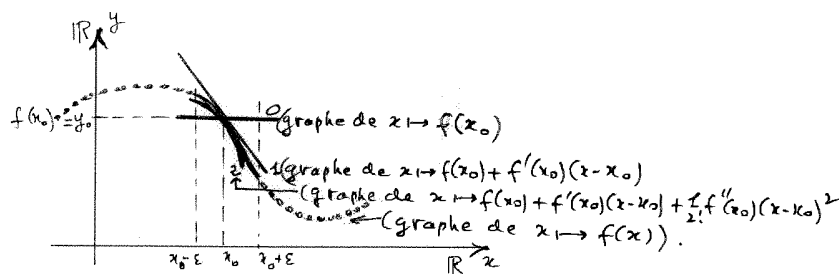


# Calcul différentiel et équations différentielles

## M2309

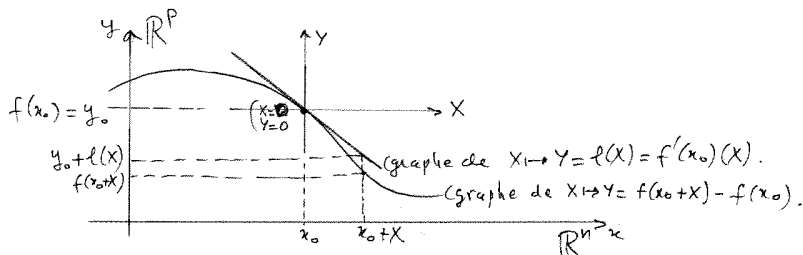
A. Chenciner, résumé de cours n<sup>o</sup>1

**Approximations polynomiales** d'une fonction  $f : (\mathbb{R}, x_0) \cdots \rightarrow (\mathbb{R}, y_0)$  au voisinage du point  $x_0$  (la notation  $\cdots \rightarrow$  signifie que  $f$  est définie sur un voisinage de  $x_0$  et que  $f(x_0) = y_0$ ) :



**Remarque : calcul différentiel sur les polynômes.** L'étude au voisinage d'un point  $x_0 \in \mathbb{R}$  d'une application polynomiale  $f$  est de nature purement algébrique. On pose  $x = x_0 + \xi$  et on réordonne  $f(x_0 + \xi)$  suivant les puissances croissantes de  $\xi = x - x_0$ .

**Généralisation.** Pour généraliser aux applications  $f : (\mathbb{R}^n, x_0) \rightarrow (\mathbb{R}^p, y_0)$  il est impératif de considérer la dérivée (ou différentielle) de  $f$  en  $x_0$  comme étant l'application linéaire  $X \mapsto Y = f'(x_0)X$ .



Sans plus d'effort, on considérera des applications  $f : (E, x_0) \cdots \rightarrow (F, y_0)$ , où  $E$  et  $F$  sont des espaces vectoriels normés, dont la dimension pourrait même être infinie. Les normes sur  $E$  et  $F$  seront toutes deux notées  $\|\cdot\|$

**Définition 1**  $f : (E, x_0) \cdots \rightarrow (F, y_0)$  est dérivable en  $x_0$  s'il existe une application linéaire continue  $l : E \rightarrow F$  telle que  $\rho(x) := f(x) - f(x_0) - l(x - x_0)$  soit un  $o(\|x - x_0\|)$  i.e. vérifie  $\lim_{x \rightarrow x_0} \frac{\|\rho(x)\|}{\|x - x_0\|} = 0$ .

**Rappel important.** 1) L'hypothèse de continuité de  $l$  n'en est une qu'en dimension infinie : les applications linéaires d'un espace vectoriel de dimension finie dans un autre sont toujours continues !

2) Sur un espace vectoriel de dimension finie, toutes les normes sont équivalentes

**Exercices.** 1) Si  $f$  est dérivable en  $x_0$ , elle est continue en  $x_0$ .

2) Si  $E$  et  $F$  sont de dimension finie, la dérivabilité ne dépend pas du choix des normes ;

3) Si  $l$  existe, elle est uniquement déterminée. On l'appelle alors la *dérivée* (ou la *différentielle*) de  $f$  en  $x_0$  et on la note  $Df(x_0)$  ou  $df(x_0)$  ou  $f'(x_0)$  ... (A une notion importante sont souvent attachées de nombreuses notations)

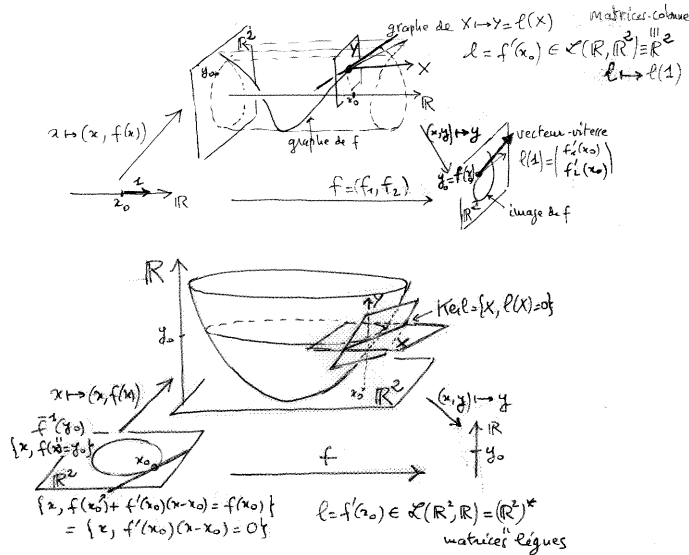
**Définition 2** Deux applications  $f, g : (E, x_0) \cdots \rightarrow (F, y_0)$  sont dites tangentes à l'ordre  $k$  en  $x_0$  si leur différence  $f - g$  est un  $o(\|x - x_0\|^k)$ . On note  $f \sim_k g$ .

**Exercice.** Montrer que  $\sim$  est une relation d'équivalence.

**Exemples.** 1)  $f \sim_0 (x \mapsto y_0)$  ssi  $f$  est continue en  $x_0$  ;

2)  $f \sim_1 (x \mapsto y_0 + l(x - x_0))$  ssi  $f$  est dérivable en  $x_0$  avec pour dérivée  $l$ .

**Courbes planes paramétrées ou définies par une équation.**



Dans les deux cas (courbe  $C$  définie comme  $f(\mathbb{R})$  ou comme  $f^{-1}(y_0)$ ), la tangente au point considéré ( $y_0 = f(x_0)$  dans le premier cas,  $x_0$  dans le second) s'obtient en remplaçant  $x \mapsto f(x)$  par  $x \mapsto f(x_0) + f'(x_0)(x - x_0)$ .

# Calcul différentiel et équations différentielles

## M2309

A. Chenciner, résumé de cours n<sup>o</sup>2

**Rappels d'algèbre linéaire en dimension finie** (*approcher localement une application par une application linéaire n'a d'intérêt que si l'on comprend bien les applications linéaires !*)

Un espace vectoriel  $E$  de dimension  $n$  est isomorphe  $\mathbb{R}^n$ . La particularité de  $\mathbb{R}^n$  est d'être muni d'une *base canonique*

$$\{e_1 = (1, 0, \dots, 0), \dots, e_n = (0, \dots, 0, 1)\}.$$

**Exercice.** Se convaincre que le plan vectoriel de  $\mathbb{R}^3$  d'équation  $x - 2y + z = 0$  ne possède pas une base plus belle que les autres.

La donnée de la base  $\{v_1, \dots, v_n\}$  de  $E$  équivaut à celle de l'isomorphisme  $\varphi : \mathbb{R}^n \rightarrow E$  défini par  $\varphi(e_i) = v_i$  pour  $i = 1, \dots, n$ .

L'espace vectoriel  $L(\mathbb{R}^n, \mathbb{R}^p)$  des applications linéaires de  $\mathbb{R}^n$  dans  $\mathbb{R}^p$  s'identifie canoniquement à celui des matrices à  $p$  lignes et  $n$  colonnes : la  $i^{\text{ème}}$  colonne de la matrice représentant  $f : \mathbb{R}^n \rightarrow \mathbb{R}^p$  est formée des coordonnées dans la base canonique de  $\mathbb{R}^p$  de l'image  $f(e_i)$  du  $i^{\text{ème}}$  vecteur de la base canonique de  $\mathbb{R}^n$ .

Soient  $\varphi : \mathbb{R}^n \rightarrow E$  une base de  $E$  et  $\psi : \mathbb{R}^p \rightarrow F$  une base de  $F$ . La matrice de l'application linéaire  $l : E \rightarrow F$  dans les bases  $\varphi, \psi$  est celle qui est canoniquement associée à  $\psi^{-1} \circ l \circ \varphi : \mathbb{R}^n \rightarrow \mathbb{R}^p$ .

**Théorème 1** *Soient  $E$  et  $F$  deux espaces vectoriels sur  $\mathbb{R}$  de dimensions respectivement  $n$  et  $p$ ; Soit  $l : E \rightarrow F$  une application linéaire de rang  $r$ . Il existe une base de  $E$  et une base de  $F$  dans lesquelles la matrice de  $l$  est formée en haut à gauche d'un bloc  $r \times r$  égal à l'identité et de zéros partout ailleurs.*

Autrement dit, **À CHANGEMENT DE BASES PRÈS À LA SOURCE ET AU BUT, IL Y A UNE SEULE APPLICATION LINÉAIRE DE RANG  $r$  FIXÉ :**

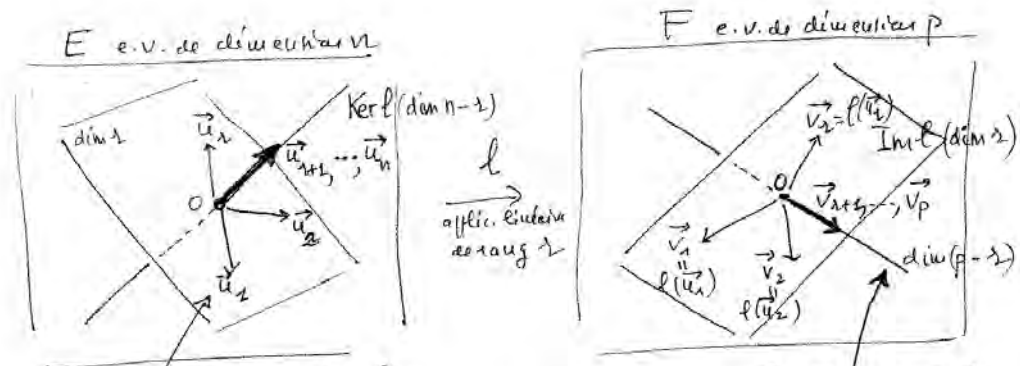
$$(x_1, \dots, x_n) \mapsto (x_1, \dots, x_r, 0, \dots, 0).$$

Cas particuliers :

1)  $n \leq p$ . Une **injection** linéaire se ramène toujours à l'injection canonique  $(x_1, \dots, x_n) \mapsto (x_1, \dots, x_n, 0, \dots, 0)$  de  $\mathbb{R}^n$  dans  $\mathbb{R}^n \times \mathbb{R}^{p-n} \cong \mathbb{R}^p$ .

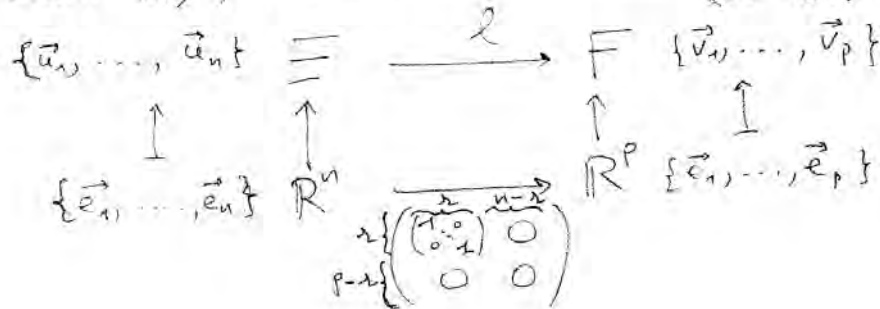
2)  $n \geq p$ . Une **surjection** linéaire se ramène toujours à la projection canonique  $(x_1, \dots, x_n) \mapsto (x_1, \dots, x_p)$  de  $\mathbb{R}^n = \mathbb{R}^p \times \mathbb{R}^{n-p}$  sur  $\mathbb{R}^p$ .

La démonstration du théorème est indiquée sur la figure au dos.

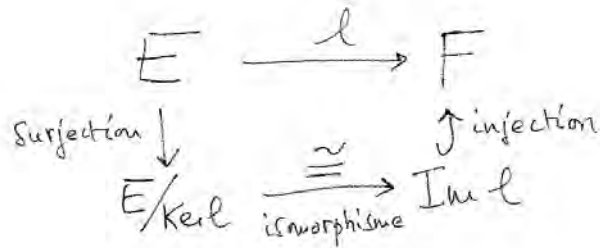


choix: (Supplémentaire de  $\text{Ker } l$   
 (base  $\vec{u}_1, \dots, \vec{u}_2$ )  $\implies$   $\vec{v}_1 = l(\vec{u}_1), \dots, \vec{v}_2 = l(\vec{u}_2)$  base de  $\text{Im } l$

choix:  $\vec{u}_{2+1}, \dots, \vec{u}_n$  base de  $\text{Ker } l$ 
choix: (Supplémentaire de  $\text{Im } l$   
 (base  $\vec{v}_{2+1}, \dots, \vec{v}_p$ )



Corollaire :  $\dim \text{Ker } l + \dim \text{Im } l = \dim E$   
 $\qquad \qquad \qquad n-2 \qquad \qquad 2 \qquad \qquad n$

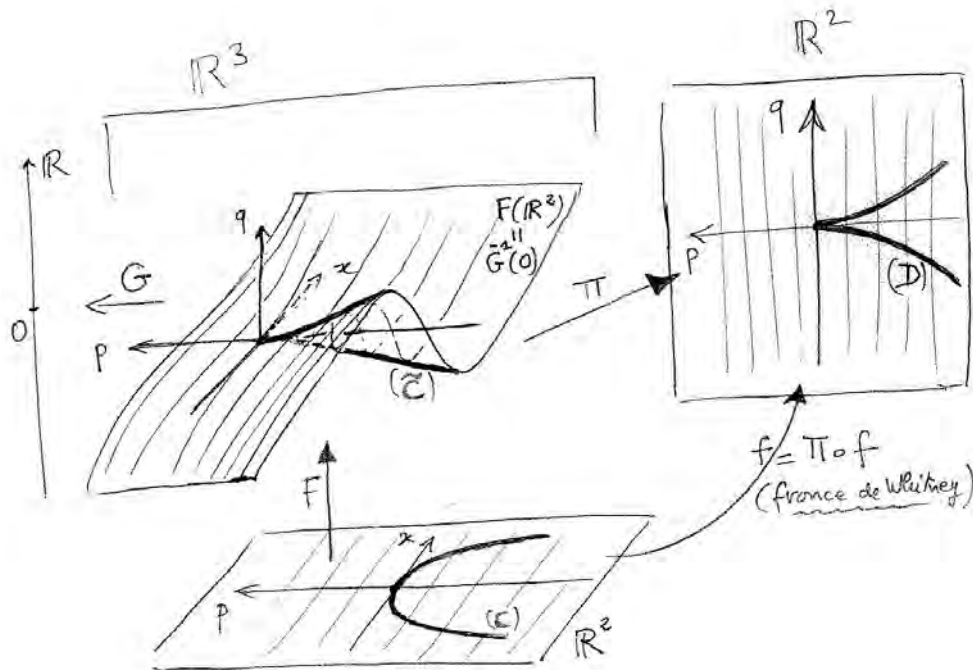


# Calcul différentiel et équations différentielles

A. Chenciner, résumé de cours n°3

**Deux exemples d'applications de  $\mathbb{R}^n$  dans  $\mathbb{R}^p$ .** Il sera bon de tester les définitions et les théorèmes du cours sur ces exemples.

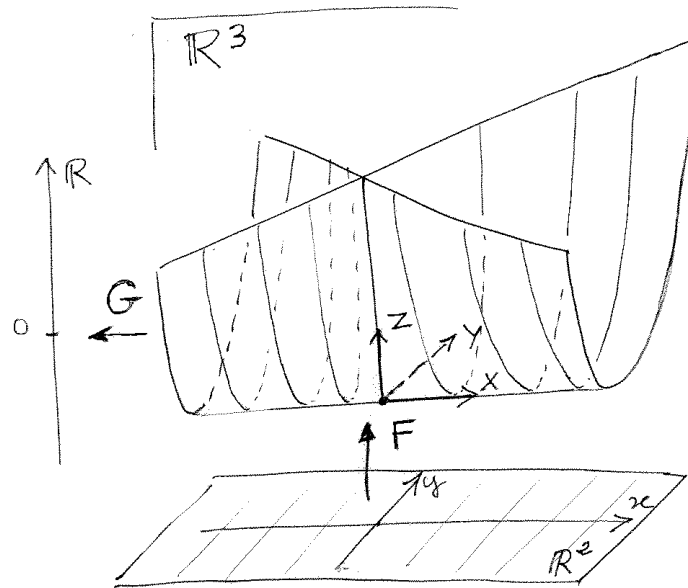
**Exemple 1 : la fronce de Whitney.** C'est une application de  $\mathbb{R}^2$  sur  $\mathbb{R}^2$  que connaissent bien couturiers et couturières. Afin de la mieux comprendre, on l'a écrite comme composée d'une injection (non linéaire) de  $\mathbb{R}^2$  dans  $\mathbb{R}^3$  avec la projection canonique de  $\mathbb{R}^3$  sur  $\mathbb{R}^2$ .



$$F(x, p) = (x, p, x^3 + px), \quad \pi(x, p, q) = (p, q), \quad f(x, p) = (p, x^3 + px), \quad G(x, p, q) = x^3 + px - q.$$

Le lieu singulier (C) de  $f$  (points en lesquels la dérivée n'est pas de rang maximum) a pour équation  $p = -3x^2$ . Son image (D) a pour équation  $4p^3 + 27q^2 = 0$  qui devrait rappeler des souvenirs.

**Exemple 2 : le parapluie de Whitney.**



$$F(x, y) = (x, xy, y^2), \quad G(X, Y, Z) = (X^2Z - Y^2).$$

Hassler Whitney est un mathématicien américain. Il a le premier étudié géométriquement les applications dérivables “générales” de  $\mathbb{R}^2$  dans  $\mathbb{R}^2$ . Avec Marston Morse qui a étudié les fonctions “générales” de  $\mathbb{R}^n$  dans  $\mathbb{R}$  (les célèbres fonctions de Morse dont nous reparlerons), il est à l’origine de la *théorie des singularités* développée par René Thom puis par Vladimir Igorevich Arnold et son école.

# Calcul différentiel et équations différentielles

A. Chenciner, résumé de cours n<sup>o</sup>4

## Applications linéaires et bilinéaires continues.

Dans ce qui suit,  $E, F, G$  sont des espaces vectoriels normés (toutes les normes sont notées  $\|\cdot\|$ ).

1) L'application linéaire  $l : E \rightarrow F$  est continue (en un point et donc en tous) si et seulement si

$$\exists C > 0, \forall X \in E, \|l(X)\| \leq C\|X\|.$$

C'est toujours le cas si  $E$  et  $F$  sont de dimension finie. On note  $L(E, F)$  l'espace vectoriel des applications linéaires continues de  $E$  dans  $F$ .

2) L'application bilinéaire  $B : E \times F \rightarrow G$  est continue (en un point et donc en tous) si et seulement si

$$\exists C > 0, \forall X, Y \in E \times F, \|B(X, Y)\| \leq C\|X\|\|Y\|.$$

C'est toujours le cas si  $E, F, G$  sont de dimension finie. On note  $L_2(E \times F, G)$  l'espace vectoriel des applications bilinéaires continues de  $E \times F$  dans  $G$ .

**Normes sur les espaces d'applications linéaires continues.** Pour parler de la continuité de  $df(x_0)$  par rapport à  $x_0$ , il faut disposer d'une topologie sur l'espace  $L(E, F)$  des applications linéaires continues de  $E$  dans  $F$ . Si  $E$  et  $F$  sont de dimension finie, respectivement  $n$  et  $p$ ,  $L(E, F)$  est un espace vectoriel de dimension  $np$ , toutes les normes dont on peut le munir sont équivalentes et définissent donc la même topologie. Une fois choisies des normes sur  $E$  et  $F$ , celle définie ci-dessous dans le cas général est la mieux adaptée aux majorations. Comme d'habitude, on a noté de la même façon les différentes normes).

**Définition 1** La norme de  $l \in L(E, F)$  est définie comme la plus petite constante  $C$  possible :

$$\|l\| = \sup_{X \neq 0} \frac{\|l(X)\|}{\|X\|} = \sup_{\|X\|=1} \|l(X)\|.$$

En particulier,

$$\forall X \in E, \|l(X)\| \leq \|l\|\|X\|.$$

Le grand intérêt de cette norme est son comportement dans les compositions : si  $E, F, G$  sont trois espaces vectoriels normés,  $l_1 \in L(E, F)$ ,  $l_2 \in L(F, G)$ , on a :

$$\|l_2 \circ l_1\| \leq \|l_2\|\|l_1\|.$$

**Exercice.** Définir de même la norme d'une application bilinéaire continue  $B$  de façon que  $\|B(X, Y)\| \leq \|B\| \|X\| \|Y\|$ .

### Dérivée dans une direction, dérivées partielles.

**Définition 2** Si  $f : (E, x_0) \cdots \rightarrow (F, y_0)$  est dérivable en  $x_0$ , on appelle dérivée de  $f$  en  $x_0$  dans la direction  $X \in E$  l'image  $df(x_0)(X) \in F$  de  $X$ . Si  $E = \mathbb{R}^n$  et  $F = \mathbb{R}^p$  et si  $X = e_i$  est le  $i^{\text{ème}}$  élément de la base canonique de  $\mathbb{R}^n$ , la dérivée de  $f$  en  $x_0$  dans la direction de  $e_i$  est appelée dérivée partielle de  $f$  par rapport à  $x_i$  en  $x_0$  et est notée  $\frac{\partial f}{\partial x_i}(x_0) \in \mathbb{R}^p$ .

**Exercice.** Montrer que, si  $x_0 = (a_1, \dots, a_n)$ ,  $\frac{\partial f}{\partial x_i}(x_0)$  est la dérivée en  $a_i$  de l'application  $(\mathbb{R}, a_i) \cdots \rightarrow (\mathbb{R}^p, y_0) : x_i \mapsto f(a_1, \dots, a_{i-1}, x_i, a_{i+1}, \dots, a_n)$  (rappelons que cette dérivée appartient à  $L(\mathbb{R}, \mathbb{R}^p) \cong \mathbb{R}^p$ ). En déduire que, si  $f = (f_1, \dots, f_p) : (\mathbb{R}^n, x_0) \rightarrow (\mathbb{R}^p, y_0)$ , la matrice de  $df(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}^p$  est

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_0) & \cdots & \cdots & \cdots & \frac{\partial f_1}{\partial x_n}(x_0) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{\partial f_j}{\partial x_1}(x_0) & \cdots & \frac{\partial f_j}{\partial x_i}(x_0) & \cdots & \frac{\partial f_j}{\partial x_n}(x_0) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{\partial f_p}{\partial x_1}(x_0) & \cdots & \cdots & \cdots & \frac{\partial f_p}{\partial x_n}(x_0) \end{pmatrix}$$

**Définition 3** Soient  $E_1, \dots, E_k, F$  des espaces vectoriels normés,  $x_0 = (a_1, \dots, a_k)$  un point de  $E_1 \times \cdots \times E_k$  et  $f : (E_1 \times \cdots \times E_k, x_0) \cdots \rightarrow (F, y_0)$  une application. On définit la dérivée partielle

$$\frac{\partial f}{\partial E_i}(x_0) \in L(E_i, F)$$

par rapport à  $E_i$  comme la dérivée de l'application

$$(E_i, a_i) \ni x_i \cdots \mapsto f(a_1, \dots, a_{i-1}, x_i, a_{i+1}, \dots, a_k) \in (F, y_0).$$

Lorsque  $f : (\mathbb{R}^n, x_0) \rightarrow F$ , l'identification canonique  $L(\mathbb{R}, F) \cong F$  fait coïncider la définition de  $\frac{\partial f}{\partial x_i}(x_0)$  avec celle de la dérivée partielle  $\frac{\partial f}{\partial E_i}(x_0)$  par rapport au  $i^{\text{ème}}$  facteur  $E_i = \mathbb{R}$  de la décomposition  $\mathbb{R}^n = \mathbb{R} \times \cdots \times \mathbb{R}$  de  $\mathbb{R}^n$  en produit de droites.

### Propriétés élémentaires de la dérivation.

**0)** Une application constante  $E \ni x \mapsto f(x) = y_0 \in F$  est dérivable en tout point  $x_0$  et sa dérivée est  $df(x_0) = 0 \in L(E, F)$ .

**1)** Une application linéaire continue  $l : E \rightarrow F$  est dérivable en tout point  $x_0$  et sa dérivée est elle-même :  $dl(x_0) = l$ .

**2)** Une application bilinéaire continue  $B : E \times F \rightarrow G$  est dérivable en tout point  $(x_0, y_0) \in E \times F$  et sa dérivée est  $dB(x_0, y_0)(X, Y) = B(x_0, Y) + B(X, y_0)$ .

**3)** Si  $f : (E, x_0) \cdots \rightarrow (F, y_0)$  et  $g : (E, x_0) \cdots \rightarrow (F, z_0)$  sont dérivables en  $x_0$ ,  $f + g$  l'est aussi et  $d(f + g)(x_0) = df(x_0) + dg(x_0) \in L(E, F)$ .

**4)** Si  $f : (E, x_0) \cdots \rightarrow (\mathbb{R}, y_0)$  et  $g : (E, x_0) \cdots \rightarrow (\mathbb{R}, z_0)$  sont dérivables en  $x_0$ ,  $fg$  l'est aussi et  $d(fg)(x_0) = f(x_0)dg(x_0) + g(x_0)df(x_0) \in L(E, \mathbb{R})$ .



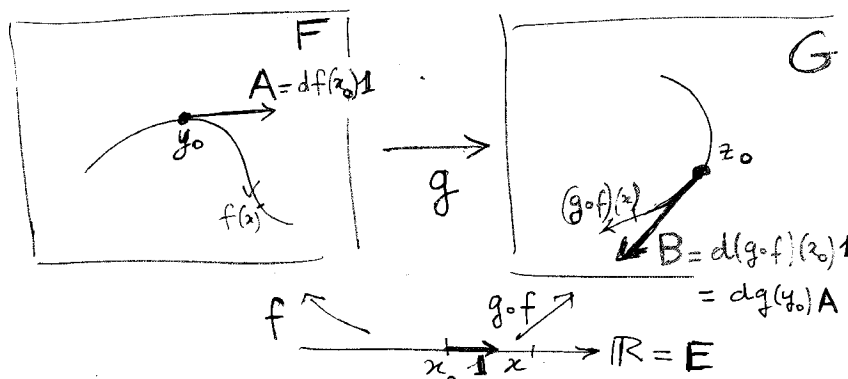
# Calcul différentiel et équations différentielles

A. Chenciner, résumé de cours n<sup>o</sup>5

## Dérivée d'une application composée : chain rule.

**Proposition 1** Si  $f : (E, x_0) \cdots \rightarrow (F, y_0)$  et  $g : (F, y_0) \cdots \rightarrow (G, z_0)$  sont dérivables respectivement en  $x_0$  et en  $y_0$ , la composée  $g \circ f : (E, x_0) \cdots \rightarrow (G, z_0)$  est dérivable en  $x_0$  et sa dérivée est  $d(g \circ f)(x_0) = dg(y_0) \circ df(x_0) \in L(E, G)$ .

## Application : interprétation mécanique de la dérivée.



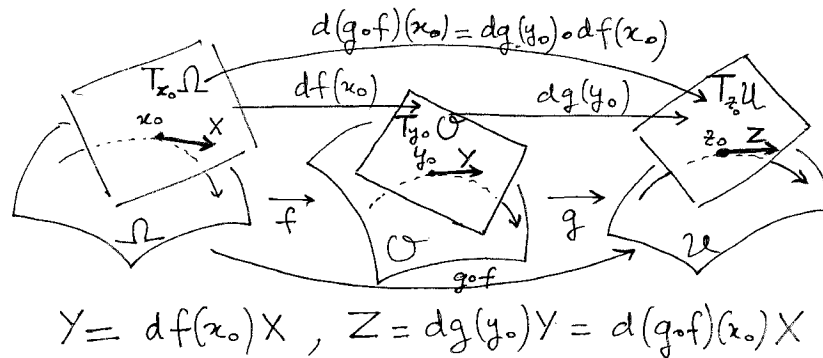
Si  $E = \mathbb{R}$ , l'application  $f : \mathbb{R} \rightarrow F$  peut être interprétée comme une courbe paramétrée dans  $F$  ou encore, si on appelle "temps" la variable dans  $E$ , comme un "mobile" dans  $F$ . Sa dérivée  $df(x_0) \in L(\mathbb{R}, F) \cong F$  s'identifie au vecteur-vitesse  $A = df(x_0)1$  en  $y_0$  de ce mobile alors que la dérivée  $d(g \circ f)(x_0) \in L(\mathbb{R}, G) \cong G$  s'identifie à la vitesse  $B = d(g \circ f)(x_0)1$  en  $z_0$  du mobile image  $g \circ f : \mathbb{R} \rightarrow G$ . Le théorème dit simplement que la dérivée  $dg(y_0)$  applique le vecteur-vitesse  $A$  du premier mobile sur le vecteur-vitesse  $B$  du mobile image :

$$dg(y_0)A = B.$$

On peut donc définir la dérivée en un point d'une application comme l'application qui transforme la vitesse d'un mobile passant par ce point, en la vitesse du mobile image lorsqu'il passe au point image.

**Espaces tangents et fibré tangent.** La formule d'apparence compliquée  $d(g \circ f)(x_0) = dg(f(x_0)) \circ df(x_0)$  devient claire si on l'interprète correctement : la difficulté vient en fait de la confusion de plusieurs notions qui, lorsqu'on remplace les domaines de définition des applications par des espaces "courbes" deviennent distinctes.

Instruits par ce qui précède, étant donné un ouvert  $\Omega$  d'un espace vectoriel normé  $E$ , appelons espace tangent à  $\Omega$  en  $x_0$  l'ensemble des vecteurs-vitesse de mobiles dans  $\Omega$  passant par  $x_0$ . Si  $f : E \supset \Omega \rightarrow \mathcal{O} \subset F$  est dérivable en  $x_0$ , sa dérivée doit être interprétée comme une application de  $T_{x_0}\Omega$  dans  $T_{y_0}\mathcal{O}$ . La translation de vecteur  $x_0$  identifie canoniquement  $T_{x_0}\Omega$  à  $E$  et la translation de vecteur  $y_0$  identifie canoniquement  $T_{y_0}\mathcal{O}$  à  $F$ , ce qui nous a permis de considérer  $df(x_0)$  comme un élément de  $L(E, F)$ . Mais l'interprétation mécanique de la dérivée nous permet de définir espace tangent et dérivée dans un cadre beaucoup plus général, par exemple lorsque  $\Omega$  et  $\mathcal{O}$  sont des surfaces. La figure ci-dessous illustre la "chain rule" dans une telle situation



**Application tangente.** Supposons  $f : E \supset \Omega \rightarrow \mathcal{O} \subset F$  dérivable en tous les points de  $\Omega$ . Puisque l'espace  $T_x\Omega$  tangent à  $\Omega$  en  $x$  s'identifie canoniquement par translation à  $E$ , la réunion  $T\Omega = \cup_{x \in \Omega} T_x\Omega$  des espaces tangents en tous les points  $x$  de  $\Omega$  peut être identifiée au produit  $\Omega \times E$ . On l'appelle le fibré tangent de  $\Omega$ . L'application  $Tf : T\Omega \rightarrow T\mathcal{O}$  définie par la formule

$$Tf(x, X) = (f(x), df(x)X)$$

est appelée "application tangente à  $f$ ". Dans ce langage, la formule de dérivation d'une application composée devient simplement l'application tangente à une composée est la composée des applications tangentés :

$$T(g \circ f) = Tg \circ Tf.$$

# Calcul différentiel et équations différentielles

A. Chenciner, résumé de cours n°6

## Applications de classe $C^1$ .

**Définition 1** Soient  $E$  et  $F$  des espaces normés,  $\Omega \subset E$  un ouvert,  $f : \Omega \rightarrow F$  une application dérivable en tout point de  $\Omega$ . On dit que  $f$  est continuellement dérivable, ou encore qu'elle est de classe  $C^1$ , si l'application

$$df : \Omega \ni x \mapsto df(x) \in L(E, F)$$

est continue de  $\Omega$  dans l'espace vectoriel normé  $L(E, F)$ .

**Exercices.** 1) Une application affine  $A : E \rightarrow F$  est de classe  $C^1$

2) La composée de deux applications de classe  $C^1$  est de classe  $C^1$ .

## Applications de classe $C^1$ et dérivées partielles.

Le produit  $E_1 \times \cdots \times E_k$  d'un nombre fini d'espaces vectoriels normés peut être muni de diverses normes qui, toutes, définissent la topologie produit, par exemple les normes  $l_p : \|(x_1, \dots, x_k)\| = (\|x_1\|^p + \cdots + \|x_k\|^p)^{\frac{1}{p}}$ . Ce qui suit ne dépend pas de la norme choisie. Rappelons qu'on a défini dans le résumé de cours n°4 les dérivées partielles  $\frac{\partial f}{\partial E_i}$  de  $f : E_1 \times \cdots \times E_k \supset \Omega \rightarrow F$ .

**Proposition 2** L'application  $f : E_1 \times \cdots \times E_k \supset \Omega \rightarrow F$  est de classe  $C^1$  si et seulement si les applications  $\frac{\partial f}{\partial E_i} : \Omega \rightarrow L(E_i, F)$  existent et sont continues.

**Le théorème fondamental du calcul différentiel et intégral** Si  $F$  un espace de Banach, c'est-à-dire un espace vectoriel normé complet (par exemple un espace vectoriel de dimension finie), on peut définir l'intégrale de Riemann  $\int_a^b \varphi(s) ds \in F$  d'une application continue  $\varphi : \mathbb{R} \supset [a, b] \rightarrow F$ .

On commence par définir de façon évidente l'intégrale d'une fonction en escalier (ou étagée), c'est-à-dire constante sur chaque intervalle d'une subdivision finie de l'intervalle  $[a, b]$  ; on montre ensuite que chaque application continue  $\varphi : [a, b] \rightarrow F$  est limite uniforme de fonctions en escalier  $E_n$ ,  $n \rightarrow \infty$ . Enfin, et c'est là que  $F$  doit être complet, on montre que pour chaque telle suite  $E_n$ , l'intégrale de  $E_n$  converge vers une limite et que cette limite (l'intégrale de  $\varphi$ ) ne dépend pas du choix de la suite.

Si  $\varphi(t) = (\varphi_1(t), \dots, \varphi_p(t)) \in F = \mathbb{R}^p$ ,  $\int_a^b \varphi(s) ds = \left( \int_a^b \varphi_1(s) ds, \dots, \int_a^b \varphi_p(s) ds \right)$ .

**Theorem 3** Soit  $F$  un espace de Banach,  $\varphi : [a, b] \rightarrow F$  une application continue. l'application  $\Phi : [a, b] \rightarrow F$  définie par  $\Phi(t) = \int_a^t \varphi(s) ds$  est de classe  $C^1$

et, si l'on identifie  $L(\mathbb{R}, F)$  à  $F$ , sa dérivée est l'application  $\varphi$ . Réciproquement, si  $\Phi : [a, b] \rightarrow F$  est une application de classe  $C^1$ , dont la dérivée en  $t$  est  $\Phi'(t) \in F \equiv L(\mathbb{R}, F)$ , on a  $\Phi(b) - \Phi(a) = \int_a^b \Phi'(s) ds \in F$ .

**Exercice.** Montrer que ceci vaut encore si l'on remplace [continu] et  $[C^1]$  respectivement par [continu par morceaux] et [continu, et  $C^1$  par morceaux].

**Formule de la moyenne et inégalité des accroissements finis.**

Soient  $E$  et  $F$  des espaces vectoriels normés,  $\Omega \subset E$  un ouvert. On suppose que  $F$  est complet afin que l'intégrale de Riemann d'une fonction à valeurs dans  $F$  soit définie (dans les applications, si  $F$  n'est pas complet, on pourra le remplacer par son complété, ce qui le plus souvent ne changera pas les énoncés).

Soient  $\gamma : [\alpha, \beta] \rightarrow \Omega$  et  $f : \Omega \rightarrow F$  des applications de classe  $C^1$  ( $\gamma$  peut n'être que continu et  $C^1$  par morceaux). On note  $a = \gamma(\alpha)$ ,  $b = \gamma(\beta)$ . Appliqué à  $f \circ \gamma$ , le théorème ci-dessus devient la

**Corollary 4** (Formule de la moyenne)  $\int_{\alpha}^{\beta} df(\gamma(t))\gamma'(t)dt = f(b) - f(a)$ .

Comme d'habitude, on a identifié la dérivée  $\gamma'(t)$  à un élément de  $E$ . Si  $\Omega$  contient le segment d'extrémités  $a$  et  $b$ , par exemple si  $\Omega$  est convexe, on peut choisir  $\gamma(t) = a + t(b - a)$ ; la formule se simplifie puisqu'alors  $\gamma'(t) = b - a \in E$ . Rappelons que, si  $F$  est complet, il en est de même de  $L(E, F)$ ; ceci permet de définir l'intégrale d'une application continue de  $[\alpha, \beta]$  dans  $L(E, F)$  et donc d'écrire :

$$\left( \int_{\alpha}^{\beta} df(a + t(b - a))dt \right) (b - a) = f(b) - f(a).$$

**Corollary 5** (Inégalité des accroissements finis) Soit  $f : E \supset \Omega \rightarrow F$  une application de classe  $C^1$  d'un ouvert d'un espace vectoriel normé à valeurs dans un espace de Banach. Soit  $K \subset \Omega$  un compact et  $M_K = \sup_{x \in K} \|df(x)\|$  (qui existe car  $f$  est  $C^1$ ). Quel que soit le chemin  $\gamma : [\alpha, \beta] \rightarrow K \subset \Omega$  continu et  $C^1$  par morceaux d'extrémités  $a, b \in K$ , on a  $\|f(b) - f(a)\| \leq M_K l(\gamma)$ , où  $l(\gamma) = \int_{\alpha}^{\beta} \|\gamma'(t)\| dt$  est la longueur de  $\gamma$ . Si de plus le segment d'extrémités  $a$  et  $b$  est tout entier contenu dans  $K$ , on a  $\|f(b) - f(a)\| \leq M_K \|b - a\|$ .

**Définition 6** (application lipschitzienne) Soient  $(X, d_X)$ ,  $(Y, d_Y)$  des espaces métriques. Une application  $f : X \rightarrow Y$  est dite lipschitzienne s'il existe  $M \geq 0$  telle que

$$\forall x_1, x_2 \in X, d_Y(f(x_1), f(x_2)) \leq M d_X(x_1, x_2).$$

Le inf des constantes  $M$  vérifiant ceci est appelée la constante de Lipschitz.

**Exercices. 1)** Soit  $f : E \supset \Omega \rightarrow F$  une application dont la dérivée est nulle en tout point de  $\Omega$ . Si  $\Omega$  est connexe,  $f$  est constante.

**2)** Soit  $f : E \supset \Omega \rightarrow F$  une application de classe  $C^1$ . Montrer que  $f$  est localement lipschitzienne (i.e. lipschitzienne en restriction à une petite boule centrée sur n'importe quel point).

**3)** Donner un exemple de  $f : [\alpha, \beta] \rightarrow \mathbb{R}^2$  de classe  $C^1$  telle qu'il n'existe aucun  $x \in [\alpha, \beta]$  vérifiant  $df(x)(\beta - \alpha) = f(\beta) - f(\alpha)$ .

# Calcul différentiel et équations différentielles

A. Chenciner, résumé de cours n<sup>07</sup>

## Le cas “facile” du théorème de Sard.

Le théorème suivant, cas particulier d’un théorème qui joue rôle très important en théorie des singularités des applications différentiables, met l’accent sur une différence fondamentale entre applications continues quelconques et applications de classes  $C^1$ .

**Définition 1** *Un sous-ensemble  $A$  de  $\mathbb{R}^p$  est dit de mesure de Lebesgue nulle si, quel que soit  $\epsilon > 0$ , on peut recouvrir  $A$  par une famille de boules (ou de cubes) dont la somme des volumes ( $p$ -dimensionnel) est majorée par  $\epsilon$ .*

**Theorem 2** *(le théorème de Sard “facile”). Si  $f : \mathbb{R}^n \supset \Omega \rightarrow \mathbb{R}^p$  est une application localement lipschitzienne (par exemple une application de classe  $C^1$ ) et si  $n < p$ , l’image  $f(\Omega)$  est de mesure de Lebesgue nulle.*

**Idée de démonstration :** 1) on commence par remarquer que, parcequ’il y a des suites convergentes, une union dénombrable d’ensembles de mesure nulle est de mesure nulle. Il suffit donc de montrer que l’image d’un cube  $K$  est de mesure nulle.  $K$  étant compact, l’hypothèse implique l’existence de  $M \geq 0$  telle que, quels que soient  $x, y \in K$ , on ait  $\|f(x) - f(y)\| \leq M\|x - y\|$ .

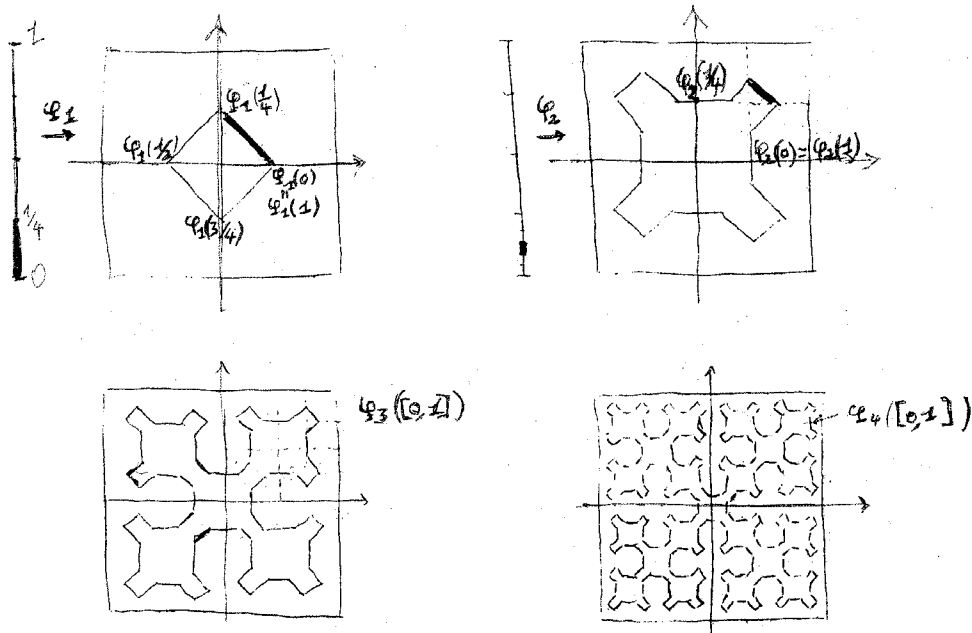
2) On découpe  $K$  (dont on appelle  $l$  la longueur des côtés) en  $N^n$  cubes  $K_i$  de longueur de côté  $l_i = \frac{l}{N}$ . Si  $x, y \in K_i$ , on a  $\|x - y\| \leq \sqrt{n} \frac{l}{N}$  (grande diagonale de  $K_i$ ) et donc  $\|f(x) - f(y)\| \leq M\sqrt{n} \frac{l}{N}$ . Ceci implique que l’image  $f(K_i)$  est contenue dans un cube de  $\mathbb{R}^p$  de côté  $2M\sqrt{n} \frac{l}{N}$  et donc de volume  $(2M\sqrt{n} \frac{l}{N})^p$ .

3) Finalement,  $f(K)$  est contenu dans une réunion de  $N^n$  cubes dont la somme des volumes est majorée par  $(2M\sqrt{nl})^p N^{n-p}$ , qui tend vers zéro lorsque  $N$  tend vers l’infini.

**Contemplation.** Dans les exemples dessinés dans le résumé de cours n<sup>03</sup>, regarder les applications auxquelles s’applique ce théorème et celles auxquelles il ne s’applique pas.

**Courbes de Peano** Le théorème ci-dessus devient faux pour les applications continues ! ! ! Il existe des applications continues surjectives d’un intervalle sur un carré. Les figures dessinées au dos indiquent la construction (en un nombre infini d’étapes) d’une telle *courbe de Peano*. La véritable raison de l’existence de ces courbes est de nature topologique : l’ensemble triadique de Cantor  $C$  est homéomorphe à son carré  $C \times C$  et il existe une application continue surjective de  $C$  sur un intervalle.

L'application continue surjective  $f : [0, 1] \rightarrow [0, 1] \times [0, 1]$  est obtenue comme limite uniforme d'applications continues  $\varphi_n$  dont les 4 premières sont représentées sur la figure ci-dessous.



Nous terminons cette feuille par un critère pour qu'une suite d'applications de classe  $C^1$  soit de classe  $C^1$  :

**Theorem 3** (suites d'applications de classe  $C^1$ ) Soit  $\{g_n : E \supset \Omega \rightarrow F\}_{n \in \mathbb{N}}$  une suite d'applications de classe  $C^1$  telle que

- 1)  $\exists x_0 \in \Omega, \exists \lambda \in F, \lim_{n \rightarrow \infty} g_n(x_0) = \lambda$  ;
- 2) La suite  $\{dg_n\}_{n \in \mathbb{N}}$  d'applications de  $\Omega$  dans  $L(E, F)$  converge localement uniformément vers  $\alpha : \Omega \rightarrow L(E, F)$ . Alors  $g_n$  tend localement uniformément vers une application  $g$  de classe  $C^1$  telle que  $dg = \alpha$ .

# Calcul différentiel et équations différentielles

A. Chenciner, résumé de cours n°8

## Espaces tangents.

Lorsqu'il existe, le sous-espace tangent en  $P$  au sous-ensemble  $\Sigma$  de l'espace vectoriel normé  $E$  est noté  $T_P\Sigma$ . Sous-espace affine de  $E$  avec un point privilégié  $P$ , il est donc muni de la structure vectorielle canonique dont l'origine est  $P$ .

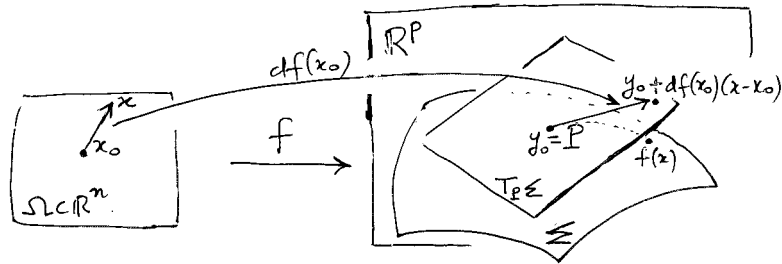
*Notations.*  $f : \mathbb{R}^n \supset \Omega \rightarrow \mathbb{R}^p$  est dérivable en  $x_0 \in \mathbb{R}^n$  et  $y_0 = f(x_0) \in \mathbb{R}^p$ .

**Cas 1)  $\Sigma = f(\Omega) \subset \mathbb{R}^p$  est défini par le paramétrage  $f$  de rang maximum (rang  $f = n \leq p$ ).**

$$\Sigma = \{(y \in \mathbb{R}^p, \exists x \in \Omega, y = f(x))\}, P = y_0.$$

$$T_P\Sigma = \{y \in \mathbb{R}^p, \exists x \in \mathbb{R}^n, y = f(x_0) + df(x_0)(x - x_0)\}.$$

$T_P\Sigma$  est un sous-espace affine de  $E = \mathbb{R}^p$  de dimension  $n = \text{rang} f$ .

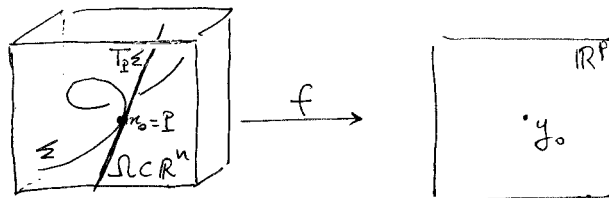


**Cas 2)  $\Sigma = f^{-1}(f(x_0)) \subset \Omega$  est défini par les équations  $f$  de rang maximum (rang  $f = p \leq n$ ).**

$$\Sigma = \{x \in \Omega, f(x) = f(x_0)\}, P = x_0.$$

$$T_P\Sigma = \{x \in \mathbb{R}^n, f(x_0) + df(x_0)(x - x_0) = f(x_0)\}.$$

$T_P\Sigma$  est un sous-espace affine de  $E = \mathbb{R}^n$  de dimension  $n - p = \dim \text{Ker} f$ .



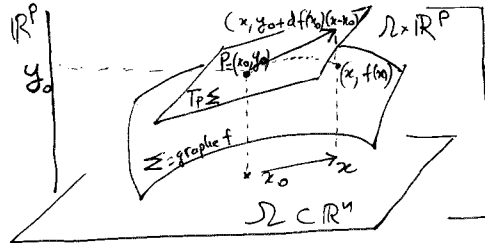
**Cas 3)**  $\Sigma \subset \Omega \times \mathbb{R}^p$  est défini comme le graphe de  $f$ .

$$\Sigma = \{(x, y) \in \Omega \times \mathbb{R}^p, y = f(x)\}, P = (x_0, f(x_0)) \in \Sigma,$$

$$T_P \Sigma = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^p, y = f(x_0) + df(x_0)(x - x_0)\}.$$

$T_P \Sigma$  est un sous-espace affine de  $E = \mathbb{R}^n \times \mathbb{R}^p$  de dimension  $n$ .

Ce dernier cas relève aussi bien du cas 1) (paramétrage  $x \mapsto (x, f(x))$ ) que du cas 2) (équations  $y - f(x) = 0$ ).



**Exercices.** 1) Appliquer ce qui précède à la recherche de la tangente en un point à une courbe plane. Donner divers exemples de *points singuliers* de tels courbes (points en lesquels la condition de rang maximum n'est pas vérifiée).

2) Considérer le cas d'une courbe  $\Sigma$  définie dans  $\mathbb{R}^3$  par deux équations  $f_1 : \mathbb{R}^3 \rightarrow \mathbb{R}$  et  $f_2 : \mathbb{R}^3 \rightarrow \mathbb{R}$ . Montrer que la condition de rang équivaut aux deux conditions suivantes : (i) l'existence de plans tangent au point  $P$  considéré aux deux surfaces d'équation  $f_1 = 0$  et  $f_2 = 0$  ; (ii) la *transversalité* de ces deux plans (i.e. le fait qu'ils se coupent suivant une droite, autrement dit, qu'ils ne sont pas confondus). Montrer que la tangente à la courbe au point  $P$  est l'intersection de ces deux plans.

**RÈGLE À RETENIR :** Pour passer d'un objet géométrique  $\Sigma \subset E$  (courbe, surface, etc) à son espace (droite, plan, etc) tangent  $T_P \Sigma$  en un point  $P$ , on remplace l'application  $f : \mathbb{R}^n \rightarrow \mathbb{R}^p$  qui le définit par la partie affine de son développement de Taylor :

$$x \mapsto f(x_0) + df(x_0)(x - x_0).$$

Cette règle s'applique dans tous les cas, que  $\Sigma$  soit défini par un paramétrage de rang maximum ( $E = \mathbb{R}^p$ ,  $P = f(x_0)$ ), des équations de rang maximum ( $E = \mathbb{R}^n$ ,  $P = x_0$ ) ou comme un graphe ( $E = \mathbb{R}^n \times \mathbb{R}^p$ ,  $P = (x_0, f(x_0))$ ).



# Calcul différentiel et équations différentielles

A. Chenciner, résumé de cours n°9

## Gradients.

La définition classique du gradient d'une fonction  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  dérivable en  $x_0$  ne met pas suffisamment en relief le choix implicite du produit scalaire euclidien de  $\mathbb{R}^n$  pour transformer le covecteur  $df(x_0)$  en le vecteur  $\text{grad}f(x_0)$ .

Une bonne façon de penser à une forme bilinéaire  $\mathcal{B} : E \times E \rightarrow \mathbb{R}$  non dégénérée sur un espace vectoriel (nous nous limiterons à la dimension finie) est de l'interpréter comme l'isomorphisme  $\varphi_{\mathcal{B}}$  de  $E$  sur le dual  $E^* := L(E, \mathbb{R})$  qui envoie  $X \in E$  sur la forme linéaire

$$E \ni Y \mapsto \varphi_{\mathcal{B}}(X)(Y) = \mathcal{B}(X, Y) \in \mathbb{R}.$$

Si par le choix d'une base on identifie  $E$  à  $\mathbb{R}^n$ , la forme bilinéaire  $\mathcal{B}$  est représentée par une matrice symétrique  $B$ , l'isomorphisme  $\varphi_{\mathcal{B}}$  s'écrit

$$\varphi_{\mathcal{B}}(X)(Y) = \langle X, BY \rangle.$$

*Noter que dans cette représentation intervient le produit scalaire euclidien canonique  $\langle, \rangle$  de  $\mathbb{R}^n$  !!!*

**Définition 1** Soit  $f : E \rightarrow \mathbb{R}$  une fonction dérivable en  $x_0$  et soit  $\mathcal{B}$  un produit scalaire (i.e. une forme bilinéaire définie positive) sur  $E$ . Le gradient  $\text{grad}_{\mathcal{B}}f(x_0) \in E$  est défini par

$$\text{grad}_{\mathcal{B}}f(x_0) = \varphi_{\mathcal{B}}^{-1}(df(x_0)),$$

*c'est-à-dire*

$$\forall Y \in E, \mathcal{B}(\text{grad}_{\mathcal{B}}f(x_0), Y) = df(x_0)(Y).$$

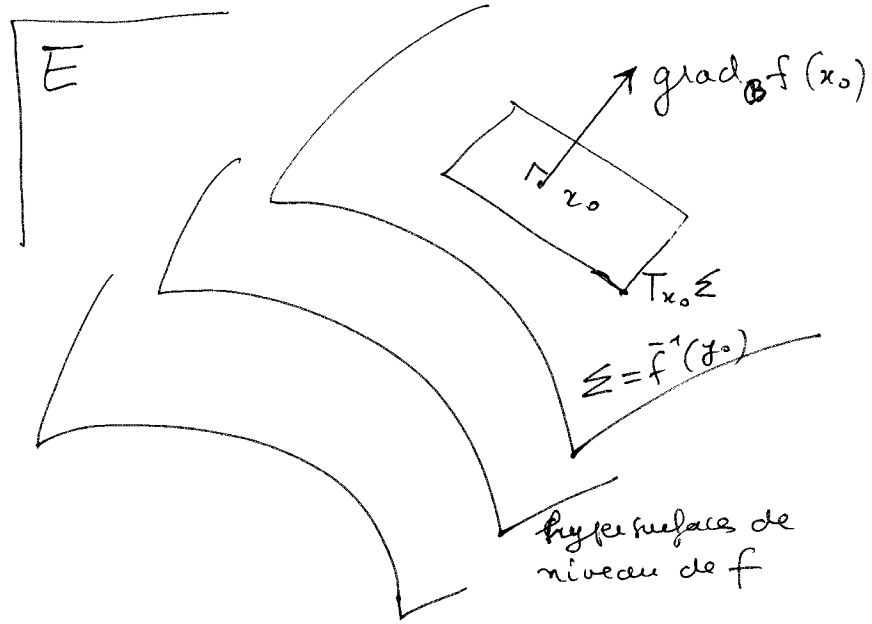
Si une base de  $E$  est choisie dans laquelle  $\mathcal{B}$  est représentée par la matrice symétrique définie positive  $B$ , la définition devient :

$$\forall Y \in \mathbb{R}^n, \langle \text{grad}_B f(x_0), BY \rangle = df(x_0).$$

Enfin, si  $B = Id$ , c'est-à-dire si  $\mathcal{B}$  est le produit scalaire canonique de  $\mathbb{R}^n$ , on retrouve la définition élémentaire :

$$\text{grad}_{Id}f(x_0) = \left( \frac{\partial f}{\partial x_1}(x_0), \dots, \frac{\partial f}{\partial x_n}(x_0) \right).$$

La figure suivante fait le lien avec la notion de plan tangent en un point  $x_0$  d'une hypersurface  $\Sigma = f^{-1}(y_0)$  définie dans un espace euclidien de dimension finie  $(E, \mathcal{B})$  par une équation de rang maximum  $f : (E, x_0) \cdots \rightarrow (\mathbb{R}, y_0)$ . La condition de rang maximum équivaut à la non annulation de  $\text{grad}_{\mathcal{B}} f(x_0)$  et ce dernier est orthogonal (pour la structure euclidienne  $\mathcal{B}$ ) au plan tangent  $T_{x_0}\Sigma$ .

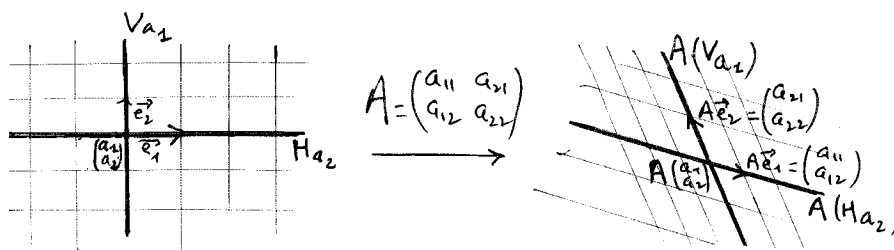


# Calcul différentiel et équations différentielles

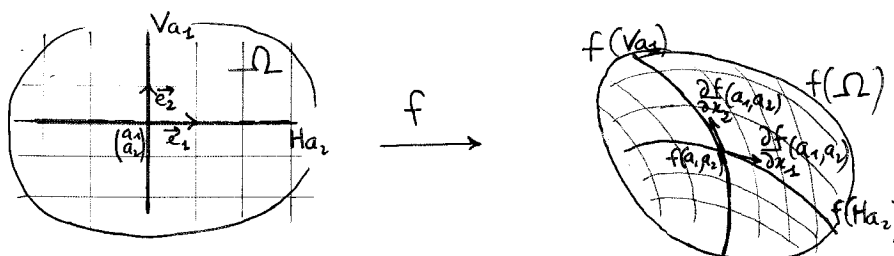
A. Chenciner, résumé de cours n°10

## Le théorème d'inversion locale

De quoi s'agit-il ? Du passage de la figure 1, *globale*, qui représente un isomorphisme linéaire de  $\mathbb{R}^2$  dans  $\mathbb{R}^2$ ...



... à la figure 2, *locale*, qui représente une application  $f : \mathbb{R}^2 \supset \Omega \rightarrow \mathbb{R}^2$  de classe  $C^1$  dont la dérivée en tout point de  $\Omega$  est inversible.



Les images par  $f$  de l'intersection avec  $\Omega$  des droites parallèles aux axes de coordonnées définissent un *système de coordonnées curvilignes* dans  $f(\Omega)$  i.e. deux familles de courbes régulières se coupant *transversalement*. En particulier, la restriction de  $f$  à  $\Omega$  est une bijection de  $\Omega$  sur  $f(\Omega)$

Si la dérivée de  $f$  est un isomorphisme en un point, elle l'est également en tous les points d'un voisinage : en effet,  $df(x_1, x_2)$  dépend continuellement de  $(x_1, x_2)$  et le groupe  $GL(\mathbb{R}^2)$  des isomorphismes linéaires de  $\mathbb{R}^2$  est ouvert dans  $L(\mathbb{R}^2, \mathbb{R}^2)$  : la non-annulation d'un déterminant est en effet une *condition ouverte*. Cette dernière propriété est encore vraie si  $\mathbb{R}^2$  est remplacé par  $\mathbb{R}^n$  ou même par un espace de Banach  $E$  : le groupe  $GL(E)$  des isomorphismes linéaires continus est ouvert dans  $L(E, E)$  ( car  $(Id - u)^{-1} = Id + u + u^2 + u^3 + \dots$ , qui converge si  $\|u\|$  est assez petit..)

C'est ce qui explique que, même en dimension infinie, le théorème d'inversion locale se contente d'une hypothèse d'inversibilité de la dérivée en un point.

**Exercice.** Comprendre ce qu'est dans  $\mathbb{R}^3$  l'analogie des figures ci-dessus.

**Définition 1** Soient  $E$  et  $F$  des espaces vectoriels normés,  $\Omega$  un ouvert de  $E$ ,  $\mathcal{O}$  un ouvert de  $F$ . Un difféomorphisme de classe  $C^1$  de  $\Omega$  sur  $\mathcal{O}$  est une application  $f : \Omega \rightarrow \mathcal{O}$  de classe  $C^1$  qui est bijective et dont la bijection réciproque  $f^{-1}$  est également de classe  $C^1$ . Si  $x_0 \in E$ , on dit que  $f : (E, x_0) \cdots \rightarrow (F, y_0)$  est un difféomorphisme local (en  $x_0$ ) s'il existe des voisinages ouverts  $\Omega$  de  $x_0$  et  $\mathcal{O}$  de  $y_0$  tels que  $f$  soit un difféomorphisme de  $\Omega$  sur  $\mathcal{O}$ . Lorsqu'on remplace  $C^1$  par  $C^0$  (i.e. continue), on parle d'homéomorphisme ou d'homéomorphisme local. Lorsqu'on remplace  $C^1$  par localement lipschitzien, on parle d'isomorphisme de Lipschitz ou d'isomorphisme local de Lipschitz.

**Theorem 2** Soient  $E$  et  $F$  des espaces de Banach,  $f : (E, x_0) \cdots \rightarrow (F, y_0)$  une application de classe  $C^1$ . Si  $df(x_0) \in L(E, F)$  est inversible,  $f$  est un difféomorphisme local de classe  $C^1$ .

Une bonne démonstration se trouve dans le livre de Chaperon (pages 230-233) ; en voici le plan :

L'outil de base est le théorème des contractions (= applications lipschitziennes de constante de Lipschitz strictement inférieure à 1) :

**Theorem 3** Une contraction  $u : X \rightarrow X$  d'un espace métrique complet  $X$  possède un unique point fixe  $x_0$ . De plus, quel que soit  $x \in X$ , la suite des itérés  $u^n(x)$  converge vers  $x_0$ .

**Lemma 4** Si  $u$  est une contraction de l'espace de Banach  $E$ ,  $Id - u$  est un homéomorphisme lipschitzien ainsi que son inverse.

**Lemma 5** Soit  $f : (E, x_0) \rightarrow (F, y_0)$  un isomorphisme local de Lipschitz entre espaces de Banach. Si  $g : (E, x_0) \rightarrow (F, z_0)$  est une application locale lipschitzienne dont la restriction à la boule  $B(x_0, r)$  a une constante de Lipschitz qui tend vers 0 avec le rayon  $r$ ,  $f + g : (E, x_0) \rightarrow (F, y_0 + z_0)$  est un isomorphisme local de Lipschitz.

**Lemma 6** Soit  $h : E \supset \Omega \rightarrow \mathcal{O} \subset F$  un isomorphisme de Lipschitz. Si  $h$  est dérivable en  $x_0 \in \Omega$ ,  $dh(x_0)$  est un isomorphisme,  $h^{-1}$  est dérivable en  $y_0 = h(x_0)$  et  $d(h^{-1})(y_0) = dh(x_0)^{-1}$ .

# Calcul différentiel et équations différentielles

A. Chenciner, résumé de cours n<sup>o</sup>11

Dans cette page,  $f : (\mathbb{R}^n, x_0) \cdots \rightarrow (\mathbb{R}^p, y_0)$  est une application de classe  $C^1$  au voisinage de  $x_0$  ; on note  $T_{x_0}^1 f$  son polynôme de Taylor à l'ordre 1 en  $x_0$  :

$$T_{x_0}^1 f(x) = y_0 + df(x_0)(x - x_0).$$

## Immersion, submersion, applications de rang constant

**Définition 1** Si  $df(x_0) \in L(\mathbb{R}^n, \mathbb{R}^p)$  est injective (resp. surjective), on dit que  $f$  est une immersion (resp. submersion) en  $x_0$ .

Pour une immersion ou une submersion,  $df(x_0)$  est de rang maximum, i.e. de rang égal à  $\inf(n, p)$ . Puisque  $f$  est  $C^1$  et que cette propriété est ouverte, elle est encore vérifiée par  $df(x)$  si  $x$  est assez voisin de  $x_0$ . Immersion et submersion sont donc les cas particuliers extrêmes d'applications de rang constant, i.e. d'applications telles que le rang de  $df(x)$  soit constant sur un voisinage de  $x_0$ .

**Théorème 2** (i) si  $f$  est une immersion en  $x_0$ , il existe un difféomorphisme local  $\psi : (\mathbb{R}^p, y_0) \cdots \rightarrow (\mathbb{R}^p, y_0)$  de classe  $C^1$  tel qu'au voisinage de  $x_0$ , on ait  $f = \psi \circ T_{x_0}^1 f$ . (ii) Si  $f$  est une submersion en  $x_0$ , il existe un difféomorphisme local  $\varphi : (\mathbb{R}^n, x_0) \cdots \rightarrow (\mathbb{R}^n, x_0)$  de classe  $C^1$  tel qu'au voisinage de  $x_0$ , on ait  $f = T_{x_0}^1 f \circ \varphi$ . (iii) Si  $f$  est de rang constant au voisinage de  $x_0$ , il existe des difféomorphismes locaux  $\varphi : (\mathbb{R}^n, x_0) \cdots \rightarrow (\mathbb{R}^n, x_0)$ ,  $\psi : (\mathbb{R}^p, y_0) \cdots \rightarrow (\mathbb{R}^p, y_0)$  de classe  $C^1$ , tels qu'au voisinage de  $x_0$ , on ait  $f = \psi \circ T_{x_0}^1 f \circ \varphi$ .

**Exercice.** Compte-tenu du théorème de structure des applications linéaires rappelé dans la feuille 2 et de l'existence de translations ramenant respectivement  $x_0$  et  $y_0$  à l'origine, on peut énoncer ce théorème sous la forme suivante :

Si  $f$  est une immersion en  $x_0$  (resp. une submersion en  $x_0$ , resp. de rang constant  $r$  au voisinage de  $x_0$ ), il existe des coordonnées locales à la source et au but, telles que  $f$  devienne l'injection canonique  $i$  de  $\mathbb{R}^n$  dans  $\mathbb{R}^p$  (resp. la surjection canonique  $\pi$ , resp. la composée  $i \circ \pi$  de l'injection canonique de  $\mathbb{R}^n$  dans  $\mathbb{R}^p$  avec la surjection canonique  $\pi$  de  $\mathbb{R}^n$  sur  $\mathbb{R}^r$ ).

L'injection canonique  $i$  de  $\mathbb{R}^a$  dans  $\mathbb{R}^{a+b}$  et la projection canonique  $\pi$  de  $\mathbb{R}^{a+b}$  dans  $\mathbb{R}^a$  sont respectivement définies par

$$i(x_1, \dots, x_a) = (x_1, \dots, x_a, 0, \dots, 0), \quad \pi(x_1, \dots, x_{a+b}) = (x_1, \dots, x_a).$$

Les démonstrations de (i) et (ii) consistent, par ajout de variables respectivement au but et à la source en la transformation de  $f$  en une application  $F$  à laquelle s'applique le théorème d'inversion locale.

**Généralisation à la dimension infinie.** Une hypothèse fondamentale est nécessaire pour que la caractérisation ci-dessus des immersions et des submersions se généralise au cas d'espaces de Banach de dimension infinie : on doit en plus de l'hypothèse que  $df(x_0)$  est injective (resp. surjective), supposer que son image (resp. son noyau) est *facteur direct* dans  $\mathbb{R}^p$  (resp. dans  $\mathbb{R}^n$ ). On dit alors que  $df(x_0)$  est une *immersion linéaire* (resp. une *submersion linéaire*).

### Le théorème des fonctions implicites

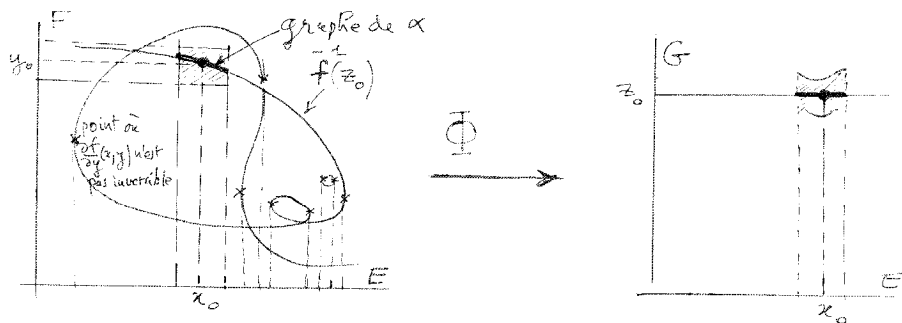
**Théorème 3** Soient  $E, F, G$  des Banach,  $f : (E \times F, (x_0, y_0)) \cdots \rightarrow (G, z_0)$  une application de classe  $C^1$  au voisinage de  $(x_0, y_0)$ . Si  $\frac{\partial f}{\partial y}(x_0, y_0) \in L(F, G)$  est un isomorphisme, il existe  $\alpha : (E, x_0) \cdots \rightarrow (F, y_0)$  de classe  $C^1$  telle que, dans un voisinage de  $(x_0, y_0)$ , le graphe de  $\alpha$  coïncide avec l'ensemble des solutions  $(x, y)$  de l'équation  $f(x, y) = z_0$ .

Au voisinage de la solution  $(x_0, y_0)$ , on a remplacé la relation *implicite* (ou l'équation)  $f(x, y) = z_0$  par la relation *explicite*  $y = \alpha(x)$ .

*Démonstration.* C'est un corollaire immédiat du théorème d'inversion locale : l'hypothèse assure que l'application

$$\Phi : (E \times F, (x_0, y_0)) \cdots \rightarrow (E \times G, (x_0, z_0)), \quad \Phi(x, y) = (x, f(x, y))$$

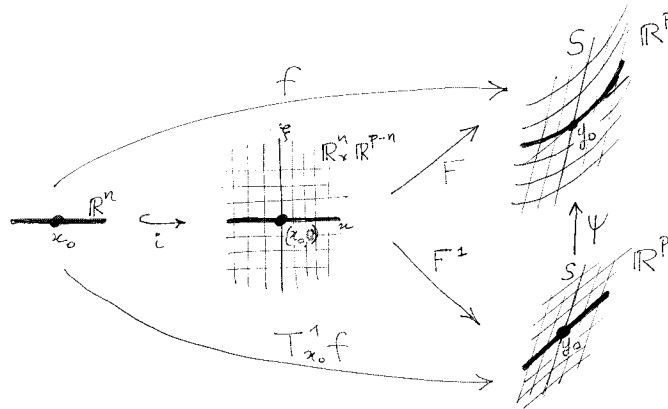
est un difféomorphisme local  $C^1$ . Il suffit alors de définir  $\alpha$  par la relation  $(x, \alpha(x)) = \Phi^{-1}(x, z_0)$ .



# Calcul différentiel et équations différentielles

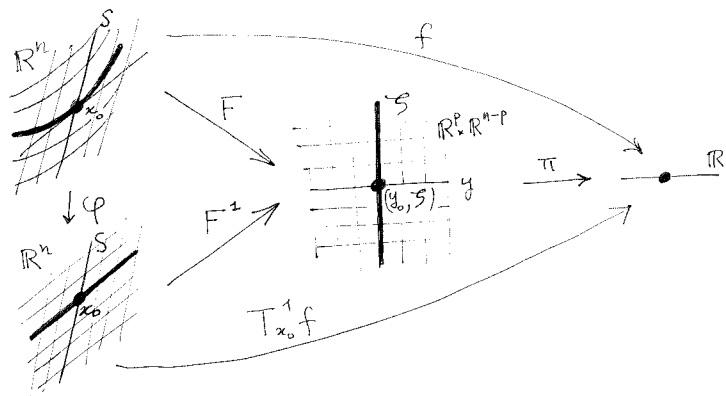
A. Chenciner, résumé de cours n°11bis

Démonstration du théorème des Immersions :  $f = \psi \circ T_{x_0}^1 f = F \circ i$ .



$F(x, \xi) = f(x) + g(\xi)$ , où  $g : \mathbb{R}^{p-n} \rightarrow \mathbb{R}^p$  est une immersion linéaire telle que  $\text{Im}g \oplus \text{Ker}df(x_0) = \mathbb{R}^p$  (on note  $S = y_0 + \text{Im}g$ ). On passe de  $F$  à  $F^1$  en remplaçant  $f$  par  $T_{x_0}^1 f$  et on pose  $\psi = F \circ (F^1)^{-1}$ . On a  $d\psi(y_0) = \text{Id}$ .

Démonstration du théorème des submersions :  $f = T_{x_0}^1 f \circ \varphi = \pi \circ F$ .



$F(x) = (f(x), g(x))$ , où  $g : \mathbb{R}^n \rightarrow \mathbb{R}^{n-p}$  est une submersion linéaire telle que  $\text{Ker}g \oplus \text{Ker}df(x_0) = \mathbb{R}^n$  (on note  $S = x_0 + \text{Ker}g$ ). On passe de  $F$  à  $F^1$  en remplaçant  $f$  par  $T_{x_0}^1 f$  et on pose  $\varphi = (F^1)^{-1} \circ F$ . On a  $d\varphi(x_0) = \text{Id}$ .

# Calcul différentiel et équations différentielles

A. Chenciner, résumé de cours n°12

**Sous-variétés de classe  $C^1$ .** Une sous-variété de classe  $C^1$  d'un Banach  $E$  est un sous-ensemble qui à *difféomorphisme local  $C^1$  près* ne se distingue pas d'un sous-espace vectoriel.

*ATTENTION :* quand nous parlerons d'un sous-espace vectoriel de l'espace de Banach  $E$ , il s'agira toujours d'un facteur direct, i.e. d'un sous-espace vectoriel  $E_1$  qui possède un supplémentaire  $E_2$ . C'est automatique en dimension finie et dans un Hilbert mais pas dans tous les Banach.

Plus précisément,

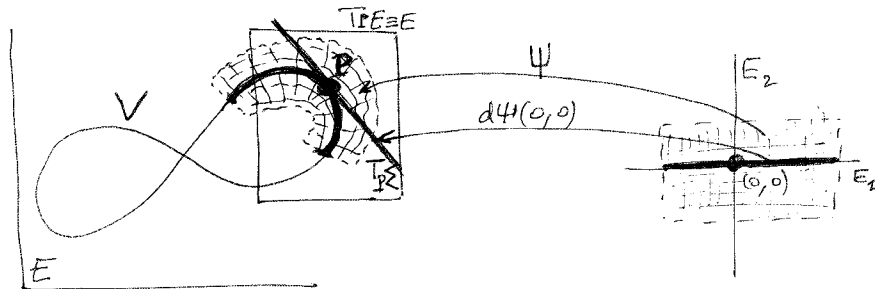
**Définition 1**  $V \subset E$  est une sous-variété de classe  $C^1$  au point  $P \in V$  s'il existe un *difféomorphisme local*  $\Psi : (E_1 \times E_2, (0,0)) \cdots \rightarrow (E, P)$  de classe  $C^1$  tel qu'au voisinage de  $P$ , on ait  $(V, P) = \Psi(E_1 \times \{0\}, (0,0))$ . On dira que  $\Psi$  est une "carte locale" de  $E$  en  $P$  "adaptée" à  $V$ .

Nous pouvons comprendre maintenant le cadre dans lequel valent les recettes données dans la feuille 8. La définition ci-dessous ne fait que généraliser celle donnée dans la feuille 5 :

**Définition 2** Si  $V \subset E$  est une sous-variété de classe  $C^1$  en  $P$ , on définit son *espace tangent en  $P$*  comme le sous-ensemble  $T_P V \subset T_P E$  des vecteurs-vitesse de chemins de classe  $C^1$  tracés sur  $V$ .

Puisque l'espace vectoriel  $T_P E$  s'identifie canoniquement à  $E$  par la translation de  $P$ ,  $T_P V$  peut donc être identifié à un sous-espace vectoriel de  $E$  ; cependant, si cette identification se fait naturellement lorsqu'on calcule, elle ne fait que troubler l'esprit quand on cherche à comprendre la géométrie.

**Lemme 3** Si  $\Psi : (E_1 \times E_2, (0,0)) \cdots \rightarrow (E, P)$  est une carte locale de  $E$  en  $P$  adaptée à  $V$ , on a  $T_P V = d\Psi(0,0)(E_1 \times \{0\})$ .

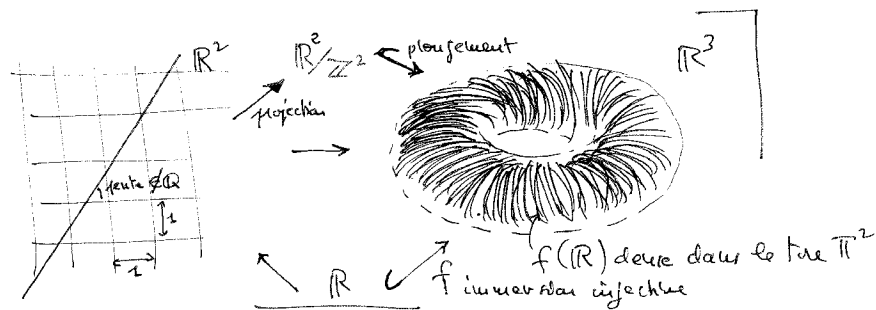




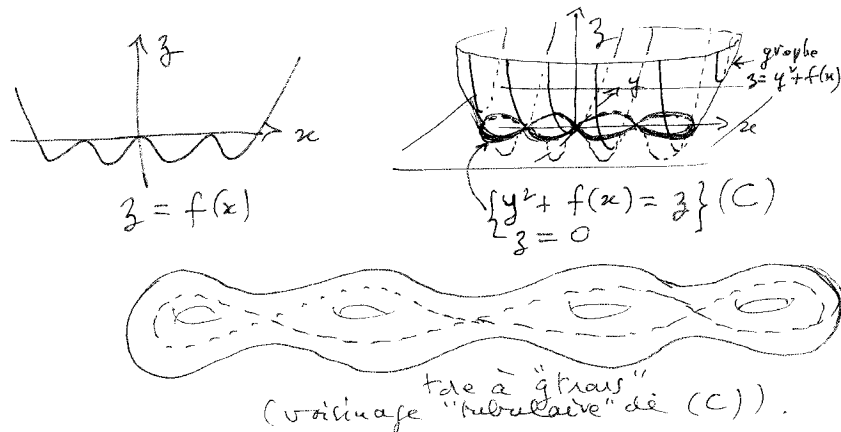
**Remarque.** On peut aussi (voir la feuille 8) considérer  $T_P V$  comme le sous-espace affine de  $E$  formé des  $P + \xi$ ,  $\xi \in T_P V$ , muni du point privilégié  $P$ .

**Exercices.** 1) Si  $f : E \supset \Omega \rightarrow F$  est une submersion en chacun des points de  $V = f^{-1}(y_0)$ ,  $V$  est une sous-variété de  $E$ . Par contre, si  $f : E \supset \Omega \rightarrow F$  est une immersion en chaque point de  $\Omega$ , chaque point de  $\Omega$  possède un voisinage dont l'image par  $f$  est une sous-variété de  $F$  mais l'image  $f(\Omega)$  elle-même n'est en général pas une sous-variété et ce, même si  $f$  est injective ! C'en est une si  $f$  est un homéomorphisme sur son image : on parle alors de *plongement* (exercice de topologie : une immersion propre est un plongement).

Des exemples intéressants d'immersions injectives de  $\mathbb{R}$  dans  $\mathbb{R}^3$  qui ne sont pas des plongements sont obtenus en composant une application affine de  $\mathbb{R}$  dans  $\mathbb{R}^2$  dont l'image est une droite de pente irrationnelle, avec la projection de  $\mathbb{R}^2$  sur le quotient  $\mathbb{R}^2/\mathbb{Z}^2$  par le réseau des entiers et enfin avec un plongement de  $\mathbb{R}^2/\mathbb{Z}^2$  dans  $\mathbb{R}^3$ .



2) Montrer en s'aidant de la figure ci-dessous qu'on peut définir une surface (=sous-variété de dimension 2) dans  $\mathbb{R}^3$  ayant la forme d'un *tore à  $g$  trous* (surface orientable de genre  $g$ ) par une submersion polynomiale  $P : \mathbb{R}^3 \rightarrow \mathbb{R}$  de la forme  $P(x, y, z) = (y^2 + f(x))^2 + z^2 = \epsilon^2$  avec  $\epsilon \neq 0$  assez petit.



# Calcul différentiel et équations différentielles

A. Chenciner, résumé de cours n°13

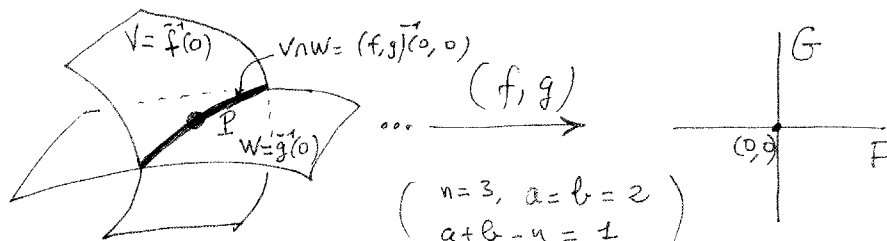
## Transversalité.

C'est la formalisation dans le cadre du calcul différentiel de la vieille notion de "position générale" : par exemple, deux plans affines choisis "au hasard" dans  $\mathbb{R}^3$  ne seront pas parallèles et se couperont suivant une droite, dans  $\mathbb{R}^4$  ils se couperont suivant un point alors que, dans  $\mathbb{R}^n$ , ils ne se rencontreront pas dès que  $n \geq 5$ . De plus, cette propriété résistera à une petite perturbation : il y a *stabilité* de l'intersection. Les théorèmes précédents impliquent une version non-linéaire de ce fait :

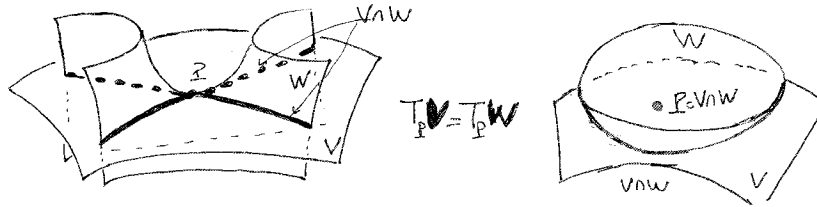
**Définition 1** Soit  $E$  un espace de Banach,  $V$  et  $W$  deux sous-variétés de  $E$  de classe  $C^1$  et  $P$  un point de  $V \cap W$ . On dit que  $V$  et  $W$  sont transverses en  $P$  si  $T_P V + T_P W = T_P E \cong E$ . On note  $V \pitchfork_P W$ .

**Proposition 2** Si deux sous-variétés  $V$  et  $W$  de  $E$  sont transverses au point  $P \in V \cap W$ , leur intersection  $V \cap W$  est une sous-variété de  $E$  au point  $P$ . Si  $E$  est de dimension finie  $n$  et  $V, W$  respectivement de dimension  $a$  et  $b$ ,  $V \cap W$  est, au voisinage de  $P$ , de dimension  $a + b - n$ .

La démonstration est basée sur la remarque suivante : si  $V \subset E$  est une sous-variété au point  $P$ , il existe une submersion  $f : (E, P) \cdots \rightarrow (F, 0)$  telle que  $(V, P) = f^{-1}(F, 0)$ . Pour le voir, il suffit de composer l'inverse  $\Psi^{-1} : (E, P) \cdots \rightarrow (E_1 \times E_2, (0, 0))$  d'une carte locale de  $E$  en  $P$  adaptée à  $V$  avec la projection canonique de  $E_1 \times E_2$  sur  $E_2$ . Si les submersions locales  $f : (E, P) \cdots \rightarrow (F, 0)$  et  $g : (E, P) \cdots \rightarrow (G, 0)$  définissent respectivement  $V$  et  $W$  au voisinage de  $P$ , l'hypothèse équivaut à ce que l'application  $(f, g) : (E, P) \cdots \rightarrow (F \times G, (0, 0))$  soit une submersion. La conclusion suit puisque  $V \cap W$  est définie au voisinage de  $P$  par l'équation  $(f, g) = (0, 0)$ .



Voici deux exemples de surfaces non transverses dans  $\mathbb{R}^3$ . Que se passe-t-il si l'on translate légèrement l'une d'elles dans la direction verticale ?



### Rudiments de calcul différentiel dans les sous-variétés.

A difféomorphisme local de classe  $C^1$  près, une sous-variété de  $E$  et un sous-espace vectoriel de  $E$  (en dimension infinie, il faut préciser “sous-espace vectoriel facteur direct”) sont indistinguables. Toutes les notions locales définies pour  $f : (E, x_0) \cdots \rightarrow (F, y_0)$  peuvent donc être étendues aux applications locales  $f : (V, P) \cdots \rightarrow (W, Q)$  d’une sous-variété de  $E$  dans une sous-variété de  $F$ .

**Définition 3** Soient  $V \subset E$ , sous-variété en  $P$  et  $W \subset F$ , sous-variété en  $Q$ . Une application  $f : (V, P) \cdots \rightarrow (W, Q)$  est dite dérivable en  $P$  (resp. de classe  $C^1$  en  $P$ ) s’il existe  $\Phi : (E_1 \times E_2, (0, 0)) \cdots \rightarrow (E, P)$  carte locale en  $P$  de  $E$  adaptée à  $V$  et  $\Psi : (F_1 \times F_2, (0, 0)) \cdots \rightarrow (F, Q)$  carte locale en  $Q$  de  $F$  adaptée à  $W$ , telles que  $F := \Psi^{-1} \circ f \circ \Phi|_{E_1} : (E_1, 0) \cdots \rightarrow (E_2, 0)$  soit dérivable en  $(0, 0)$  (resp. de classe  $C^1$  au voisinage de  $(0, 0)$ ). La même propriété est alors vraie quel que soit le couple de cartes choisi (exercice !).

**Définition 4** Si  $f : (V, P) \cdots \rightarrow (W, Q)$  est dérivable en  $P$ , sa dérivée, qui est une application linéaire  $df(P) : T_P V \rightarrow T_Q W$ , est définie par la formule suivante, elle aussi indépendante du choix des cartes  $\Phi$  et  $\Psi$  ( $F$  est définie dans la définition précédente) :

$$df(P)X = d\Psi(0, 0) \circ dF(0, 0) \circ (d\Phi(0, 0)|_{E_1 \times \{0\}})^{-1}X.$$

On peut alors définir immersions, submersions et difféomorphismes (locaux ou globaux) d’une sous-variété dans une autre et étendre à ce cadre les théorèmes locaux qui les concernent (feuilles 10 et 11).

### Exercices.

**1** Montrer que si  $f : (V, P) \cdots \rightarrow (W, Q)$  est de classe  $C^1$ , elle est la restriction à  $V$  d’une application  $\Phi : (E, P) \cdots \rightarrow (F, Q)$  de classe  $C^1$  d’image contenue dans  $W$ . Montrer que  $df(P)$  est la restriction de  $d\Phi(P)$  à  $T_P V \subset T_P E$ .

**2** Justifier les assertions de la feuille 8. Interpréter dans ce cadre les figures de la feuille 3 et celle à la fin de la feuille 5.

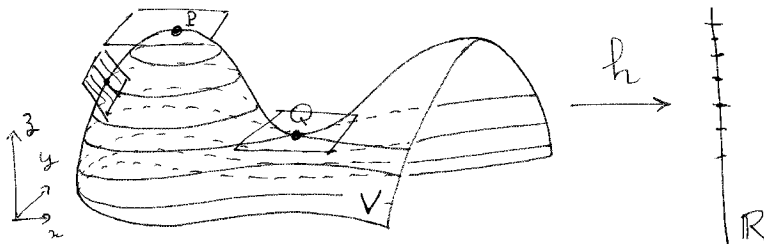
# Calcul différentiel et équations différentielles

A. Chenciner, résumé de cours n°14

## Extrema liés et multiplicateurs de Lagrange.

**Définition 1** Soit  $\Omega \subset E$  un ouvert d'un espace de Banach,  $V \subset \Omega$  une sous-variété de classe  $C^1$ ,  $f : \Omega \rightarrow \mathbb{R}$  une fonction de classe  $C^1$  à valeurs réelles. On appelle "point critique" de  $f|_V$  un point  $P$  en lequel s'annule la dérivée  $d(f|_V)(P) : T_P V \rightarrow T_{f(P)} \mathbb{R} \cong \mathbb{R}$  de la restriction  $f|_V : V \rightarrow \mathbb{R}$  de  $f$  à  $V$ .

Il y a beaucoup de raisons de s'intéresser aux points critiques : ce sont les *extrema* de  $f|_V$ , sur lesquels nous reviendrons dès que nous disposerons de la dérivée seconde. La figure ci-dessous montre les *courbes de niveau*  $h|_V^{-1}(a)$  de la restriction à une sous-variété  $V$  de  $\mathbb{R}^3$  de la *fonction hauteur* (ou *fonction altitude*)  $h(x, y, z) = z$ . On y voit deux points critiques : un *maximum local*  $P$  et un *col*  $Q$ , en lesquels le plan tangent à  $V$  est horizontal ; ceux qui aiment se promener en montagne penseront aux *cartes d'état-major* de l'IGN.



**Définition 2** Soit  $P$  un point critique de  $f|_V$ . Si  $V$  est localement définie par l'équation  $g = a$ , où  $g : (\Omega, P) \rightarrow (F, a)$  est une submersion de classe  $C^1$ , on dit encore que  $P$  est un *extremum de  $f$  lié par la contrainte  $g = a$* .

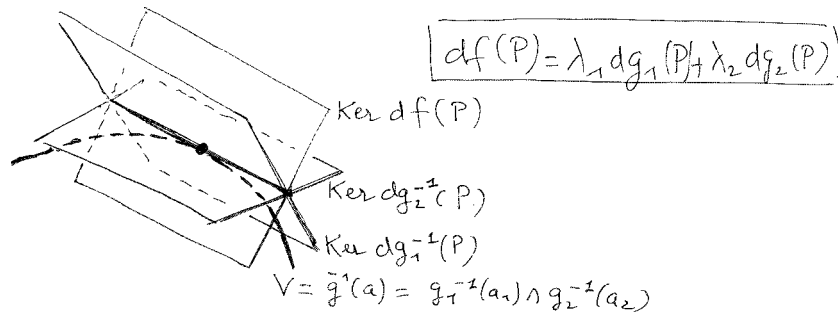
**Lemme 3** Avec les notations ci-dessus,  $P$  est un point critique de  $f|_V$  (i.e. un *extremum de  $f$  lié par la contrainte  $g = a$* ) si et seulement s'il existe une (unique) forme linéaire continue  $\lambda \in F' := L(F, \mathbb{R})$  telle que  $df(P) = \lambda \circ dg(P)$ . La forme linéaire  $\lambda$  est appelée "multiplicateur de Lagrange".

$$\begin{array}{ccc}
 & & T_{f(P)} \mathbb{R} \cong \mathbb{R} \\
 & \nearrow df(P) & \\
 E \cong T_P E & & \\
 & \searrow dg(P) & \\
 & & T_a F \cong F \\
 & & \uparrow \lambda
 \end{array}$$

Lorsque  $F = \mathbb{R}^p$  est de dimension finie (i.e. lorsqu'il y a  $p$  contraintes scalaires  $g_1 = a_1, \dots, g_p = a_p$ ), le multiplicateur de Lagrange équivaut à la donnée de  $p$  nombres réels  $\lambda_1, \dots, \lambda_p$  tels que  $df(P) = \sum_{i=1}^p \lambda_i dg_i(P)$ .

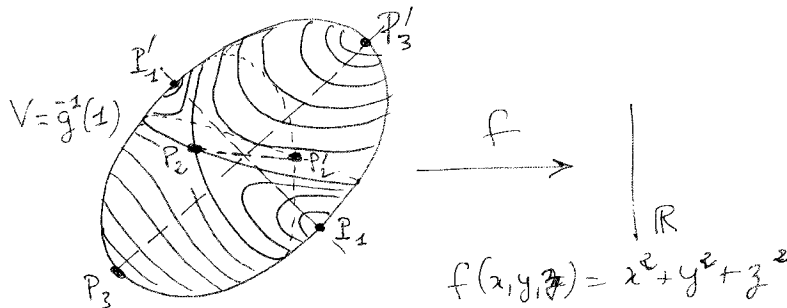
Nous dirons indifféremment "le" multiplicateur de Lagrange  $\lambda = (\lambda_1, \dots, \lambda_p)$  ou "les" multiplicateurs de Lagrange  $\lambda_1, \dots, \lambda_p$ .

La figure ci-dessous montre le cas d'une courbe  $V$  définie dans  $\mathbb{R}^3$  au voisinage de  $P$  par deux équations  $g_1 = a_1, g_2 = a_2$  indépendantes (i.e. telles que  $g = (g_1, g_2)$  soit une submersion en  $P$ , ce qui équivaut à ce que les formes linéaires  $dg_1(P)$  et  $dg_2(P)$  soient linéairement indépendantes).



La démonstration est un exercice d'algèbre linéaire : il faut caractériser les formes linéaires continues sur  $E$  qui s'annulent sur le noyau  $E_1$  d'une submersion linéaire de  $E$  dans  $F$  ; or une telle submersion se ramène, modulo des isomorphismes affines à la source et au but, à la projection canonique  $E_1 \times F \rightarrow F$ , ce qui rend l'assertion évidente.

**Un exemple d'application, les axes propres d'un ellipsoïde.** Réfléchir sur la figure ci-dessous, dans laquelle  $V$  est un ellipsoïde défini par l'équation  $g = 1$  où  $g$  est un polynôme homogène du second degré, et  $f(x, y, z) = x^2 + y^2 + z^2$ . Interpréter le multiplicateur de Lagrange associé à chacun des 6 points critiques.



# Calcul différentiel et équations différentielles

A. Chenciner, résumé de cours n°15

## Dérivées d'ordre supérieur.

**Définition 1** L'application  $f : (E, x_0) \cdots \rightarrow (F, y_0)$  est dite deux fois dérivable en  $x_0$  si, d'une part, elle est dérivable sur un voisinage  $\Omega$  de  $x_0$  et si, d'autre part, l'application  $\Omega \ni x \mapsto df(x) \in L(E, F)$  est dérivable en  $x_0$ . On appelle dérivée seconde  $d^2f(x_0) \in L^2(E, F)$  de  $f$  en  $x_0$  l'application bilinéaire de  $E \times E$  dans  $F$  canoniquement associée à  $d(df)(x_0) \in L(E, L(E, F))$  ; autrement dit,  $d^2f(x_0)(X', X'') \in F$  est la valeur en  $X'' \in E$  de  $[d(df)(x_0)](X') \in L(E, F)$ . L'application  $f$  est dite de classe  $C^2$  en  $x_0$  si  $d^2f(x)$  est définie pour  $x$  voisin de  $x_0$  et si l'application  $d^2f : E \mapsto L^2(E, F)$  est continue en  $x_0$ . On définit de même les dérivées d'ordre supérieur  $d^k f(x_0) \in L^k(E, F)$  comme applications  $k$ -linéaires de  $E \times E \times \cdots \times E$  dans  $F$ .

**Théorème 2 (Schwarz)** Si  $d^2f(x_0)$  existe, elle est symétrique :

$$d^2f(x_0)(X', X'') = d^2f(x_0)(X'', X') ; \quad \text{on note } d^2f(x_0) \in L_s^2(E, F).$$

La démonstration, plus simple si l'on suppose  $f$  de classe  $C^2$ , est basée sur l'évaluation de l'expression "du second ordre" en  $\epsilon$

$$\Delta = f(x_0 + \epsilon X + \epsilon Y) + f(x_0) - f(x_0 + \epsilon X) - f(x_0 + \epsilon Y),$$

qu'on écrit de deux façons comme différence d'expressions du premier ordre.

Plus généralement, on déduit par récurrence du théorème de Schwarz la symétrie des dérivées d'ordre supérieur :  $d^k f(x_0)(X^1, \dots, X^k)$  est invariante par permutation des  $X_i$ .

**Notations.** Si  $u \in L_s^k(E, F)$  est une application  $k$ -linéaire symétrique, on notera  $u(X^k)$  ou même  $uX^k$  à la place de  $u(X, X, \dots, X)$  et plus généralement  $u(X_1^{j_1} \cdots X_n^{j_n})$  pour la valeur de  $u$  sur le  $k$ -uplet formé de  $j_1$  exemplaires de  $X_1, \dots, j_n$  exemplaires de  $X_n$ . Avec cette notation, on vérifie la *formule du multinôme* :

$$u(X_1 + \cdots + X_n)^k = \sum_{j_1 + \cdots + j_n = k} \frac{k!}{j_1! \cdots j_n!} u(X_1^{j_1} \cdots X_n^{j_n}),$$

dans laquelle on convient que les  $X_i$  affectés de l'exposant 0 n'apparaissent pas.

**Cas de la dimension finie.** Si  $E = \mathbb{R}^n$  et  $X = \sum_{j=1}^n X_j \vec{e}_j$ , où comme d'habitude on a noté  $\vec{e}_1, \dots, \vec{e}_n$  la base canonique de  $\mathbb{R}^n$ ,

$$u(X^k) = \sum_{j_1 + \dots + j_n = k} \frac{k!}{j_1! \dots j_n!} X_1^{j_1} \dots X_n^{j_n} u(\vec{e}_1^{j_1} \dots \vec{e}_n^{j_n}),$$

et donc

$$d^k f(x_0)(X^k) = \sum_{j_1 + \dots + j_n = k} \frac{k!}{j_1! \dots j_n!} \frac{\partial^k f}{\partial x_1^{j_1} \partial x_2^{j_2} \dots \partial x_n^{j_n}}(x_0) X_1^{j_1} \dots X_n^{j_n},$$

où les  $\frac{\partial^k f}{\partial x_1^{j_1} \partial x_2^{j_2} \dots \partial x_n^{j_n}}(x_0) = d^k f(x_0)(\vec{e}_1^{j_1} \dots \vec{e}_n^{j_n})$  sont les *dérivées partielles*.

En particulier,  $d^2 f(x_0)$  est représentée par une matrice carrée symétrique, la *Hessienne* en  $x_0$ , dont les coefficients sont les *dérivées partielles secondes*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) = d^2 f(x_0)(\vec{e}_i, \vec{e}_j) = \frac{\partial^2 f}{\partial x_j \partial x_i}(x_0) \in F.$$

**Remarque importante.** Contrairement à la dérivée première, la *dérivée seconde n'est pas une notion intrinsèque* : sa définition fait appel à l'identification canonique à  $E$  de l'espace tangent  $T_x E$  en un point quelconque  $x \in E$ . Or une telle identification canonique n'existe plus si  $E$  est remplacé par une (sous-)variété. Pour la restaurer, il faut une *connexion*. Seule la notion de *jet d'ordre  $k$*  (i.e. de développement de Taylor en un point) peut être définie intrinsèquement. Pour comprendre cette remarque, il faut absolument faire l'exercice suivant

**Exercice.** Soient  $f : (E, x_0) \dots \rightarrow (F, y_0)$  et  $\varphi : (E, x_0) \dots \rightarrow (E, x_0)$  deux applications de classe  $C^2$  (on pensera à  $\varphi$  comme étant un difféomorphisme local, c'est-à-dire un "changement de coordonnées non linéaire"). Calculer la dérivée seconde en  $x_0$  de  $f \circ \varphi$ . Comparer aux formules de transformation d'une forme quadratique par changement de base.

**Polynômes de Taylor.** Si  $f : (E, x_0) \dots \rightarrow (F, y_0)$  est  $k$  fois dérivable en  $x_0$ , son *polynôme de Taylor d'ordre  $k$  en  $x_0$*  est l'application  $T_{x_0}^k f$  de  $E$  dans  $F$  définie par

$$x \mapsto f(x_0) + df(x_0)(x - x_0) + \frac{1}{2!} d^2 f(x_0)(x - x_0)^2 + \dots + \frac{1}{k!} d^k f(x_0)(x - x_0)^k.$$

**Formule de Taylor avec reste intégral.** Si  $f : E \supset \Omega \rightarrow F$  est de classe  $C^{k+1}$ , et si le segment  $[x_0, x]$  est contenu dans  $\Omega$ ,

$$f(x) = T_{x_0}^k f(x) + \left[ \int_0^1 \frac{(1-t)^k}{k!} d^{k+1} f(x_0 + t(x - x_0)) dt \right] (x - x_0)^{k+1}.$$

On en déduit la

**Formule de Taylor-Young.** Si  $f : (E, x_0) \dots \rightarrow (F, y_0)$  est  $k$  fois dérivable au point  $x_0$ ,

$$f(x) = T_{x_0}^k f(x) + o(|x - x_0|^k)$$

# Calcul différentiel et équations différentielles

A. Chenciner, résumé de cours n°16

**Développements limités.** Une application polynomiale  $P : E \rightarrow F$  est une application de la forme  $P(x) = \sum_{i=0}^k u_i(x^i)$ , où  $u_i \in L_s^i(E, F)$  (avec la convention que  $L_s^0(E, F) = F$ ).

**Définition 1** On dit que l'application  $f : (E, x_0) \cdots \rightarrow (F, y_0)$  possède un développement limité à l'ordre  $k$  (ou encore un  $dl_k$ ) en  $x_0$  s'il existe un polynôme  $P(x)$ , qu'on écrit en général sous la forme  $P(x) = \sum_{i=0}^k u_i((x - x_0)^i)$ , tel que

$$f(x) = P(x) + o(|x - x_0|^k).$$

**ATTENTION !** Une reformulation de Taylor-Young est que l'existence de  $d^k f(x_0)$  implique celle d'un  $dl_k$  de  $f$  en  $x_0$ . Réciproquement, l'existence en  $x_0$  d'un  $dl_1$  implique l'existence de  $df(x_0)$  mais celle d'un  $dl_k$  pour  $k \geq 2$  n'implique pas même l'existence de  $d^2 f(x_0)$ . Les exemples classiques sont les  $f : \mathbb{R} \rightarrow \mathbb{R}$  définies par  $f(x) = x^{k+1} \sin \frac{1}{x^m}$ .

**Compositions.** Une composée d'applications de classe  $C^k$  est encore de classe  $C^k$ . Cela découle de la

**Proposition 2 (formule de Faa di Bruno).** Si  $f : (E, x_0) \cdots \rightarrow (F, y_0)$  et  $g : (F, y_0) \cdots \rightarrow (G, z_0)$  sont  $k$  fois dérivables respectivement en  $x_0$  et en  $y_0$ ,  $g \circ f$  est  $k$  fois dérivable en  $x_0$  et l'on a, en notant  $p! = \Pi(p_j!)$  et  $|p| = \sum p_j$ ,

$$\frac{1}{k!} d^k (g \circ f)(x_0) X^k = \sum \frac{1}{p!} d^{|p|} g(f(x_0)) \left( (df(x_0)X)^{p_1}, \dots, \left( \frac{1}{k!} df^k(x_0) X^k \right)^{p_k} \right),$$

où la somme porte sur tous les  $p = (p_1, \dots, p_k)$  tels que  $\sum j p_j = k$ .

**De  $C^1$  à  $C^k$ .** De façon cavalière (mais correcte), il nous suffira d'affirmer que si, dans les théorèmes du type inversion locale, submersion, immersion, sous-variétés que l'on a énoncés, on remplace  $C^1$  par  $C^k$  dans l'hypothèse de régularité faite sur l'application considérée, on peut également remplacer  $C^1$  par  $C^k$  dans la conclusion. Il faut bien entendu pour cela définir, par exemple, ce qu'est une sous-variété de classe  $C^k$  mais le lecteur le fera sans peine.

Prenons comme exemple le théorème d'inversion locale (feuille 10, lemme 6) et supposons que  $h$  soit de classe  $C^k$  avec  $k \geq 2$ . Montrons par récurrence que  $h^{-1}$  est de classe  $C^k$ . Nous savons déjà que l'inverse  $h^{-1}$  est de classe  $C^1$ , ce qui initie la récurrence. S'il est de classe  $C^{k-1}$ , il en est de même de sa dérivée car celle-ci est la composée de trois applications de classe  $C^{k-1}$ , à savoir  $h^{-1}$ ,  $Dh$ , et enfin l'application qui à un isomorphisme linéaire fait correspondre son inverse. On en déduit que  $h$  est un difféomorphisme local de classe  $C^k$ .



**Extrema.** La notion de point critique étant locale, le cas d'une fonction à valeurs réelles définie sur une (sous)-variété  $V$  d'un espace de Banach se ramène, par le choix d'une carte locale à la considération du cas où  $f : \Omega \rightarrow \mathbb{R}$  est définie sur un ouvert  $\Omega$  d'un espace de Banach  $E$ . Ce n'est bien entendu pas une raison pour oublier la méthode des multiplicateurs de Lagrange dans la recherche des points critiques mais une fois ceux-ci identifiés on peut faire ce choix de travailler dans une carte locale.

**Définition 3** Une fonction  $f : \Omega \rightarrow \mathbb{R}$  a un minimum local (resp. un maximum local) en  $x_0$  s'il existe un voisinage  $\mathcal{V}$  de  $x_0$  sur lequel  $f(x_0) = \inf f|_{\mathcal{V}}$  (resp.  $f(x_0) = \sup f|_{\mathcal{V}}$ ). Si  $x_0$  est seul de son espèce dans  $\mathcal{V}$ , on dit que c'est un minimum local (resp. maximum local) strict.

On déduit de la formule de Taylor-Young à l'ordre deux le

**Théorème 4** Supposons  $f$  deux fois dérivable en  $x_0$ . Si  $x_0$  est un minimum local de  $f$ , c'est un point critique (i.e.  $df(x_0) = 0$ ) et on a  $d^2f(x_0) \geq 0$  (i.e.  $d^2f(x_0)$  est une forme quadratique positive). Si  $x_0$  est un point critique et si  $d^2f(x_0)$  est définie positive,  $x_0$  est un minimum strict. Pour les maxima, remplacer positive par négative.

**Exercice.** Généraliser le théorème ci-dessus au cas de fonctions dont le développement de Taylor en  $x_0$  commence à l'ordre  $k > 2$ .

**Un coup d'œil sur le lemme de Morse.** C'est une généralisation du théorème des submersions.

Rappelons que ce dernier implique qu'au voisinage d'un point régulier  $x_0$  (i.e. d'un point *non critique*, on dit aussi *non singulier*) une fonction de classe  $C^1$ ,  $f : (\mathbb{R}^n, x_0) \cdots \rightarrow (\mathbb{R}, y_0)$  s'écrit  $f = T_{x_0}^1 f \circ \varphi$ , où  $\varphi : (E, x_0) \cdots \rightarrow (E, x_0)$  est un difféomorphisme local.

Si maintenant  $x_0$  est un point critique *non dégénéré* de  $f$ , supposée de classe  $C^k$ ,  $k \geq 3$ , i.e. si  $df(x_0) = 0$  et  $d^2f(x_0)$  est une forme quadratique non dégénérée (ou encore si  $X \mapsto [Y \mapsto d^2f(x_0)(X, Y)]$  est un isomorphisme de  $E$  sur son dual  $E' = L(E, \mathbb{R})$ ),  $f$  s'écrit localement  $f = T_{x_0}^2 f \circ \varphi$ , où  $\varphi : (E, x_0) \cdots \rightarrow (E, x_0)$  est un difféomorphisme local de classe  $C^{k-2}$ .

L'une des démonstration de ce lemme est une jolie conséquence du théorème d'existence des solutions d'équations différentielles.

. . . **et les singularités.** L'étude du comportement local des fonctions de classe  $C^k$  au voisinage d'un point critique est une branche très riche des mathématiques, appelée *théorie des singularités des fonctions (plus généralement des applications) différentiables*. On trouvera une introduction à cette théorie dans l'article que j'ai écrit pour l'*Encyclopédie Universalis*.

# Calcul différentiel et équations différentielles

A. Chenciner, résumé de cours n°17

## 2<sup>ème</sup> partie : ÉQUATIONS DIFFÉRENTIELLES.

Considéré comme la traduction mathématique du concept de *déterminisme*, le théorème d'existence et d'unicité (locales) des solutions d'équations différentielles joue un rôle fondamental en mathématique et en physique mais l'on sait depuis plus d'un siècle, en particulier depuis les recherches de Poincaré, que

– d'une part la plupart des équations ne peuvent être intégrées (adieu les formules explicites et les recettes magiques),

– d'autre part ce déterminisme mathématique n'implique en général pas la possibilité pratique de prédire l'évolution à long terme des solutions : c'est la *dépendance sensible des données initiales*, stupidement appelée *théorie du chaos* (voir cependant le recueil *Chaos et déterminisme* dans "Points Sciences"), qui explique l'importance des méthodes probabilistes (*théorie ergodique*).

Une exception notable à ce sombre tableau est donnée par les équations linéaires "à coefficients constants" dans  $\mathbb{R}^n$ . mais même dans ce cas, la compréhension de la situation passe par la géométrie. Ce sera l'occasion d'introduire diverses notions telles *champ de vecteurs*, *portrait de phase*, *flot*, *intégrale première*, *fonction de Lyapunov*, etc...

**Champs de vecteurs et équations différentielles dans un espace de Banach.** Soit  $\Omega \subset E$  un ouvert d'un espace de Banach. Un *champ de vecteurs (autonome)*  $X$  de classe  $C^k$  sur  $\Omega$  est la donnée en chaque point  $x \in \Omega$  d'un vecteur tangent  $X(x) \in T_x\Omega$ . Puisque chaque espace tangent  $T_x\Omega$  est canoniquement identifié à  $E$  par translation (feuille 5), on peut considérer  $X$  comme une application de  $\Omega$  dans  $E$ , ce qui donne un sens à la régularité  $C^k$ ,  $k \geq 0$ , Lipschitz, etc... À un champ de vecteurs  $X$  sur  $\Omega$  est associée l'équation différentielle  $\frac{dx}{dt} = X(x)$ . Une *solution* d'une telle équation est une application  $t \mapsto x(t)$  d'un intervalle  $[a, b]$  à valeurs dans  $\Omega$  telle que pour tout  $t \in [a, b]$  on ait :  $\frac{dx}{dt} = X(x(t))$ , où  $\frac{dx}{dt}$  désigne le *vecteur-vitesse*  $dx(t) \cdot 1$ , image de la dérivée par l'isomorphisme canonique de  $L(\mathbb{R}, E)$  sur  $E$ . Autrement dit, une solution est une courbe paramétrée dans  $\Omega$  (un mobile si l'on pense à la variable  $t$  comme désignant le temps) qui, lorsqu'elle passe par le point  $x$  a pour vitesse le vecteur  $X(x) \in T_x\Omega \cong E$ .

Plus généralement, à un *champ de vecteurs dépendant du temps* (i.e. non autonome)  $X(t, x)$ , on associe l'équation différentielle  $\frac{dx}{dt} = X(t, x)$ , dont les solutions sont les courbes paramétrées  $t \mapsto x(t)$  qui, lorsqu'elles passent au temps  $t$  par le point  $x$ , y ont la vitesse  $X(t, x)$ .

Si  $E = \mathbb{R}$ , on a une “équation différentielle scalaire” ; si  $E = \mathbb{R}^n$ , on a un “système de  $n$  équations différentielles scalaires”  $\frac{dx_i}{dt} = X_i(t, x_1, \dots, x_n)$ ,  $i = 1, \dots, n$ . Un exemple typique s’obtient à partir d’une équation différentielle scalaire d’ordre  $n$ ,

$$\frac{d^n x}{dt^n} = f\left(t, x, \frac{dx}{dt}, \dots, \frac{d^{n-1}x}{dt^{n-1}}\right)$$

en posant  $x_1 = x, x_2 = \frac{dx}{dt}, \dots, x_n = \frac{d^{n-1}x}{dt^{n-1}}$ .

**Existence et d’unicité pour le “problème de Cauchy”.** Dans ce qui suit,  $X(t, x)$  est un champ de vecteurs dépendant du temps défini sur un ouvert  $\mathcal{U}$  de  $\mathbb{R} \times E$ . On appelle *problème de Cauchy* la recherche d’une solution  $x(t)$  de l’équation  $\frac{dx}{dt} = X(t, x)$ , qui soit définie pour  $t$  voisin de  $t_0$  et vérifie  $x(t_0) = x_0$  (le couple  $(t_0, x_0)$  est la *donnée initiale*). Ces énoncés seront précisés par la suite en ce qui concerne la régularité des solutions, leur dépendance des données initiales et d’éventuels paramètres et les cas d’existence et unicité globales.

**Théorème 1 (Cauchy-Lipschitz)** *Si l’application  $(t, x) \mapsto X(t, x)$  est continue et localement (en  $(t, x)$ ) lipschitzienne en la variable  $x$  (i.e. si chaque point  $(t_0, x_0) \in \mathcal{U}$  possède un voisinage  $I \times \Omega$  tel qu’une même constante de Lipschitz vaille pour les applications  $\Omega \ni x \mapsto X(t, x)$  où  $t \in I$ ), il y a existence et unicité locale d’une solution du problème de Cauchy.*

**Théorème 2 (Peano)** *Si  $E$  est de dimension finie, la seule continuité de l’application  $(t, x) \mapsto X(t, x)$  suffit à assurer l’existence locale, mais pas forcément l’unicité, d’une solution du problème de Cauchy.*

**Solutions maximales.** Ce sont les solutions  $I \ni t \mapsto x(t) \in E$  qui ne se prolongent pas à un intervalle plus grand que  $I$ . Par exemple,  $t \mapsto \frac{1}{t-t_0}$  de  $I = ]t_0, +\infty[$  dans  $\mathbb{R}$  est une solution maximale de l’équation  $\frac{dx}{dt} = -x^2$ .

**Remarques. 1) Des équations non autonomes aux équations autonomes.**

On peut toujours transformer un champ de vecteurs dépendant du temps dans  $E$  en un champ de vecteurs autonome dans  $\mathbb{R} \times E$  en prenant le temps lui-même comme variable : on associe à l’équation non autonome  $\frac{dx}{dt} = X(t, x)$  l’équation autonome  $\frac{d\xi}{dt} = \Xi(\xi)$  où  $\xi = (\tau, x)$  et  $\Xi(\xi) = (1, X(\tau, x))$ , c’est-à-dire  $\frac{d\tau}{dt} = 1, \frac{dx}{dt} = X(\tau, x)$ . Attention, les solutions de la nouvelle équation sont non seulement les solutions  $t \mapsto x(t)$  de l’équation originale mais encore toutes les applications  $t \mapsto x(t - t_0)$  obtenues en changeant l’origine du temps.

**2) Equations multiformes.** Une équation différentielle générale se présente sous la forme

$$f\left(t, x, \frac{dx}{dt}, \dots, \frac{d^n x}{dt^n}\right) = 0.$$

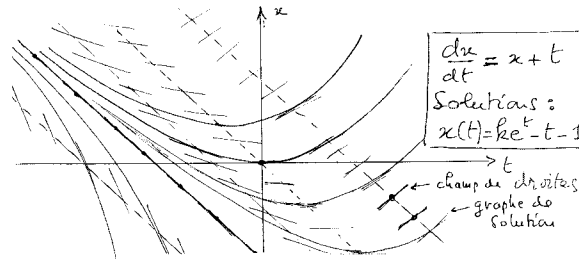
Si cette équation ne détermine pas univoquement la dérivée d’ordre maximal  $\frac{d^n x}{dt^n}$  comme fonction de  $t, x, \frac{dx}{dt}, \dots, \frac{d^{n-1}x}{dt^{n-1}}$ , elle ne se ramène pas à un champ de vecteurs dépendant du temps et il ne peut y avoir unicité, même si  $f$  est très régulière. C’est le cas pour l’équation  $x^2 + \left(\frac{dx}{dt}\right)^2 - 1 = 0$ , dont on résoudra le problème de Cauchy.

# Calcul différentiel et équations différentielles

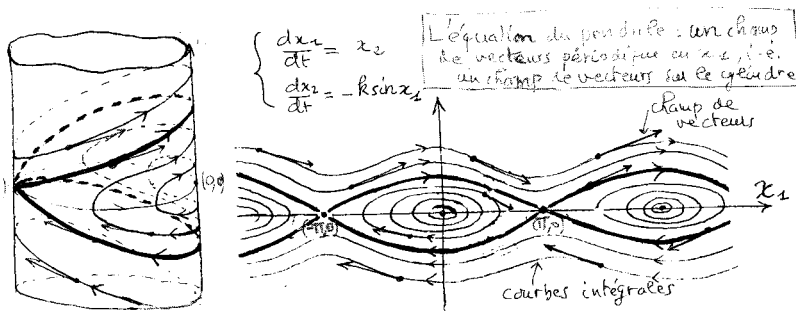
A. Chenciner, résumé de cours n°18

## Portraits de phase.

**1) Dans l'espace temps.** Un champ de vecteurs dépendant éventuellement du temps  $X(t, x)$  définit un champ de directions (non orientées) dans l'espace temps. Les graphes des solutions sont tangents en chaque point à la direction attachée à ce point. Le champ de vecteurs est autonome si et seulement si le feuilletage de l'espace-temps défini par ces graphes est invariant par les translations  $t \mapsto t - t_0$ .

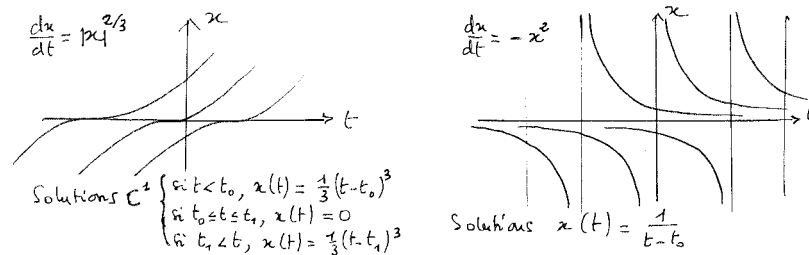


**2) Dans l'espace.** (Cette représentation n'a d'intérêt que dans le cas autonome car sinon les courbes intégrales s'intersectent. Ex :  $\ddot{x} = 0$  dans le plan.) On confondra souvent la solution  $x(t)$  et la courbe intégrale qu'elle définit dans  $E$ . Il faut cependant noter que, si  $t \mapsto x(t)$  est solution de l'équation différentielle autonome  $\frac{dx}{dt} = X(x)$ , il en est de même de  $t \mapsto x(t - t_0)$  et toutes ces solutions définissent la même courbe intégrale. On en déduit que, dans le cas autonome, s'il y a existence et unicité, les différentes courbes intégrales sont ou bien disjointes ou bien confondues ; elles définissent en particulier un feuilletage du domaine de définition de  $X$ .



**Remarque.** Les graphes des solutions dans l'espace temps de  $\frac{dx}{dt} = X(t, x)$  sont exactement les courbes intégrales du champ autonome  $\Xi(\tau, x) = (1, X(\tau, x))$  qu'on a associé à  $X(t, x)$ .

**Exemples et contre-exemples en dimension 1.** Les deux figures sont dans l'espace-temps. La première est un exemple de champ non lipschitzien, avec existence (globale) des solutions mais non unicité ; la deuxième un exemple de champ  $C^\infty$  dans lequel la solution maximale n'est pas définie sur  $\mathbb{R}$  tout entier (un exemple encore plus simple de ce dernier fait est un champ constant dans  $\mathbb{R}^n$  privé d'un point).



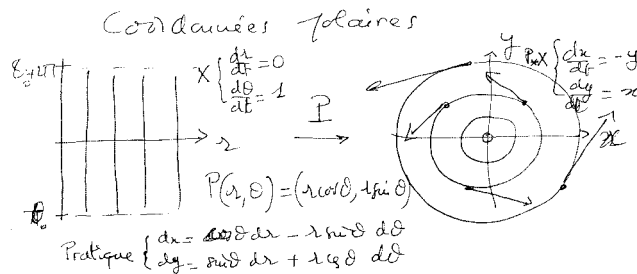
**Changements de coordonnées : image directe d'un champ de vecteurs par un difféomorphisme.**

**Définition 1** Soit  $\varphi : E \supset \Omega \dots \rightarrow \mathcal{O} \subset F$  un difféomorphisme,  $X$  un champ de vecteurs autonome défini sur  $\Omega$ . L'image directe  $\varphi_*X$  de  $X$  est le champ de vecteurs sur  $\mathcal{O}$  défini par

$$(\varphi_*X)(\varphi(x)) = d\varphi(x)X(x).$$

Rappelons la définition mécanique de la dérivée donnée dans la feuille 5 : l'image par la dérivée  $d\varphi(x_0)$  de la vitesse d'un mobile passant au point  $x_0$  est la vitesse au point  $y_0 = \varphi(x_0)$  du mobile image. On en déduit immédiatement

**Proposition 2** Les courbes intégrales du champ  $\varphi_*X$  sont les images par  $\varphi$  des courbes intégrales de  $X$ .



**Remarque.** Si  $X$  dépend du temps, on définit  $\varphi_*X$  comme ci-dessus ; bien entendu, on peut également faire des changements de coordonnées "dépendant du temps".

# Calcul différentiel et équations différentielles

A. Chenciner, résumé de cours n°19

## Équations différentielles linéaires.

Ce sont les équations de la forme

$$(L) \quad \frac{dx}{dt} = A(t)x + b(t),$$

où  $t \mapsto A(t)$  et  $t \mapsto b(t)$  sont définies sur un intervalle  $I$  de  $\mathbb{R}$ , à valeurs respectivement dans  $L(E, E)$  et  $E$ . Il faudrait plutôt parler d'équations affines mais la terminologie classique est "équation sans (ou avec) second membre" suivant que  $b(t)$  est (ou non) identiquement nulle" (dans le premier cas, on parle aussi d'"équation homogène").

Si  $A$  ne dépend pas du temps (en dimension finie : (systèmes d')équations linéaires "à coefficients constants"), l'équation se résout explicitement :

**Théorème 1** *Le problème de Cauchy  $\frac{dx}{dt} = Ax + b(t)$ ,  $x(t_0) = x_0$  admet une solution unique, définie sur  $\mathbb{R}$  tout entier :*

$$x(t) = e^{(t-t_0)A} \left( x_0 + \int_{t_0}^t e^{-(t-t_0)A} b(t) dt \right).$$

*Démonstration :* La série  $Id + tA + \frac{1}{2!}t^2A^2 + \dots + \frac{1}{n!}t^nA^n + \dots$  de fonctions de  $\mathbb{R}$  dans  $L(E, E)$  converge normalement, donc uniformément, sur tout compact car la série des normes est majorée par la série définissant  $e^{T\|A\|}$ , où  $T$  est le max de  $|t|$  sur le compact considéré. Il en est de même de la série obtenue par dérivation terme à terme, ce qui montre que  $t \mapsto e^{tA}$  est de classe  $C^1$  avec pour vecteur-dérivée en  $t$ ,  $Ae^{tA} = e^{tA}A$ . On en déduit que si  $x(t)$  est solution de l'équation  $\frac{dx}{dt} = Ax$ , la dérivée de  $e^{-tA}x(t)$  est identiquement nulle. On conclut par *variation des constantes*, i.e. en cherchant les solutions sous la forme  $x(t) = k(t)e^{tA}$ .

On déduit en particulier de ce théorème que l'ensemble des solutions forme un sous-espace affine de  $C^0(\mathbb{R}, E)$  (de même dimension que  $E$  si cette dimension est finie). Les algébristes auront plaisir à résumer la situation dans la *suite exacte*

$$0 \rightarrow E \rightarrow C^k(\mathbb{R}, E) \rightarrow C^{k-1}(\mathbb{R}, E) \rightarrow 0,$$

qui signifie que la deuxième flèche,  $x_0 \mapsto e^{tA}x_0$ , est injective, que la troisième,  $x(t) \mapsto \frac{dx}{dt}(t) - Ax(t)$ , est surjective et que son noyau coïncide avec l'image de la deuxième.

**Théorème 2 Cauchy-Lipschitz linéaire** Si  $t \mapsto A(t)$  et  $t \mapsto b(t)$  sont continues, le problème de Cauchy pour l'équation (L) possède une solution unique, définie sur  $I$  tout entier.

La démonstration est une simple application du théorème du point fixe des contractions. En effet, les solutions  $x(t) = x_0 + y(t)$  de l'équation correspondent aux points fixes  $y(t)$  de l'application  $\Phi$  :

$$C^0(I, E) \ni y(t) \mapsto \int_{t_0}^t [A(s)(x_0 + y(s)) + b(s)] ds \in C^0(I, E).$$

Si  $I$  est compact,  $C^0(I, E)$  est un espace de Banach (pour la norme du sup, c'est-à-dire la norme de la convergence uniforme) et on a une majoration explicite

$$\|\Phi^n(y_1) - \Phi^n(y_2)\|_{C^0} \leq \frac{|I|^n K^n}{n!} \|y_1 - y_2\|_{C^0},$$

où  $|I|$  est la longueur de l'intervalle compact  $I$  et  $K = \sup_{s \in I} A(s)$  ; on en déduit qu'un itéré de  $\Phi$  est une contraction. Si  $I$  n'est pas compact, on l'écrit comme réunion croissante de compacts et on conclut en utilisant l'unicité.

Je laisse au lecteur le plaisir d'écrire la même suite exacte que dans le cas "à coefficients constants".

**Remarque.** Lorsque les opérateurs  $A(s)$  commutent deux à deux, par exemple lorsque  $E = \mathbb{R}$ , la solution du problème de Cauchy est donnée par une formule explicite : il suffit de remplacer  $(t - t_0)A$  par  $\int_{t_0}^t A(s) ds$  dans la formule du Théorème 1.

**Flot.** Le flot de l'équation (L) est la famille  $\varphi_{t_0}^t \in GL(E)$  d'isomorphismes affines (linéaires si  $b \equiv 0$ ) continus définis, pour  $t_0$  et  $t$  appartenant à  $I$ , par la condition que, si  $x \in E$ ,  $\varphi_{t_0}^t(x) \in E$  soit la valeur à l'instant  $t_1$  de l'unique solution qui vaut  $x$  à l'instant  $t_0$ . En particulier,

$$(\varphi_{t_0}^{t_1})^{-1} = \varphi_{t_1}^{t_0}, \quad \varphi_{t_1}^{t_2} \circ \varphi_{t_0}^{t_1} = \varphi_{t_0}^{t_2}.$$

Lorsque l'équation est autonome ( $A$  et  $b$  indépendants du temps),  $I = \mathbb{R}$  et  $\varphi_{t_0}^{t_1}$ , qui ne dépend que de la différence  $t = t_1 - t_0$ , se note en général  $\varphi_t$ . Si par exemple  $b \equiv 0$ , on obtient un groupe à un paramètre d'éléments de  $GL(E)$ , c'est-à-dire un homomorphisme de groupes  $t \mapsto \varphi_t$  du groupe additif  $\mathbb{R}$  dans le groupe (pour la composition)  $GL(E)$  des isomorphismes linéaires continus de l'espace  $E$ .

# Calcul différentiel et équations différentielles

A. Chenciner, résumé de cours n°20

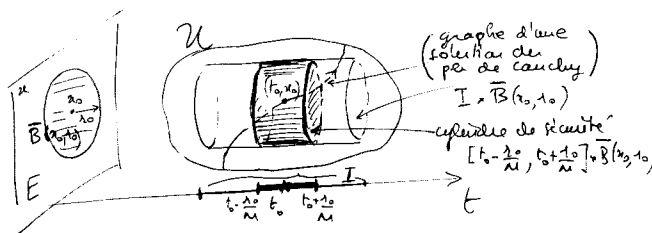
## Trois démonstrations du théorème de Cauchy-Lipschitz.

**Théorème 1 Cauchy-Lipschitz** Soit  $\mathcal{U}$  un ouvert de  $\mathbb{R} \times E$ ,  $X : \mathcal{U} \rightarrow E$  un champ de vecteurs dépendant du temps. On suppose que l'application  $X$  est continue et localement lipschitzienne en  $x$  (i.e. à tout  $(t_0, x_0) \in \mathcal{U}$  on peut associer un réel  $k > 0$  et un voisinage  $I \times \Omega$  tel que, quel que soit  $t \in I$ , l'application  $\Omega \ni x \mapsto X(t, x) \in E$  soit  $k$ -lipschitzienne). Alors, le problème de Cauchy

$$\frac{dx}{dt} = X(t, x), \quad x(t_0) = x_0$$

admet une unique solution locale de classe  $C^1$ ,  $I \ni ]t_0 - \epsilon, t_0 + \epsilon[ \ni t \mapsto x(t) \in \Omega$ . Plus précisément, on peut choisir  $\epsilon = \frac{r_0}{M}$ , où  $r_0$  est le rayon d'une boule fermée  $\bar{B}(x_0, r_0)$  de centre  $x_0$  contenue dans  $\Omega$  et  $M = \sup_{(t,x) \in I \times \bar{B}(x_0, r_0)} \|X(t, x)\|$ .

**Préliminaire : cylindres de sécurité.** La simple continuité de  $X$  implique que si  $x(t)$  est une solution définie sur  $[t_0 - \epsilon, t_0 + \epsilon]$  du problème de Cauchy,  $x(t)$  reste dans la boule fermée  $\bar{B}(x_0, r_0)$  tant que  $\epsilon \leq \frac{r_0}{M}$  (regarder le premier temps  $t$  pour lequel  $x(t)$  appartient au bord de la boule). On dit que le cylindre  $C = [t_0 - \epsilon, t_0 + \epsilon] \times \bar{B}(x_0, r_0)$  est un *cylindre de sécurité* pour le champ de vecteurs  $X$ .



### Démonstration 1 : le théorème de Cauchy-Lipschitz comme conséquence du théorème du point fixe (Picard).

Un fois déterminé un cylindre de sécurité, c'est exactement la démonstration exposée dans le cas linéaire, à ceci près qu'on doit remplacer  $C^0(I, E)$  par  $C^0([t_0 - \epsilon, t_0 + \epsilon], \bar{B}(x_0, r_0))$ , l'application  $\Phi$  par

$$y(t) \mapsto x_0 + \int_{t_0}^t X(s, x_0 + y(s)) ds,$$

et utiliser l'hypothèse faite sur  $X$  dans l'estimation lipschitzienne de  $\Phi$ .



**Démonstration 2 : le théorème de Cauchy-Lipschitz comme conséquence du théorème d'inversion locale (Robbin).** Au lieu de commencer par chercher un cylindre de sécurité, ce qui permet de travailler dans un espace d'applications dont le domaine de définition  $[t_0 - \epsilon, t_0 + \epsilon]$  est fixé, on considère  $\epsilon$  comme une inconnue du problème et on se ramène à l'intervalle  $[-1, +1]$  par un changement d'échelle. Autrement dit, on cherche une solution du problème de Cauchy sous la forme

$$x(t) = x_0 + y((t - t_0)/\epsilon),$$

où  $\tau \mapsto y(\tau)$  est une application continue de  $[-1, +1]$  dans  $\Omega$ .

Il s'agit donc de résoudre par rapport à  $y$  au voisinage de  $(0, \hat{x}_0, 0)$  l'équation implicite  $F(\epsilon, x_0, y) = 0$ , où

$$F : I \times \Omega \times C_0^1([-1, +1], \Omega) \ni (\epsilon, x_0, y) \mapsto \frac{dy}{d\tau} - \epsilon X(t_0 + \epsilon\tau, x_0 + y) = 0 \in C^0([-1, +1], E).$$

L'espace  $C_0^1([-1, +1], \Omega)$  est l'ensemble des applications de classe  $C^1$  de  $[-1, +1]$  dans  $\Omega$  qui s'annulent en 0, muni de la norme

$$\|y\|_1 := \max_{\tau \in [-1, +1]} \{ \|y(\tau)\|, \left\| \frac{dy}{d\tau}(\tau) \right\| \}$$

qui en fait un espace de Banach. Il reste à voir que le théorème des fonctions implicites est applicable ici. Si l'on suppose que  $X$  est de classe  $C^1$  en  $(t, x)$ , il suffit de remarquer que la dérivée  $\frac{\partial F}{\partial y}(0, \hat{x}_0, 0)$  est l'application linéaire

$$C_0^1([-1, +1], E) \ni y \mapsto \frac{dy}{d\tau} \in C^0([-1, +1], E),$$

qui est un isomorphisme continu. On trouve ainsi une solution locale unique  $y_{\epsilon, x_0}$  qui dépend de manière  $C^1$  de la donnée initiale  $x_0$  et de  $\epsilon$ . Le théorème d'inversion locale qu'on a démontré dans le cadre lipschitzien permet de traiter le cas général (exercice !). Un autre exercice consiste en la démonstration de ce que si  $\epsilon' < \epsilon$ , la solution  $y_{\epsilon', x_0}$  est la restriction à l'intervalle  $[t_0 - \epsilon', t_0 + \epsilon']$  de la solution  $y_{\epsilon, x_0}$ .

**Démonstration 3 : le théorème de Cauchy-Lipschitz comme conséquence de la construction de solutions approchées (Euler) et d'un lemme de comparaison (Gronwall).**

ATTENTION : la convergence de la méthode d'Euler n'est assurée que si

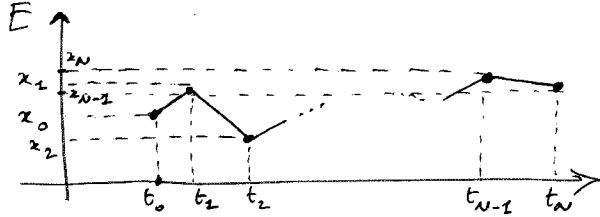
- ou bien  $X(t, x)$  est localement Lipschitzien en toute les variables  $(t, x)$  ;
- ou bien  $X(t, x)$  n'est que continue mais  $E$  est de dimension finie ; dans ce cas, grâce au *Théorème d'Ascoli*, cette méthode prouve le *théorème de Peano* assurant l'existence locale des solutions mais pas forcément leur unicité.

1<sup>ère</sup> étape : *construction de solutions approchées*. La méthode la plus élémentaire est celle d'Euler : on divise le temps en  $N$  petits intervalles sur chacun desquels on remplace  $X(t, x)$  par un champ constant ; autrement dit, on définit le graphe

d'une solution approchée  $\xi : [t_0 - \epsilon, t_0 + \epsilon] \rightarrow E$  en interpolant linéairement entre les points  $(t_n, x_n)$  définis par récurrence à partir de  $(t_0, x_0)$  par

$$x_{n+1} = x_n + (t_{n+1} - t_n)X(t_n, x_n),$$

c'est-à-dire : si  $t_n \leq t \leq t_{n+1}$ ,  $\xi(t) = x_n + (t - t_n)X(t_n, x_n)$ .



*Majoration de l'erreur :*

– Supposons  $X$  Lipschitzien en  $(t, x)$ , c'est-à-dire

$$\|X(t, x) - X(t', x')\| \leq k(|t - t'| + \|x - x'\|) \text{ dans le cylindre } C,$$

Après avoir vérifié que  $\xi(t)$  est à valeurs dans la boule  $\bar{B}(t_0, x_0)$ , on constate qu'en tous ses points de dérivabilité,  $\xi$  vérifie :

$$\left\| \frac{d\xi}{dt}(t) - X(t, \xi(t)) \right\| \leq k(1 + M) \max_{0 \leq i \leq N-1} |t_i - t_{i+1}|.$$

– Si la dimension de  $E$  est finie, on déduit de la compacité de  $C$  que  $X_C$  est uniformément continue. Si  $\omega_X : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  est le module de continuité de  $X$ ,

$$\omega_X(u) = \max\{\|X(t, x) - X(t', x')\|; |t - t'| + \|x - x'\| \leq u\},$$

on montre (exercice !) que

$$\left\| \frac{d\xi}{dt}(t) - X(t, \xi(t)) \right\| \leq \omega_X(1 + M) \max_{0 \leq i \leq N-1} |t_i - t_{i+1}|.$$

*2<sup>ème</sup> étape : le lemme de Gronwall.*

**Lemme 2** *On suppose vérifiées les hypothèses du théorème de Cauchy-Lipschitz. Soient  $u_i : I \rightarrow \Omega$ ,  $i = 1, 2$ , deux solutions approchées de l'équation différentielle  $\frac{dx}{dt} = X(t, x)$  ; plus précisément,  $u_1$  et  $u_2$  sont deux applications continues, dérivables par morceaux, telles que*

$$\left\| \frac{du_i}{dt}(t) - X(t, u_i(t)) \right\| \leq \epsilon_i, \quad i = 1, 2, \quad \|u_1(t_0) - u_2(t_0)\| \leq \delta.$$

Alors, pour tout  $t \in I$ ,

$$\|u_1(t) - u_2(t)\| \leq \delta e^{k|t-t_0|} + \frac{\epsilon_1 + \epsilon_2}{k} (e^{k|t-t_0|} - 1).$$

**Esquisse de démonstration.** Supposons  $t - t_0 > 0$ . Ecrivaint

$$u_1(t) - u_2(t) = u_1(t_0) - u_2(t_0) + \int_{t_0}^t \left( \frac{du_1}{dt}(s) - \frac{du_2}{dt}(s) \right) ds,$$

on déduit des hypothèses l'inéquation intégrale

$$\|u_1(t) - u_2(t)\| \leq \delta + (\epsilon_1 + \epsilon_2)(t - t_0) + k \int_{t_0}^t \|u_1(s) - u_2(s)\| ds.$$

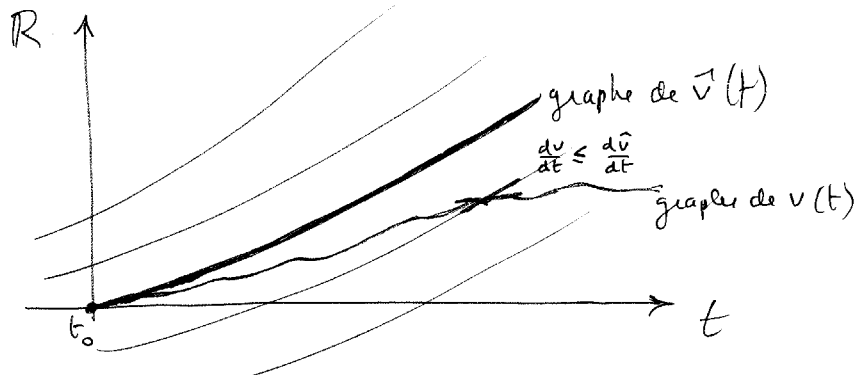
Posant  $v(t) = \int_{t_0}^t \|u_1(s) - u_2(s)\| ds$ , on obtient l'inéquation différentielle

$$\frac{dv}{dt}(t) \leq \delta + (\epsilon_1 + \epsilon_2)(t - t_0) + kv(t).$$

L'équation différentielle linéaire obtenue en remplaçant  $\leq$  par  $=$  s'intègre explicitement par "variation des constantes" ; la solution nulle en  $t = t_0$  s'écrit

$$\hat{v}(t) = \frac{\delta}{k}(e^{k(t-t_0)} - 1) + \frac{\epsilon_1 + \epsilon_2}{k^2} (e^{k(t-t_0)} - 1 - k(t - t_0)).$$

On conclut car toute solution  $v(t)$  de l'inéquation différentielle telle que  $v(t_0) = 0$  vérifie  $v(t) \leq \hat{v}(t)$  (figure).



**Exercice.** Montrer à l'aide du lemme de Gronwall,

- 1) l'existence dès qu'existe une suite de solutions approchées avec  $\delta = 0$  telle que  $\epsilon_1 + \epsilon_2$  tende vers 0 : une telle suite est en effet de Cauchy pour la norme uniforme et sa limite est nécessairement une solution  $C^1$  de l'équation différentielle ;

- 2) l'unicité : ( $\delta = \epsilon_1 = \epsilon_2 = 0$ ) ;

- 3) la dépendance  $C^0$  par rapport aux données initiales ( $\epsilon_1 = \epsilon_2 = 0$ ).

# Calcul différentiel et équations différentielles

A. Chenciner, résumé de cours n°21

## FLOTS

**Solutions maximales et résolvante.** On considère comme précédemment un champ de vecteurs dépendant du temps  $X(t, x)$  défini sur un ouvert  $\mathcal{U}$  de  $\mathbb{R} \times E$  et vérifiant les hypothèses du théorème de Cauchy-Lipschitz.

**Lemme 1** *Chaque solution locale du problème de Cauchy*

$$\frac{dx}{dt} = X(t, x), \quad x(t_0) = x_0$$

*se prolonge en une unique solution maximale (i.e. ne pouvant être prolongée à un intervalle de temps plus grand). De plus, l'intervalle de définition d'une solution maximale est nécessairement ouvert.*

On déduit en effet de l'unicité locale d'une part, de la continuité des solutions d'autre part, que l'ensemble des temps en lesquels deux solutions coïncident est à la fois ouvert et fermé. Deux solutions coïncidant en un point coïncident donc sur toute l'intersection (qui est un intervalle) de leurs intervalles de définition. L'ouverture de l'intervalle maximal vient de l'existence locale de solutions.

**Définition 2** *On appelle "résolvante" l'application  $(t_0, x_0, t) \mapsto \varphi_{t_0}^t(x_0)$  qui, à un temps initial  $t_0$ , une donnée initiale  $x_0$  et un temps  $t$ , associe la valeur au temps  $t$  de la solution maximale qui vaut  $x_0$  au temps  $t_0$ . Le domaine de définition de la résolvante est un ouvert  $\mathcal{R}$  de  $\mathcal{U} \times \mathbb{R}$ .*

On a donc

$$\frac{d\varphi_{t_0}^t(x_0)}{dt} = X(t, \varphi_{t_0}^t(x_0)), \quad \varphi_{t_0}^{t_0}(x_0) = x_0.$$

De l'unicité, on déduit la propriété fondamentale de la résolvante : partout où elle est définie, on a la formule

$$\varphi_{t_1}^t \circ \varphi_{t_0}^{t_1} = \varphi_{t_0}^t.$$

On en déduit que  $\varphi_{t_0}^t$  est une bijection sur son image, de bijection réciproque  $\varphi_t^{t_0}$ . De plus, on déduit par exemple du théorème de continuité par rapport aux paramètres du point fixe d'une famille de contractions dont la constante de Lipschitz est localement bornée par une même constante  $k < 1$ , que l'unique solution locale dépend continuellement de la donnée initiale  $x_0$  et donc que l'application

résolvante est continue sur son domaine de définition. Les  $\varphi_{t_0}^t$  sont donc des homéomorphismes. C'est cette famille d'homéomorphismes qu'on appelle souvent le *flot* de l'équation différentielle. D'aucuns réservent ce mot pour le cas autonome où  $\varphi_{t_0}^t$  ne dépend que de la différence  $\theta = t - t_0$  et est alors noté  $\varphi_\theta$ . On définit ainsi un *groupe local à un paramètre* : là où la formule a un sens, on a  $\varphi_{\theta_1} \circ \varphi_{\theta_2} = \varphi_{\theta_1 + \theta_2}$ .

**Régularité.** Sous les seules hypothèses du théorème de Cauchy-Lipschitz, on peut montrer (voir le livre de Chaperon) que l'application résolvante est localement lipschitzienne de  $\mathcal{R}$  dans  $E$  et donc que les  $\varphi_{t_0}^t$  sont des homéomorphismes de Lipschitz (bien entendu,  $\varphi_{t_0}^t$  est  $C^1$  par rapport à  $t$  ou  $t_0$ ). Plus généralement, on a le

**Théorème 3** *Si les applications  $(t, x) \mapsto X$  et  $(t, x) \mapsto \frac{\partial X}{\partial x}$  sont de classe  $C^k$  (par ex. si  $X$  est de classe  $C^{k+1}$ ), l'application résolvante  $(t_0, x_0, t) \mapsto \varphi_{t_0}^t(x_0)$  est de classe  $C^{k+1}$ .*

On raisonne par récurrence. Basée sur le théorème des fonctions implicites, la démonstration de Robbin nous assure que, si  $X$  est de classe  $C^1$ , l'application qui à  $x_0$  fait correspondre la solution  $t \mapsto \varphi_{t_0}^t(x_0)$ , ou plutôt sa restriction, considérée comme élément de l'espace de Banach des fonctions de classe  $C^1$  de  $[t_0 - \epsilon, t_0 + \epsilon]$  à valeurs dans  $E$ , qui valent  $x_0$  en  $t_0$ , est de classe  $C^1$ . On en déduit (exercice) que l'application résolvante est également de classe  $C^1$  (on connaissait a priori la dépendance  $C^1$  par rapport à  $t$  et donc par rapport à  $t_0$  puisque  $\varphi_{t_0}^t = (\varphi_t^{t_0})^{-1}$ ). Si l'on suppose seulement  $X$  continue ainsi que  $\frac{\partial X}{\partial x}$  (hypothèses du théorème pour  $k = 0$ ), il faut travailler un peu plus (voir Chaperon). On raisonne alors par récurrence, i.e. en supposant l'application résolvante de classe  $C^k$ . La dérivée  $\frac{\partial \varphi_{t_0}^t(x_0)}{\partial t} = X(t, \varphi_{t_0}^t(x_0))$  est donc de classe  $C^k$  et l'argument ci-dessus montre qu'il en est de même de la dérivée  $\frac{\partial \varphi_{t_0}^t(x_0)}{\partial t_0}$ . Enfin, la dérivée  $\frac{\partial \varphi_{t_0}^t(x_0)}{\partial x_0}$  est de classe  $C^k$  car le couple  $(\varphi_{t_0}^t(x_0), \frac{\partial \varphi_{t_0}^t(x_0)}{\partial x_0})$  est solution de l'équation différentielle (dans un ouvert de  $\mathbb{R} \times E \times L(E, E)$ ) :

$$\frac{d(x, y)}{dt} = (X(t, x), \frac{\partial X}{\partial x}(t, x)y),$$

qui satisfait aux hypothèses du théorème correspondant à  $k - 1$ .

### Flots globaux.

**Théorème 4 (Un cas d'existence globale)** *Soit  $X(t, x)$ , défini et continu sur  $\mathbb{R} \times E$  tout entier, vérifiant les hypothèses du théorème de Cauchy-Lipschitz. On suppose qu'à  $t$  fixé, le champ de vecteurs  $x \mapsto X(t, x)$  est globalement lipschitzien, avec une constante de Lipschitz  $k_t$  uniformément bornée si  $t$  reste dans un intervalle compact. Dans ce cas, l'unique solution du problème de Cauchy est définie sur  $\mathbb{R}$  tout entier. Un exemple typique est celui des équations linéaires non autonomes  $\frac{dx}{dt} = A(t)x$ , où  $t \mapsto A(t)$  est continue de  $\mathbb{R}$  dans  $L(E, E)$ .*

# Calcul différentiel et équations différentielles

A. Chenciner, résumé de cours n<sup>o</sup>22

## FLOTS (suite)

ATTENTION : le domaine de définition de  $\varphi_{t_0}^t$  dépend de  $t_0$  et  $t$ . Un exemple simple est donné par le champ  $X(x_1, x_2) = (1, 0)$  (encore noté  $\frac{\partial}{\partial x_1}$ ) sur  $\Omega = \mathbb{R}^2 \setminus \{0\}$ . On voit facilement que le domaine de définition de  $\varphi_{t_0}^t$  est le complémentaire dans  $\mathbb{R}^2$  de l'intervalle fermé  $[t_0 - t, 0]$  (si  $t > t_0$ ,  $[0, t_0 - t]$  sinon). On peut rendre cet exemple moins artificiel en envoyant à l'infini le "trou" à l'origine, par exemple en considérant l'application  $f$  de  $\mathbb{R}^2$  sur  $\mathbb{R}^2 \setminus \{0\}$  définie par  $f(x, y) = (e^x \cos y, e^x \sin y)$  (qui se factorise en un "difféomorphisme" du cylindre  $\mathbb{R} \times \mathbb{R}/\mathbb{Z}$  sur  $\mathbb{R}^2 \setminus \{0\}$ ) et en remarquant que le champ de vecteurs  $X(x, y) = (e^{-x} \cos y, e^{-x} \sin y)$  a pour image directe par  $f$  (voir feuille 18, bien que l'application  $f$  ne soit un difféomorphisme que localement, l'image directe est bien définie à cause de la périodicité en  $y$ ) le champ  $\frac{\partial}{\partial x_1}$ .

**Le cas autonome.** Si le champ de vecteurs  $X(x)$  ne dépend pas de  $t$  et si  $t \mapsto x(t)$  est solution de  $\frac{dx}{dt} = X(x)$ , il en est de même de  $t \mapsto x(t - t_0)$ . Et donc,  $\varphi_{t_0}^t = \varphi_{t-t_0}$ . On appelle alors *flot* de l'équation la famille des difféomorphismes  $\varphi_t$ . Celle-ci est un *groupe (local) à un paramètre de difféomorphismes (locaux) de  $E$* , ce qui signifie que  $\varphi_{t_1} + \varphi_{t_2} = \varphi_{t_1+t_2}$ .

Lorsque  $X$  est défini sur  $E$  tout entier et que les solutions maximales sont également définies sur  $\mathbb{R}$  tout entier, on obtient un *groupe à un paramètre de difféomorphismes de  $E$* , c'est-à-dire un homomorphisme de groupes (ou représentation non linéaire de  $(\mathbb{R}, +)$ ) :

$$(\mathbb{R}, +) \rightarrow (\text{Diff}^{k+1}(E, E), \text{composition}).$$

On dit que le champ de vecteurs autonome  $X$  est le *générateur infinitésimal* du groupe à un paramètre (ou du flot)  $(\varphi_t)_{t \in \mathbb{R}}$ . Par exemple, le champ de vecteurs  $X(x_1, x_2) = (-x_2, x_1)$  dans  $\mathbb{R}^2$  est le générateur infinitésimal du groupe à un paramètre des rotations (après identification de  $(x_1, x_2) \in \mathbb{R}^2$  à  $z = x_1 + ix_2 \in \mathbb{C}$ , l'équation s'écrit  $\frac{dz}{dt} = iz$  et donc  $\varphi_t(z) = e^{it}z$ ).

**De  $\mathbb{R}$  à  $\mathbb{Z}$ .** Si l'on n'observe le système qu'à des intervalles de temps discrets  $\tau$  (stroboscopie), on obtient une représentation du groupe additif  $\mathbb{Z}$ , qui revient à l'itération du difféomorphisme  $\varphi_\tau$ . Ainsi le flot, qui est la version globale (ou intégrée) de l'objet infinitésimal qu'est le champ de vecteurs apparaît-il comme la version continue de l'itération d'un difféomorphisme. C'est le point de vue "Systèmes dynamiques" sur la théorie qualitative des équations différentielles (voir mon article dans l'*Encyclopedia Universalis*).

## Dérivée de Lie : intégrales premières et fonctions de Lyapunov.

**Définition 1** *Étant données une fonction dérivable  $f : \Omega \rightarrow \mathbb{R}$  définie sur un ouvert de  $E$  et un champ de vecteurs autonome  $X$  sur  $\Omega$ , on appelle dérivée de Lie de  $f$  suivant  $X$  la fonction  $x \mapsto df(x)X(x)$ . On la note  $L_X f$ , ou  $df \cdot X$ , ou  $X \cdot f$ , cette dernière notation indiquant clairement que  $X$  est considéré comme un opérateur différentiel agissant sur  $f$ .*

La dérivée de Lie n'est rien d'autre que la dérivée le long des solutions de l'équation différentielle  $\frac{dx}{dt} = X(x)$  ; en effet, si  $x(t)$  est une solution,

$$\frac{d}{dt}(f(x(t))) = df(x(t))\frac{dx}{dt}(t) = df(x(t))X(x(t)).$$

**Définition 2** *Une fonction constante (resp. monotone) le long de solutions, i.e. telle que  $L_X f \equiv 0$  (resp.  $L_X f$  de signe constant) est appelée une intégrale première (resp. une fonction de Lyapunov) de l'équation différentielle.*

**Exemple : le pendule.** Dans l'espace des phases (i.e. l'espace des couples "position-vitesse"  $x = (\theta, v)$ ), considérons l'équation du second ordre

$$\frac{d^2\theta}{dt^2} + \alpha \frac{d\theta}{dt} + \omega^2 \sin \theta = 0, \text{ i.e. } \frac{dx}{dt} = X(x), \text{ où } X(\theta, v) = (v, -\alpha v - \omega^2 \sin \theta).$$

L'énergie totale

$$H(\theta, v) = \frac{1}{2}v^2 - \omega^2 \cos \theta,$$

somme de l'énergie cinétique  $\frac{1}{2}v^2$  et de l'énergie potentielle  $-\omega^2 \cos \theta$  est une intégrale première si  $\alpha = 0$  (pendule conservatif) et une fonction de Lyapunov (au sens large car  $L_X H$  peut s'annuler) sinon. En effet,  $L_X H(\theta, v) = -\alpha v^2$ . Le coefficient  $\alpha$  représente un frottement s'il est positif, une excitation sinon.

## Evolution du volume sous un flot en dimension finie.

**Lemme 3** *Le flot d'un champ  $X(t, x)$  de classe  $C^k$  ( $k \geq 1$ ) sur  $\mathbb{R}^n$  vérifie*

$$\frac{d}{dt} \det(d\varphi_{t_0}^t(x_0)) = \det(d\varphi_{t_0}^t(x_0)) \text{trace} \frac{\partial X}{\partial x}(t, \varphi_{t_0}^t(x_0)).$$

La démonstration suit de la formule donnant la dérivée de la fonction déterminant (notée "det") en  $A \in GL(\mathbb{R}^n)$  :

$$d\det(A)\Delta = \det A \times \text{trace}(A^{-1}\Delta).$$

**Champs hamiltoniens.** Le pendule conservatif est un exemple de *champ hamiltonien*, i.e. d'un champ sur  $\mathbb{R}^{2n}$  (coordonnées  $x_1, \dots, x_n, y_1, \dots, y_n$ ), pouvant dépendre du temps, de la forme

$$\frac{dx_i}{dt} = \frac{\partial H}{\partial y_i}, \quad \frac{dy_i}{dt} = -\frac{\partial H}{\partial x_i}, \quad i = 1, \dots, n, \quad \text{où } H : \mathbb{R} \times E \rightarrow \mathbb{R}.$$

Il se déduit du gradient (euclidien) de  $x \mapsto H(t, x)$  par une opération de rotation de  $-\frac{\pi}{2}$  des projections de ce dernier sur chacun des plans  $(x_i, y_i)$ . On déduit du lemme ci-dessus que  $\det(d\varphi_{t_0}^t(x_0))$  est indépendant de  $t$ , donc égal à 1 puisque  $\varphi_{t_0}^{t_0} = Id$ . Le flot d'un champ hamiltonien sur  $\mathbb{R}^{2n}$  préserve donc le volume (i.e. la mesure de Lebesgue  $2n$ -dimensionnelle). Et celui d'un gradient ?

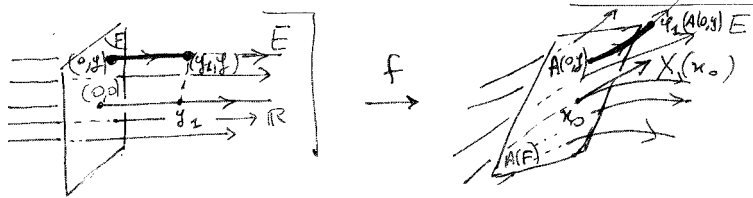
# Calcul différentiel et équations différentielles

A. Chenciner, résumé de cours n°23

## ÉTUDE LOCALE D'UN CHAMP DE VECTEURS.

**Théorème 1 (Redressement local, ou “flow box”).** Soit  $X$  un champ de vecteurs autonome sur (un ouvert de)  $E$ ,  $x_0 \in E$  tel que  $X(x_0) \neq 0$ . Il existe un difféomorphisme local, de classe  $C^{k+1}$  si le champ  $X$  est de classe  $C^{k+1}$ ,  $f : (\mathbb{R} \times F, (0, 0)) \ni (y_1, y) \cdots \mapsto x = f(y_1, y) \in (E, x_0)$ , tel qu'au voisinage de  $x_0$  on ait  $X = f_* \frac{\partial}{\partial y_1}$ , i.e.  $X(f(y_1, y)) = \frac{\partial f}{\partial y_1}(y_1, y)$ .

**Démonstration.** Ayant choisi un isomorphisme affine  $A : \mathbb{R} \times F \rightarrow E$  tel que  $A(0, 0) = x_0$  et  $A(\mathbb{R} \times \{0\}) = x_0 + \mathbb{R}X(x_0)$  (en dimension infinie, cela demande une démonstration), on pose  $f(y_1, y) = \varphi_{y_1}(A(0, y))$ , où  $\varphi_t$  est le flot de  $X$ . De la transversalité de  $X(x_0)$  et  $A(0, F)$ , il suit que  $f$  est un difféomorphisme local.



**Exercice.** En déduire que, par un changement de coordonnées local dépendant du temps, on peut annuler un champ  $X(t, x)$  au voisinage d'un point donné.

**Linéarisation.** Au voisinage d'un équilibre  $x_0$ , i.e. lorsque  $X(x_0) = 0$ , la première approximation est donnée par la *linéarisation du champ*, c'est-à-dire le champ linéaire  $\xi \mapsto dX(x_0)\xi$ . Il est donc important de comprendre les portraits de phase des champs linéaires (voir la feuille décrivant le cas où  $E = \mathbb{R}^2$ ).

**Stabilité.** Soit  $\frac{dx}{dt} = X(x)$  une équation autonome dans  $\mathbb{R}^n$  et  $x_0$  une position d'équilibre, c'est-à-dire un point de  $\mathbb{R}^n$  tel que  $X(x_0) = 0$ .

**Définition 2 (Stabilité au sens de Lyapunov).** La position d'équilibre  $x_0$  est dite stable au sens de Lyapunov si, pour tout  $\epsilon > 0$ , il existe  $\delta > 0$  tel que toute solution  $x(t)$  telle que  $\|x(0) - x_0\| < \delta$  soit définie sur tout  $\mathbb{R}_+$  et vérifie  $\|x(t) - x_0\| < \epsilon$  pour tout  $t > 0$ .

**Définition 3 (Stabilité asymptotique).** La position d'équilibre  $x_0$  est dite asymptotiquement stable si elle est stable au sens de Lyapunov et si de plus toute solution  $x(t)$  telle que  $\|x(0) - x_0\|$  soit assez petit vérifie  $\lim_{t \rightarrow +\infty} x(t) = x_0$ .



La contemplation des portraits de phase des champs de vecteurs linéaires autonomes dans le plan rend naturel le théorème suivant :

**Théorème 4** Soit  $\frac{dx}{dt} = X(x)$  une équation différentielle de classe  $C^{k+1}$  avec  $k \geq 1$ ,  $x_0$  une position d'équilibre et  $A = dX(x_0)$ . Si toutes les valeurs propres de  $A$  ont leur partie réelle strictement négative, la position d'équilibre  $x_0$  est asymptotiquement stable

1) On commence par montrer le théorème pour l'équation "linéarisée en  $x_0$ ",  $\frac{d\xi}{dt} = A\xi$ , obtenue en remplaçant  $X(x)$  par la partie affine  $dX(x_0)(x - x_0)$  de son développement de Taylor et en posant  $\xi = x - x_0$ . On construit pour cela une fonction de Lyapunov :

i) le cas diagonalisable. Si  $\{\vec{v}_1, \dots, \vec{v}_n\}$  est une base propre (éventuellement complexe) de  $A$ , la forme quadratique  $Q : \mathbb{C}^n \rightarrow \mathbb{R}$

$$Q\left(\sum_{i=1}^n z_i \vec{v}_i\right) = \sum_{i=1}^n z_i \bar{z}_i$$

convient. En effet, si  $\sum_{i=1}^n z_i(t) \vec{v}_i$  est une solution (réelle ou complexe),

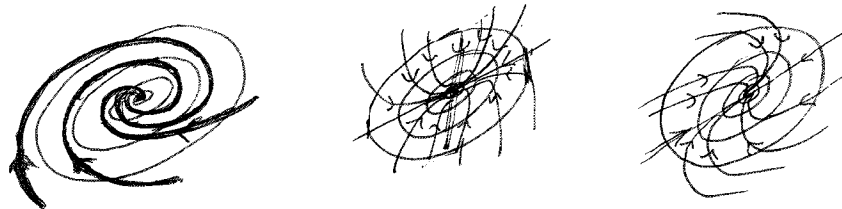
$$\frac{d}{dt} \left( Q\left(\sum_{i=1}^n z_i(t) \vec{v}_i\right) \right) = 2 \sum_{i=1}^n \operatorname{Re} \lambda_i |z_i|^2.$$

Puisque  $A$  est réelle, il est naturel d'associer à chaque couple  $\lambda_i, \bar{\lambda}_i$  de valeurs propres conjuguées un couple de vecteurs propres conjugués  $\vec{v}_i, \bar{\vec{v}}_i$ , ce qui revient à identifier  $\mathbb{R}^n$  à  $\mathbb{C}^p \times \mathbb{R}^q$  s'il y a  $p$  couples  $\lambda_i, \bar{\lambda}_i$  et  $q$  valeurs propres réelles.

Si, pour tout  $i$ ,  $\operatorname{Re} \lambda_i < -m < 0$ ,  $\frac{d}{dt} (Q(\sum_{i=1}^n z_i(t) \vec{v}_i)) < -nm Q(\sum_{i=1}^n z_i(t) \vec{v}_i)$ , et donc  $Q(\sum_{i=1}^n z_i(t) \vec{v}_i) < e^{-nmt}$ , ce qui prouve la stabilité de l'équilibre  $\xi = 0$ .

ii) le cas non diagonalisable. Une remarque nous ramène pratiquement au cas précédent : quel que soit  $\epsilon > 0$ , il existe une base dans laquelle la matrice de  $A$  est triangulaire avec des termes non diagonaux majorés par  $\epsilon$  (exercice !).

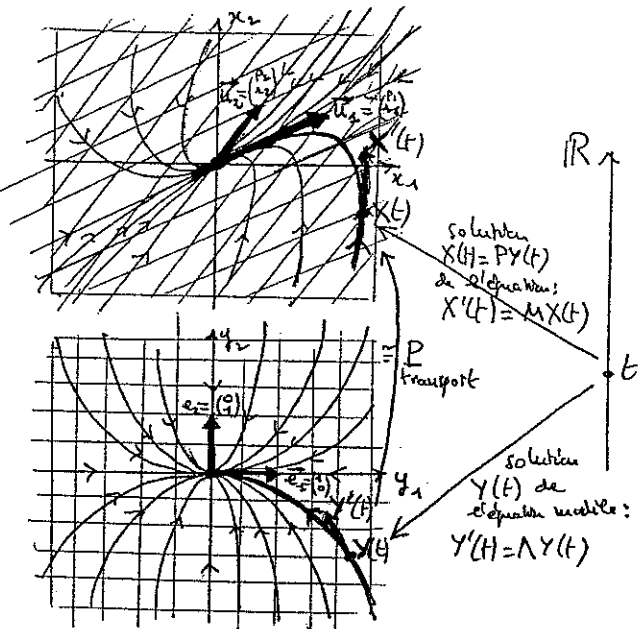
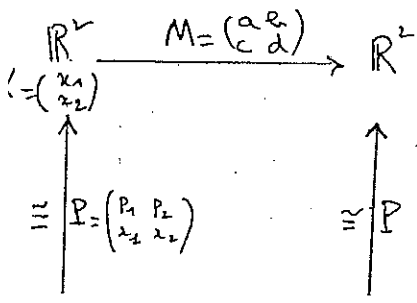
2) Il reste à remarquer que, suffisamment près de la singularité  $x_0$ , les termes non linéaires ne remettent pas en question la propriété de  $Q$  d'être une fonction de Lyapunov. Ainsi, les courbes intégrales issues d'un petit voisinage de  $x_0$  restent dans ce petit voisinage tant qu'elles sont définies. La borne  $r_0/M$  du temps de définition de la solution donné par le théorème de Cauchy-Lipschitz est donc valable en tout point d'une telle courbe intégrale, ce qui montre que celle-ci est définie pour tout temps  $t \in \mathbb{R}_+$  et termine la démonstration du théorème.



$$X' = MX$$

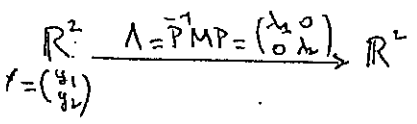
CAS RÉEL DIAGONALISABLE

$$\begin{cases} x_1' = ax_1 + bx_2 \\ x_2' = cx_1 + dx_2 \end{cases} \textcircled{1}$$



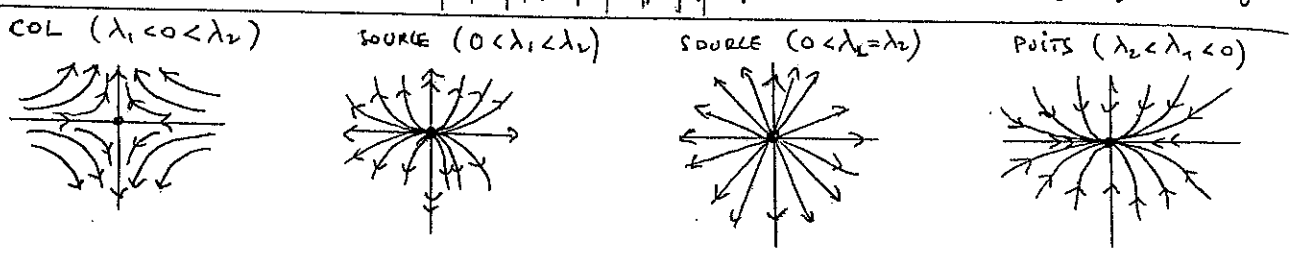
Niveau où habite l'équation différentielle structurée.

Niveau où habite l'équation "modèle".



$$Y' = \Lambda Y$$

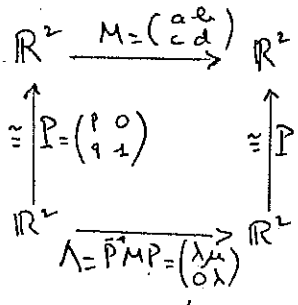
$$\begin{cases} y_1' = \lambda_1 y_1 \\ y_2' = \lambda_2 y_2 \end{cases}$$



$$X' = MX$$

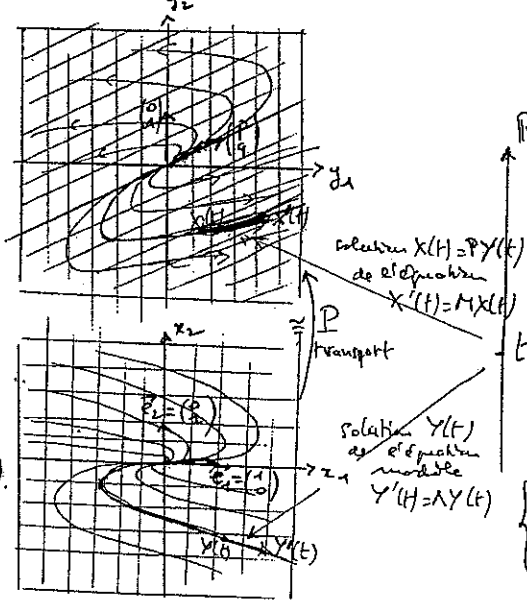
CAS RÉEL NON DIAGONALISABLE

$$\begin{cases} x_1' = ax_1 + bx_2 \\ x_2' = cx_1 + dx_2 \end{cases} \textcircled{2}$$



On a supposé que l'unique direction propre n'est pas verticale ce qui a permis de compléter la base par  $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$  (choix purement arbitraire).

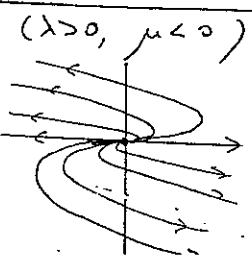
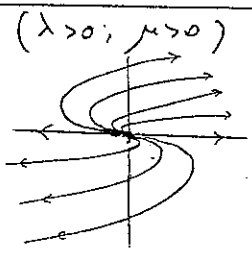
$$Y' = \Lambda Y$$



Niveau où habite l'équation différentielle structurée.

Niveau où habite l'équation modèle.

$$\begin{cases} y_1' = \lambda y_1 + \mu y_2 \\ y_2' = \lambda y_2 \end{cases}$$



Hélice entre source et puits spirale

$\textcircled{3}$

$$\begin{cases} y_1 = (\lambda_1 + \mu k_2 t) e^{\lambda t} \\ y_2 = k_2 e^{\lambda t} \end{cases}$$

qui se résout par variation des constantes :

$$\begin{cases} y_1 = (k_1 + \mu k_2 t) e^{\lambda t} \\ y_2 = k_2 e^{\lambda t} \end{cases}$$



# Calcul différentiel et équations différentielles

*A. Chenciner, résumé de cours n°24*

**UNE APPLICATION DU TH. DE CAUCHY-LIPSCHITZ : LE LEMME DE MORSE.**