d\psi_{x_0} = Id$ et tel que, localement, $\psi(M) = N$. Plus précisément, il existe des voisinages ouverts U de x_0 et V de 0 tels que $\psi : U \longrightarrow V$ est un difféomorphisme et tel que $\psi(M \cap U) = N \cap V$.

4. Il existe une immersion $j : N \longrightarrow E$ dont l'image locale est M. Plus précisément, il existe un voisinage ouvert U de 0 dans N et une immersion $j : U \longrightarrow E$ telle que $j(U)$ est un voisinage de x_0 dans M, et telle que $dj(0)$ est l'inclusion de N dans E.

5. Pour tout supplémentaire fermé E_1 de N dans E, il existe une fonction $g : N \longrightarrow E_1$ telle que $dg_0 = 0$ et telle que M est localement le graphe de g.

Démonstration. Si M est localement le graphe de g, on pose $P(x) = \pi_{E_1}(x) - g \circ \pi_N(x)$.

Si $M = \{P = 0\}$, on pose $\psi(x) = \pi_N(x) + P(x)$.

Si $\psi(M) = N$, on pose $j = \psi^{-1}_{|N}$.

Si $M = j(N)$, on remarque que $\pi_N \circ j$ est un difféomorphisme local de N et on pose $g = \pi_{E_1} \circ (\pi_N \circ j)^{-1}$. □

7.6 Théorème des fonctions implicites

Théorème 12. *Soit $f : E \times F \longrightarrow G$ une application C^1 au voisinage de (x_0, y_0). Si $\partial_y f(x_0)$ est un isomorphisme de Banach, alors il existe un voisinage ouvert U de x_0 dans E, un voisinage ouvert V de y_0 dans F application $g : U \longrightarrow F$ qui est C^1 et telle que*
$$f(x, y) = 0 \quad \Leftrightarrow \quad y = g(x) \quad \textit{pour} \quad (x, y) \in U \times V.$$
De plus, on a $dg(x_0) = -(\partial_y f(x_0))^{-1} \circ \partial_x f(x_0)$.

Démonstration. L'application f est une submersion en (x_0, y_0), donc l'ensemble $M := \{(x, y) \in U \times V : f(x, y) = f(x_0, y_0)\}$ est, localement au voisinage de (x_0, y_0), une sous-variété tangente au noyau de $df(x_0, y_0)$. Ce noyau N est le graphe de l'application linéaire $L := -(\partial_y f(x_0))^{-1} \circ \partial_x f(x_0) : E \longrightarrow F$. On a $E \times F = N \oplus (\{0\} \times F)$. Au vu des caractérisations des sous-variétés, il existe un voisinage $\tilde{U}$ de 0 dans N, une application $\tilde{g} : \tilde{U} \longrightarrow F$ et un voisinage $\tilde{V}$ de y_0 dans F telle que $M \cap ((x_0, y_0) + \tilde{U} + (\{0\} \times \tilde{V})) = \{(x_0, y_0) + n + (0, \tilde{g}(n)), n \in \tilde{U}\}$.

En posant $U = \{x_0 + x : x + (0, L \cdot x) \in \tilde{U}\}$ et $g(x) = y_0 + \tilde{g}(x - x_0, L \cdot (x - x_0)) + L \cdot (x - x_0)$, on a alors $M = \{(x, g(x)), x \in U\}$. De plus, $\tilde{g}$ est C^1 et $d\tilde{g}(0) = 0$, donc g est C^1 et $dg(x_0) = L$.

Pour retrouver facilement l'expression de $dg(x_0)$, on remarque que $f(x, g(x))$ est identiquement nulle, donc en différentiant en x_0,

$$0 = \partial_x f(x_0, y_0) + \partial_y f(x_0, y_0) \circ dg(x_0).$$

□

8 Équations différentielles

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

$$x'(t) = F(t, x(t)) \tag{1}$$

avec $F : \mathbb{R} \times E \longrightarrow E$. C'est une équation *non-autonome* du premier ordre. Rappelons qu'on peut formellement réduire les équations non-autonomes à des équations autonomes (c'est à dire de la forme $y'(t) = G(y(t))$. Il suffit pour ceci de poser $y(t) = (t, x(t))$ et $G(y) = G(t, x) = (1, F(t, x))$. Même si cette réduction est fort utile, il y a diverses raisons de préférer considérer une équation non autonome. Par exemple, les hypothèses dans l'énoncé du théorème de Cauchy-Lipschitz que nous donnons ci-dessous ne sont pas symétriques en t et x. Cet énoncé ne pourrait être déduit facilement d'un énoncé autonome.

L'équation (1) seule ne détermine pas une unique courbe $x(t)$. Pour déterminer une solution, il faut lui adjoindre une condition initiale, c'est à dire prescrire sa valeur en un temps initial, que l'on prendra souvent nul.

Le Couple constitué de l'équation (1) et de la condition initiale $x(0) = x_0$ (ou, plus généralement, $x(t_0) = x_0$) est appelé problème de Cauchy.

Théorème 13 (Cauchy-Lipschitz)**.** *Supposons que F est continue et Lipschitz en x dans un voisinage de (t_0, x_0) (avec une constante de Lipschitz indépendante de t). Alors il existe un temps $\tau > 0$ tel que, pour tout $T \in]0, \tau]$, le problème de Cauchy admet une unique solution $x(t) :]t_0 - T, t_0 + T[\longrightarrow E$.*

On remarque en suivant la preuve que la continuité en t n'est pas aussi important que le caractère Lipschitz en x, on pourrait affaiblir cette hypothèse.

Démonstration. On fixe un voisinage $J \times B(x_0, r)$ de (t_0, x_0) et des constantes M et L telles que

$$\|F(t, x)\| \leqslant M \quad , \quad \|F(t, x) - F(t, y)\| \leqslant L\|x - y\|$$

pour tous $t \in J$, $x \in B(x_0, r)$ et $y \in B(x_0, r)$. La courbe $x(t) :]t_0 - T, t_0 + T[\longrightarrow E$ est une solution du problème de Cauchy si et seulement si elle est continue et

$$x(t) = x_0 + \int_{t_0}^{t} F(s, x(s))ds$$

pour tout $t \in]t_0 - T, t_0 + T[$. Soit $\mathcal{F}$ l'application de $C^0(]t_0 - T, t_0 + T[, E)$ qui, à la courbe x associe la courbe

$$\mathcal{F}(x)(t) = x - 0 + \int_{t_0}^{t} F(s, x(s))ds.$$

Si x prend ses valeurs dans $\bar{B}(x_0, r)$, et si $]t_0 - T, t_0 + T[\subset J$, alors $\|F(s, x(x))\| \leqslant M$ pour tout $s \in]t_0 - T, t_0 + T[$, et donc $\|\mathcal{F}(x)(t) - x_0\| \leqslant |t - t_0|M$. Si on fait sur T l'hypothèse supplémentaire que

$$TM < r,$$

on conclut que $\mathcal{F}$ préserve le fermé $C^0(]t_0 - T, t_0 + T[, \bar{B}(x_0, r))$ (muni de la norme uniforme). Si l'on considère deux courbes x et y de $C^0(]t_0 - T, t_0 + T[, \bar{B}(x_0, r))$, on a

$$\|\mathcal{F}(x)(t) - \mathcal{F}(y)(t)\| = \left\| \int_{t_0}^{t} F(s, x(s)) - F(s, y(s)) ds \right\| \leqslant \int_{t_0}^{t} L\|x(s) - y(s)\| ds \leqslant TL\|x - y\|.$$

En imposant sur T la condition supplémentaire que

$$TL < 1$$

on constate donc que $\mathcal{F}$ est une contraction de l'espaces complet $C^0(]t_0 - T, t_0 + T[, \bar{B}(x_0, r))$. Il y a donc un unique point fixe, donc une unique solution à valeurs dans $\bar{B}(x_0, r)$. Pour démontrer l'unicité des solutions à valeurs dans E, on vérifie qu'une telle solution $x(t)$ prend necessairement ses valeurs dans $B(x_0, r)$. On note pour ceci $]S^-, S^+[$ la composante connexe de t_0 dans l'ensemble $\{t \in]t_0 - T, t_0 + T[: x(t) \in B(x_0, r)\}$. Alors, comme $\|F(s, x(s))\| \leqslant M$ sur $]S^-, S^+[$, on a $\|x(S^+) - x_0\| \leqslant (S^+ - t_0)M \leqslant TM < r$, donc $x(S^+) \in B(x_0, r)$. De la même façon, $x(S^-) \in B(x_0, r)$. Ceci montre que $S^{\pm} = t_0 \pm T$. □

Supposons maintenant que la fonction $F : \mathbb{R} \times E \longrightarrow E$ vérifie les hypothèses du Théorème de Cauchy-Lipschitz en chaque (t_0, x_0), c'est à dire qu'elle est continue et localement Lipschitz par rapport à la variable x, localement uniformément en t. On peut supposer pour simplifier que F est localement Lipschitz en (t, x).

Propriété 35. *Deux solutions x et y définies sur un même intervalle I coïncident si il existe $t \in I$ tel que $x(t) = y(t)$.*

Démonstration. L'ensemble $\{t \in I : x(t) = y(t)\}$ est ouvert, fermé et non-vide. □

On dit qu'une solution $x(t)$ définie sur l'intervalle ouvert I est maximale si x ne peut être prolongée en une solution définie sur un intervalle J contenant strictement I.

Propriété 36. *Toute solution (définie sur un intervalle) admet un unique prolongemet maximal.*

Démonstration. Considérons une solution $x : I \longrightarrow E$, et considérons la réunion J de tous les intervalles ouverts sur lesquels il existe une solution qui prolonge x. Comme deux prolongements différents coïncident sur l'intersection de leurs domaines de définition, il existe une extension définie sur J, elle est donc maximale. □

Si on se donne un temps initial t_0 et une condition initiale x_0, il existe donc une unique solution maximale de l'équation différentielle vérifiant $x(t_0) = x_0$. Voici un critère qui doit être satisfait par les solutions maximales définies sur un intervalle borné:

Propriété 37. *Soit $x :]T^-, T^+[\longrightarrow E$ une solution maximale. Si $T^+ < \infty$ (resp. $T^- > -\infty$), alors la fonction $t \longmapsto F(t, x(t))$ n'est pas bornée au voisinage de T^+ (resp de T^-).*

Démonstration. Si la fonction $t \longrightarrow F(t, x(t))$ est bornée, alors la courbe $x(t)$ est Lipschitz, donc elle admet un prolongement par continuité en T^+. On considère alors une solution y de l'équation différentielle définie au voisinage de T^+ et telle que $y(T^+) = x(T^+)$. Étendons alors la fonction $x(t)$ au-delà du temps T^+ par la fonction y. On appelle $\tilde{x}(t)$ cette extension. Montrons que $\tilde{x}$ est une solution de l'équation différentielle, ce qui contredit la maximalité de T^+. Seul le temps T^+ pose problème. La courbe $\tilde{x}$ est dérivable à droite en T^+ et sa dérivée à droite est $y'(T^+) = F(T^+, x(T^+))$. A gauche en T^+, on a

$$(x(T^+) - x(T^+ - t)) = \int_{T^+ - t}^{T^+} F(s, x(s)) ds.$$

On conclut que x, et donc $\tilde{x}$ est dérivable à gauche en T^+ de dérivée $F(T^+, x(T^+))$. Comme les dérivées à gauche et à droite sont égales, $\tilde{x}$ est dérivable en T^+, et elle y vérifie l'équation. $\square$

9 Lemme de Gronwall et estimations

Nous donnons ici une méthode générale pour établir des inégalités sur les solutions des équations différentielles.

Lemme 38. *Soit $f(t, x) : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R}$ une fonction continue, croissante par rapport à x. Soient $x(t)$ et $y(t)$ des fonctions satisfaisant*

$$x(t) \leqslant x_0 + \int_0^t f(s, x(s))ds$$
$$y(t) = y(0) + \int_0^t f(s, y(s))ds.$$

Alors, si $x_0 < y(0)$, on a $x(t) < y(t)$ pour tout $t \geqslant 0$.

Démonstration. Soit $[0, T[$ le plus grand intervalle de la forme $[0, t[$ sur lequel $x < y$. Alors

$$y(T) = y(0) + \int_0^T f(s, y(s))ds > x_0 + \int_0^T f(s, x(s))ds \geqslant x(T)$$

puisque $x(s) \leqslant y(s)$ sur $[0, T]$. Ceci contredit la maximalité de T. $\square$

Voici quelques cas particuliers particulièrement utiles:

Propriété 39. *Si*

$$x(t) \leqslant x_0 + \int_0^t ax(s)ds,$$

alors $x(t) \leqslant x_0 e^{at}$ pour tout $t \geqslant 0$.

Démonstration. On applique le lemme avec $f(t, x) = ax$ et $y(t) = y_0 e^{at}$, $y_0 > x_0$. On remarque en effet que $y' = ay$, et donc $y(t) = y_0 + \int_0^t ay(s)ds$. On conclut que $x(t) < y_0 e^{at}$ pour tout $y_0 > x_0$, et donc que $x(t) \leqslant x_0 e^{at}$. $\square$

Propriété 40. *Si*

$$x(t) \leqslant x_0 + \int_0^t (ax(s) + b)ds,$$

alors, pour tout $t \geqslant 0$,

$$x(t) \leqslant x_0 e^{at} + \frac{b}{a}(e^{at} - 1).$$

Démonstration. Il suffit de remarquer que la fonction $y(t) = y_0 e^{at} + b/a(e^{at} - 1)$ est la solution de l'équation $y' = ay + b$ vérifiant $y(0) = y_0$, et d'appliquer le lemme. $\square$

On a une version encore plus générale:

Propriété 41 (Lemme de Gronwall). *Soient $a(t) : [0,\infty) \longrightarrow [0,\infty)$ et $b(t) : [0,\infty) \longrightarrow \mathbb{R}$ des fonctions continues. Si*

$$x(t) \leqslant x_0 + \int_0^t \big(a(s)x(s) + b(s)\big)ds$$

pour tout $t \geqslant 0$, alors

$$x(t) \leqslant x_0 e^{A(t)} + \int_0^t b(s)e^{A(t)-A(s)}ds$$

pour tout $t \geqslant 0$, où $A(t) = \int_0^t a(s)ds$.

Démonstration. Il suffit de vérifier que la fonction $y(t) = y_0 e^{A(t)} + \int_0^t b(s)e^{A(t)-A(s)}ds$ est la solution de l'équation $y'(t) = a(t)x(t) + b(t)$ vérifiant $y(0) = y_0$. On applique alors le lemme, qui nous assure que $x(t) < y(t)$ pour $y_0 > x_0$. □

Ces propiétés vont nous permettre d'obtenir de nombreuses inégalités sur les solutions des équations différentielles. On commence par estimer comment une variation de la condition initiale affecte la solution:

Propriété 42. *Soient $x(t)$ et $y(t)$ deux solutions de l'équation $x'(t) = F(t, x(t))$ définies sur le même intervalle I. Si F est L-Lipschitz en x, alors*

$$\|x - y\|(t) \leqslant \|x - y\|(0)e^{L|t|}.$$

Ceci redonne l'unicité, puisque si $x(0) = y(0)$, alors $\|x - y\|$ est identiquement nul.
Démonstration. On a

$$x(t) = x(0) + \int_0^t F(s, x(s))ds \quad , \quad y(t) = y(0) + \int_0^t F(s, y(s))ds,$$

donc

$$\|x - y\|(t) = \left\| x_0 - y_0 + \int_0^t F(s, x(s)) - F(s, y(s))ds \right\| \leqslant \|x - y\|(0) + \int_0^t L\|x - y\|(s)ds.$$

On conclut par le Lemme de Gronwall que

$$\|x - y\|(t) \leqslant \|x - y\|(0)e^{Lt}$$

pour tout $t \geqslant 0$. L'inégalité pour $t \leqslant 0$ s'obtient par exemple en considérant les fonctions $x(-t)$ et $y(-t)$, qui résolvent l'équation $z'(t) = -f(t, z(t))$. □

Ayant estimé comment la solution dépend de la condition initiale, on va maintenant estimer comment elle dépend de l'équation:

Propriété 43. *Soient $x(t)$ et $y(t)$ des solutions respectivement des équations $x'(t) = F(t, x(t))$ et $y'(t) = G(t, x(t))$ définies sur le même intervalle I. Supposons que F et G sont L-Lipschitz par rapport à x, et que $\|F - G\|$ est fini. Alors,*

$$\|x - y\|(t) \leqslant \|x - y\|(0)e^{L|t|} + \|F - G\|\frac{e^{L|t|} - 1}{L}.$$

En particulier, si G converge uniformément vers F, alors le solution $y(t)$ de l'équation $y'(t) = G(t,y(t))$ issue de x_0 converge vers la solution $x(t)$ de l'équation $y'(t) = G(t,y(t))$ issue du même point, uniformément sur tout intervalle borné.

Démonstration. Il suffit de remarquer que

$$\|x-y\|(t) \leqslant \|x-y\|(0) + \int_0^t \|F-G\| + L\|x-y\|(s)ds$$

et d'appliquer le Lemme de Gronwall pour obtenir l'inégalité pour $t \geqslant 0$. On procède comme ci-dessus en considérant les courbes $x(-t)$ et $y(-t)$ pour obtenir la conclusion pour $t \leqslant 0$. □

Propriété 44. *Si $F(t,x)$ est continue et si il existe une fonction continue $L(t)$ telles que*

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

pour tous t, x et y, alors les solutions de l'équation différentielle sont définies sur $\mathbb{R}$ et satisfont l'inégalité

$$\|x(t) - x(0)\| \leqslant \left| \int_0^t \|F(s,x(0))\| e^{\left|\int_s^t L(\sigma)d\sigma\right|} ds \right|$$

La preuve utilise une méthode très importante en analyse consistant à déduire un résultat d'existence d'une inégalité *a priori*.

Démonstration. Soit $x(t) :]T^-,T^+[\longrightarrow E$ une solution maximale. Pour $t \in [0,T^+[$, on a

$$\|x(t) - x(0)\| \leqslant \int_0^t \|F(s,x(0))\| + L(s)\|x(s) - x(0)\| ds.$$

Par le Lemme de Gronwall, on déduit que

$$\|x(t) - x(0)\| \leqslant \int_0^t \|F(s,x(0))\| e^{\int_s^t L(\sigma)d\sigma} ds$$

Si T^+ était fini, la solution serait bornée sur $[0,T^+[$, et donc aussi la foction $F(t,x(t))$, ce qui contredirait la maximalité de T^+. Le raisonnement est identique en T^-. □

10 Systèmes différentiels linéaires.

10.1 Systèmes autonomes et exponentielle d'opérateurs

On considère ici un espace de Banach E et une application linéaire $A \in \mathcal{L}(E,E)$. On s'intéresse à l'équation différentielle linéaire

$$x'(t) = A \cdot x(t).$$

On déduit de la théorie générale des équations différentielles que:

Propriété 45. *Pour chaque condition initial $x_0 \in E$, il existe une unique solution $x(t) : \mathbb{R} \longrightarrow E$ vérifiant $x(0) = x_0$. Cette solution vérifie*

$$\|x(t)\| \leqslant \|x_0\| e^{\|A\||t|}.$$

Notons $e^A(x_0)$ la valeur au temps 1 de la solution issue de x_0. On vérifie facilement que l'application $x_0 \longmapsto e^A(x_0)$ est linéaire, et l'estimation ci-dessus montre que e^A est continue, avec

$$\|e^A\| \leqslant e^{\|A\|}.$$

Propriété 46. *La solution de l'équation $x' = Ax$ issue de x_0 est la courbe $t \longmapsto e^{tA} \cdot x_0$.*

DÉMONSTRATION. Soit $x(s)$ la solution étudiée. Pour chaque $t \in \mathbb{R}$, la courbe $y : s \longmapsto x(ts)$ vérifie $y'(s) = tAy(s)$ et $y(0) = x_0$. On a donc $x(t) = y(1) = e^{tA} \cdot x_0$. □

On peut exprimer cette propriété comme suit:

Proposition 47. *La courbe $L : t \longmapsto e^{tA}$ vérifie l'équation différentielle linéaire sur $\mathcal{L}(E)$*

$$L'(t) = A \circ L(t).$$

C'est l'unique solution de cette équation vérifiant $L(0) = Id$.

Propriété 48. *On a les relations*

$$e^{(t+s)A} = e^{tA} \circ e^{sA}.$$

et, si $A \circ B = B \circ A$,

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

DÉMONSTRATION. Pour chaque $s \geqslant 0$, considérons les courbes $Y(t) = e^{(t+s)A}$ et $Z(t) = e^{tA} \circ e^{sA}$. On voit que $Y' = A \circ Y$, et $Z' = A \circ Z$. De plus $Y(0) = Z(0) = e^{sA}$. Par unicité de la solution de l'équation $X' = A \circ X$ issue de e^{sA}, on a $Y(t) = Z(t)$ pour tout t.

Pour montrer la seconde égalité, on considère cette fois les courbes $Y(t) = e^{t(A+B)}$ et $Z(t) = e^{tA} \circ e^{tB}$. On a $Y(0) = Z(0) = Id$. De plus, on a les équations

$$Y'(t) = (A + B) \circ Y(t) \quad , \quad Z'(t) = A \circ Z(t) + e^{tA} \circ B \circ e^{tB}.$$

Si on montre que $e^{tA} \circ B = B \circ e^{tA}$, on conclut que $Y = Z$ par unicité de la solution du problème de Cauchy. Pour montrer cette formule, on constate que les courbes $e^{tA} \circ B$ et $B \circ e^{tA}$ satisfont toutes deux l'équation $X' = AX$, (on utilise ici que $A \circ B = B \circ A$) et qu'elles sont toutes deux issues de la condition initiale B. □

Finalement, on retrouve la définition classique de l'exponentielle:

Propriété 49.

$$a^A = \sum_{k \geqslant 0} \frac{1}{k!} A^k.$$

DÉMONSTRATION. Pour démontrer la formule, nous allons utiliser la construction de e^{tA} comme solution de l'équation intégrale

$$e^{tA} = I + \int_0^t Ae^{sA} ds.$$

On introduit donc l'application $\mathcal{A} : X(t) \longmapsto I + \int_0^t AM(s)ds$ sur $C([\mathbb{R}), \mathcal{L}(E))$. Définissions alors la suite $X_i(t)$ telle que $X_0(t) = I$ et $X_i = \mathcal{A}(X_{i-1})$. On calcule que $X_1(t) = I + tA$, $X_2(t) = I + tA + t^2A^2/2$, et, par récurrence, que

$$X_i(t) = \sum_{k=0}^{i} \frac{t^k}{k!} A^k.$$

La série (qui est normalement convergente) est donc un point fixe de l'équation intégrale, c'est donc la solution de l'équation différentielle. $\square$

Dans cette preuve, l'application $\mathcal{A}$ n'est pas forcément contractante sur $C(\mathbb{R}, \mathcal{L}(E))$, ou même sur $\mathcal{C}([-1,1], \mathcal{L}(E))$. Pour comprendre directement (sans en calculer les termes) que la suite $(X_i)_{|[-1,1]}$ converge, on peut démontrer par récurrence que $\|\mathcal{A}^j(X)(t) - \mathcal{A}^j(Y)(t)\| \leqslant t^j/j!$. Ceci montre que $\mathcal{A}^j$ est contractante sur $\mathcal{C}([-1,1], \mathcal{L}(E))$ pour j assez grand et suffit à garantir la convergence de la suite X_i sur $[-1,1]$, et l'unicité du point fixe de $\mathcal{A}$.

Exercice 50. *Soit $F : E \longrightarrow E$ une application. Si il existe une $j \in \mathbb{N}$ tel que F^j est contractante, alors F a un unique point fixe, qui est la limite de toutes les suites $x_{i+1} = F(x_i)$.*

10.2 Systèmes linéaires non autonomes

On considère une application continue $A(t) : \mathbb{R} \longrightarrow \mathcal{L}(E)$, et on s'intéresse à l'équation

$$x'(t) = A(t) \cdot x(t).$$

On commence par appliquer le Lemme de Gronwall:

Propriété 51. *On a l'inégalité a priori*

$$\|x(t)\| \leqslant \|x_0\| e^{\left|\int_0^t \|A(s)\| ds\right|}$$

donc les solutions se prolongent à $\mathbb{R}$.

On considère alors, pour chaques t_0 et t_1, l'application $R(t_0, t_1)$ qui, au point x_0 de E, associe la valeur en t_1 de la solution $x(t)$ de l'équation qui vérifie $x(t_0) = x_0$. On vérifie que $R(t_0, t_1)$ est une application linéaire, les estimées ci-dessus montrent qu'elle est de plus continue, avec

$$\|R(t_0, t_1)\| \leqslant e^{\left|\int_{t_0}^{t_1} \|A(s)\| ds\right|}.$$

C'est la *résolvante* de l'équation linéaire.

Propriété 52. *On a la propriété de Markov*

$$R(t_0, t_2) = R(t_1, t_2) \circ R(t_0, t_1)$$

pour tous t_0, t_1, t_2.

DÉMONSTRATION. La courbe $t \longmapsto R(t_0, t)x_0$ est solution de l'équation différentielle, et elle vaut $R(t_0, t_1)x_0$ au temps t_1. La courbe $t \longmapsto R(t_1, t) \cdot (R(t_0, t_1)x_0)$ est solution de l'équation, et elle vaut $R(t_0, t_1)x_0$ au temps t_1. Ces deux courbes sont donc égales pour tout x_0. $\square$

La résolvante $R(t_0, t)$ est la solution de l'équation différentielle linéaire $X'(t) = A(t) \circ X(t)$ sur $\mathcal{L}(E)$ qui vaut Id en t_0.

Attention, par analogie au cas scalaire, on peut être tenté de penser que $R(t_0, t) = e^{\int_0^t A(s) ds}$. Mais ce n'est pas le cas en général (sauf si les matrices $A(t)$ commutent entre elles).

Propriété 53.

$$R(0, t) = \lim_{n \longrightarrow \infty} e^{tA((n-1)t/n)/n} \circ \cdots \circ e^{tA(t/n)/n} \circ e^{tA(0)/n}.$$

Démonstration. Supposons que $t \geqslant 0$ pour fixer les idées et on considère pour l'application $A_n(s)$ qui vaut $A(kT/n)$ pour $s \in [kT/n, (k+1)T/n[$. On définit alors Les opérateurs $R_n(s)$ tels que

$$R_n(s) = e^{(s-kt/n)A((n-1)t/n)} \circ \cdots \circ e^{tA(t/n)/n} \circ e^{tA(0)/n}$$

pour $s \in [kT/n, (k+1)T/n[$. On a

$$R_n(s) = Id + \int_0^s A_n(\sigma) \circ R_n(\sigma) d\sigma.$$

On a donc

$$\begin{aligned}
\|R_n(s) - R(0,s)\| &\leqslant \int_0^s \|(A_n - A)(\sigma)\| \|R(0,\sigma)\| + \|A_n(\sigma)\| \|R_n(\sigma) - R(0,\sigma)\| d\sigma \\
&\leqslant \int_0^s \epsilon_n + C\|R_n(\sigma) - R(0,\sigma)\| d\sigma,
\end{aligned}$$

où ϵ_n est une suite qui tend vers 0, et C est une constante, qui ne dépendent pas de $s \in [0,t]$. Par le lemme de Gronwall, on conclut que

$$\|R_n(0,t) - R(t)\| \leqslant \epsilon_n \frac{e^{Ct} - 1}{C}.$$

□

Considérons finalement l'équation non homogène

$$x'(t) = b(t) + A(t) \cdot x,$$

où $b : \mathbb{R} \longrightarrow E$ et $A : \mathbb{R} \longrightarrow \mathcal{L}(E)$ sont continues et bornées. Les solutions sont définies sur $\mathbb{R}$, et données par la formule de Duhamel:

$$x(t) = R(0,t) \cdot x_0 + \int_0^t R(s,t) \cdot b(s) ds,$$

où R est la résolvante de l'équation homogène associé ($\dot{x} = A(t)x$). Cette formule se démontre par la méthode de variation de la constante: On pose $y(t) = R(t,0) \cdot x(t)$, de sorte que $x(t) = R(0,t) \cdot y(t)$, et on calcule

$$b(t) + A(t) \cdot x(t) = x'(t) = A(t) \circ R(0,t) \cdot y(t) + R(0,t) \cdot y'(t) = A(t) \cdot x(t) + R(0,t) \cdot y'(t),$$

dont on déduit que $y'(t) = R(t,0)b(t)$. On a alors $y(t) = y(0) + \int_0^t R(s,0) \cdot b(s) ds$ et

$$x(t) = R(0,t) \cdot x_0 + \int_0^t R(0,t) \circ R(s,0) \cdot b(s) ds.$$

Réciproquement, cette formule donne une solution de l'équation.

11 Flot, dépendance

On s'intéresse ici à l'équation non-linéaire

$$x'(t) = F(t, x(t))$$

et l'on suppose pour fixer les idées que F est continue et que

$$\|F(t,0)\| \leqslant M \quad , \quad \|F(t,x)-F(t,y)\| \leqslant L\|x-y\|$$

pour tout $t \in \mathbb{R}$ et pour tous x et y dans E. Pour tout $(t_0,x_0) \in \mathbb{R}\times E$, il existe une unique solution $x(t) : \mathbb{R} \longrightarrow E$ vérifiant $x(t_0) = x_0$. On note $\varphi_{t_0}^t(x_0)$ la valeur $x(t)$ de cette solution au temps t. Pour chaque $x \in E$, la courbe $t \longmapsto \varphi_{t_0}^t(x)$ est une solution de l'équation différentielle. Comme nous l'avons fait pour les équations linéaires en définissant la résolvante, nous allons nous intéresser à l'application $x \longmapsto \varphi_{t_0}^t(x)$, que l'on appelle le flot de l'équation différentielle.

Théorème 14.

- $\varphi_{t_0}^{t_2} = \varphi_{t_1}^{t_2} \circ \varphi_{t_0}^{t_1}$.
- *L'application $\varphi_{t_0}^{t_1} : E \longrightarrow E$ est un homéomorphisme bi-Lischitz, dont la constante de Lipschitz est majorée par $e^{L|t_1-t_0|}$.*
- *Si F est différentiable par rapport à x, et si $\partial_x F$ est uniformément continue (en (t,x)), alors $\varphi_{t_0}^{t_1}$ est C^1 et sa différentielle $d\varphi_{t_0}^{t_1}(x_0)$ est la résolvante $R(t_0,t_1)$ associée au système linéarisé*

$$v'(t) = \partial_x F(t, \varphi_{t_0}^t(x_0)) \cdot v(t).$$

On pourrait en fait se contenter de continuité de $\partial_x F$ dans le troisième point.

Démonstration. Le premier point se prouve exactement comme dans le cas linéaire.

Le second point découle du lemme de Gronwall.

Montrons la différentiabilité du flot. On peut supposer que $t_0 = 0$. On fixe une solution $x(t) = \varphi_0^t(x_0)$ et on considère une autre solution $y(t)$, on note $v(t) := y(t) - x(t)$. On a

$$\|v(t)\| \leqslant \|v(0)\| e^{L|t|}.$$

Posons $A(t) := \partial_x F(t, x(t))$, et notons $R(s,t)$ la résolvante de l'équation linéaire associée. On a $\|A(t)\| \leqslant L$ et donc

$$\|R(s,t)\| \leqslant e^{L|t-s|}.$$

On va montrer que $v(t_1) = R(0,t_1)\cdot v(0) + \|v(0)\|\epsilon(\|v(0)\|)$, ce qui implique que $R(0,t_1)$ est la différentielle de $\varphi_0^{t_1}$ en x_0.

L'uniforme continuité de $\partial_x F$ implique (via l'inégalité des accroissements finis) l'existence d'un module de continuité ρ tel que

$$\|F(t,y) - F(t,x(t)) - A(t)\cdot(y-x(t))\| \leqslant \rho(\|y-x(t)\|)\|y-x(t)\|$$

pour tout t et tout y. On écrit alors

$$v'(t) = F(t,y(t)) - F(t,x(t)) = A(t)\cdot v(t) + B(t)$$

où

$$\|B(t)\| \leqslant \rho(\|v(t)\|)\|v(t)\|.$$

Le formule de Duhamel

$$v(t_1) = R(0,t_1)\cdot v(0) + \int_0^{t_1} R(s,t_1)\cdot B(s)ds$$

implique (en estimant très grossièrement) que

$$\|v(t_1) - R(0,t_1)\cdot v(0)\| \leqslant |t_1| e^{2|t_1|L} \rho\big(e^{|t_1|L}\|v(0)\|\big)\|v(0)\| = \|v(0)\|\epsilon(\|v(0)\|).$$

Nous avons montré que $d\varphi_0^{t_1}(x_0) = R(0, t_1)$.

Montrons maintenant la continuité de cette différentielle. On considère comme plus haut une seconde solution $y(t)$, et on note $B(t) := \partial_x F(t, y(t))$. On rappelle que

$$\|B(t) - A(t)\| \leqslant \rho(\|y(t) - x(t)\|) \leqslant \rho(e^{|t_1|L}\|x_0 - y_0\|)$$

pour $t \in [-|t_1|, |t_1|]$. On a

$$d\varphi_0^t(x_0) = I + \int_0^t A(s) \circ d\varphi_0^s(x_0)ds \quad , \quad d\varphi_0^t(y_0) = I + \int_0^t B(s) \circ d\varphi_0^s(y_0)ds$$

donc, pour $t \in [-|t_1|, |t_1|]$,

$$\begin{aligned}\|d\varphi_0^t(x_0) - d\varphi_0^t(y_0)\| &\leqslant \int_0^t \|A(s) - B(s)\|\|d\varphi_0^s(x_0)\| + \|A(s)\|\|d\varphi_0^s(x_0) - d\varphi_0^s(y_0)\|ds \\ &\leqslant \epsilon(\|x_0 - y_0\|) + \int_0^t C\|d\varphi_0^s(x_0) - d\varphi_0^s(y_0)\|ds\end{aligned}$$

où ϵ est un module de continuité et C est une constante, qui ne dépendent que de t_1 et de F. Par le lemmme de Gronwall, on conclut que

$$\|d\varphi_0^t(x_0) - d\varphi_0^t(y_0)\| \leqslant \epsilon(\|x_0 - y_0\|)e^{C|t_1|}.$$

Ceci démontre que $d\varphi_0^t$ est uniformément continue, avec un module de continuité uniforme sur tout intervalle de temps borné. □

Addendum 1. *Sous les hypothèses du dernier point du théorème, l'application* $\Phi : (t, x) \longmapsto \varphi_0^t(x)$ *est* C^1.

Démonstration. Montrons d'abord la continuité de Φ. On a vu que , pour chaque t, le flot Φ est Lipschitz de constante $e^{|t|L}$ par rapport à la variable x. De plus, pour $t' \geqslant t$, on a

$$\|\Phi(t', x) - \Phi(t, x)\| \leqslant |t' - t| \sup_{s\in[t,t']} \|F(s, \Phi(s, x))\|.$$

Comme (propriété 44)

$$\|\Phi(s, x)\| \leqslant \|x\|e^{Ls} + M\frac{e^{Ls} - 1}{L}$$

et comme F est bornée sur les bornés, on conclut que Φ est localement Lipschitz.

On a $\partial_t\varphi_0^t(x) = F(t, \varphi_0^t(x))$, qui est continu. On a vu que, pour tout $T > 0$, il existe un module de continuité ρ tel que

$$\|\partial_x\Phi(t, x) - \partial_x\Phi(t, x')\| \leqslant \rho(\|x' - x\|)$$

pour tout $t \in [-T, T]$ et pour tous x et x' dans E. En notant comme plus haut $R(0, s) = \partial_x\Phi(s, x)$, on a l'inégalité, pour $t' \geqslant t$,

$$\|R(0, t') - R(0, t)\| \leqslant (t' - t) \sup_{s\in[t,t']} \|A(s)\|\|R(0, s)\| \leqslant (t' - t)e^{2L(t'-t)}.$$

On conclut que $\partial_x\Phi$ est continue, et donc que Φ est C^1. □

12 Champs de vecteurs et difféomorphismes

On considère ici le cas particulier d'une équation différentielle autonome

$$x'(t) = F(x(t)),$$

ou $F : E \longrightarrow E$ est une application, que l'on appellera un champ de vecteurs. On le supposera ici C^1 et Lipschitz. Il existe alors un flot φ_s^t, qui est C^1.

Propriété 54. $\varphi_s^t = \varphi_0^{t-s}$

Démonstration. Pour tout $x_0 \in E$, les courbes $t \longmapsto \varphi_s^t(x_0)$ et $t \longmapsto \varphi_0^{t-s}(x_0)$ satisfont l'équation $x' = F(x)$ et la condition initiale $x(s) = x_0$. Elles sont donc égales. □

On note dans ce cas φ^t le flot φ_0^t, on a

$$\varphi^{t+s} = \varphi^t \circ \varphi^s.$$

Si ψ est un difféomorphisme C^1, on note $\psi_* F$ le champ de vecteur

$$\psi_* F(x) := d\psi(\psi^{-1}(x)) \cdot F(x).$$

Le flot de $\psi_* F$ est donné par $\psi \circ \varphi^t \circ \psi^{-1}$. Ceci revient à dire que, si $x(t)$ est est une solution de $x' = F(x)$, alors $y(t) := \psi(x(t))$ est une solution de $y' = \psi_* F(y)$. En effet,

$$y't) = d\psi(x(t)) \cdot x'(t) = d\psi(x(t)) \cdot F(x(t)) = \psi_* F(y(t)).$$

Au voisinage d'un point régulier ($F(x_0) \neq 0$), tout champ de vecteurs peut être réduit à une champ de vecteurs constant par un changement de variables:

Théorème 15. *Soit F un champ de vecteurs C^1 sur l'espace de Banach E, et x_0 un point tel que $F(x_0) \neq 0$. Il existe alors un espace de Banach G et un difféomorphisme local*

$$\psi : (E, x_0) \longrightarrow (\mathbb{R} \times G, 0)$$

tel que $\psi_* F = (1, 0)$ au voisinage de 0.

Démonstration. Comme l'énoncé est local, on peut commencer par modifier le champ de vecteurs F en dehors d'un voisinage de x_0 afin de le rendre Lipschitz. Nous supposerons donc qu'il l'est.

On considère un supplémentaire fermé G de la droite $\mathbb{R}F(x_0)$ dans E. On considère alors l'application $\Phi : \mathbb{R} \times G \longrightarrow E$ donnée par $\Phi(t, x) = \varphi^t(x_0 + x)$. L'application Φ est C^1, et

$$\partial_t \Phi(0, 0) = F(x_0) \quad , \quad \partial_x \Phi(0, 0) = I_G$$

où I_G est l'inclusion de G dans E. On conclut que $d\Phi(0, 0)$ est un isomorphisme de Banach, et donc que Φ est un difféomorphisme local. En notant ψ le difféomorphisme inverse, on a

$$\psi \circ \varphi^s \circ \psi^{-1}(t, x) = \psi \circ \varphi^s \circ \Phi(t, x)\psi \circ \Phi(t + s, x) = (t + s, x).$$

Donc ψ conjugue le flot de F avec celui de $(1, 0)$. □

Ce théorème de conjugaison implique l'existence et l'unicité locale des solutions de l'équation $x' = F(x)$, puisque l'existence et l'unicité des solutions est évidente pour l'équation redressée.