HAL open science

Calcul différentiel et équations différentielles

Alexandre Munnier

To cite this version:

Alexandre Munnier. Calcul différentiel et équations différentielles. Licence. France. 2021. hal-03135214

HAL Id: hal-03135214

https://hal.science/hal-03135214

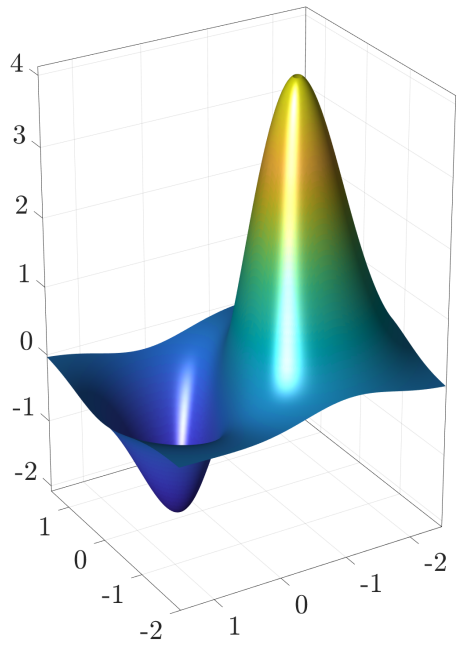
Submitted on 8 Feb 2021

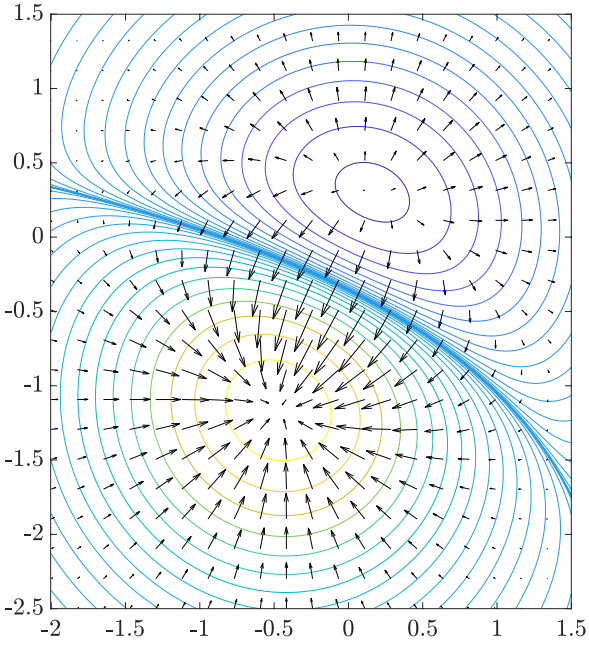
HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Calcul différentiel et Équations différentielles

Alexandre Munnier
alexandre.munnier@univ-lorraine.fr





7 février 2021

Ce cours est enseigné depuis 2018 à la faculté des sciences de l'Université de Lorraine à Nancy. Il est en grande partie tiré du livre :

- Jean–Christophe Yoccoz. *Cours de topologie : calcul différentiel, équations différentielles pour la licence MAF.* Université Paris-Sud, 1994,

disponible gratuitement ici.

La section portant sur les équations différentielles est également inspirée de :

- Coddington E-.A et Levinson L. *Theory of Ordinary Differential Equations.* McGraw-Hill Book Company, Inc., New York-Toronto-London, 1955. xii+429 pp.
- Zuily Cl. et Queffélec H. *Éléments d'analyse pour l'agrégation.* Masson, Paris-Milan-Barcelone, 1995. 478 pp.

et dans une moindre mesure de :

- Coddington E-. A. et Carlson R. *Linear Ordinary Differential Equations.* Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997. xii+341 pp. ISBN 0-89871-388-9.
- Reinhard H. *Equations différentielles, fondements et applications.* Dunod, Paris, 1982. xiv+446 pp.

De nombreux exercices ont été glanés à droite à gauche sur internet (sur le site exo7 en particulier). Je prie les personnes qui reconnaitraient des exercices de leur cru de bien vouloir m'excuser de ne pas les citer.

Tous les graphes ont été réalisés avec le logiciel MATLAB.

Table des matières

1 Prérequis au calcul différentiel **3**
1.1 Normes et espaces vectoriels normés . . . 3
1.2 Applications continues . . . 4
1.3 Applications linéaires . . . 4
1.4 Applications bilinéaires . . . 5
1.5 Intégration d'une fonction à valeurs dans un evn . . . 6

2 Différentielle d'une fonction **9**
2.1 Différentiabilité . . . 9
2.2 Opérations sur les différentielles . . . 11
2.3 Dérivées directionnelles, dérivées partielles . . . 12
2.4 Représentations graphiques . . . 17

3 Théorème de la moyenne **21**
3.1 Rappels sur les fonctions de $\mathbb{R}$ dans $\mathbb{R}$ . . . 21
3.2 Fonctions de E dans F . . . 21
3.3 Remarques . . . 24

4 Théorèmes d'inversion **27**
4.1 Théorème d'inversion locale . . . 27
4.2 Théorème des fonctions implicites . . . 30

5 Différentielles d'ordre supérieur **35**
5.1 Différentielle seconde . . . 35
5.2 Formule de Taylor . . . 38

6 Points critiques et extrema **41**
6.1 Points réguliers et points critiques . . . 41
6.2 Formes quadratiques . . . 42
6.3 Extrema . . . 44
6.4 Extrema liés . . . 46
6.4.1 Points réguliers . . . 46
6.4.2 Points critiques de H sur $\mathcal{M}$ . . . 48

7 Équations différentielles **53**
7.1 Définitions et notations . . . 53
7.2 Réduction de l'ordre . . . 54
7.3 Équations différentielles autonomes . . . 55
7.4 Existence et unicité locale de solutions . . . 55
7.5 Unicité globale, solutions maximales . . . 58
7.6 Cas particulier des edo linéaires : Premières propriétés . . . 62
7.7 Continuité par rapport aux conditions initiales . . . 63
7.8 Différentiabilité par rapport aux conditions initiales . . . 67
7.9 Flot d'une edo autonome . . . 69
7.10 Équations différentielles linéaires . . . 73
7.10.1 Rappels d'algèbre linéaire, exponentielle de matrices . . . 73
7.10.2 Équations linéaires homogènes . . . 76
7.10.3 Équations linéaires avec second membre . . . 81
7.10.4 Équations linéaires à cœfficients constants . . . 82
7.10.5 Équations linéaires d'ordre n à cœfficients constants . . . 85
7.10.6 L'équation de Sturm–Liouville . . . 89
7.11 Stabilité des équations différentielles autonomes . . . 94
7.11.1 Stabilité des edo linéaires à cœfficients constants . . . 97
7.11.2 Stabilité des edo autonomes non linéaires . . . 102
7.12 Portrait de phase . . . 105

Chapitre 1

Prérequis au calcul différentiel

1.1 Normes et espaces vectoriels normés

Soient E un espace vectoriel (ev) de dimension finie n (que l'on pourra donc identifier avec $\mathbb{R}^n$). Une norme sur E est une application

$$\begin{array}{rcl} \|\cdot\| : E & \to & \mathbb{R}_+ \\ x & \mapsto & \|x\| \end{array}$$

vérifiant :

1. Pour tout $x \in E$, $\|x\| = 0$ entraîne $x = 0$;
2. Pour tout $(x,\lambda) \in E \times \mathbb{R}$, $\|\lambda x\| = |\lambda|\|x\|$;
3. Pour tout $(x,y) \in E \times E$, $\|x+y\| \leqslant \|x\| + \|y\|$ (inégalité triangulaire).

Une norme est qualifiée d'euclidienne lorsqu'elle est définie par :

$$\|x\| = \sqrt{\langle x,x\rangle} \quad \text{pour tout } x \in E,$$

où $\langle\cdot,\cdot\rangle$ est un produit scalaire sur E.

On peut par exemple munir $\mathbb{R}^n$ des normes usuelles suivantes, définies pour tout $p \in [1,\infty[$ et tout $x = (x_1,\ldots,x_n) \in \mathbb{R}^n$ par :

$$\|x\|_p = \left(\sum_{j=1}^{n} |x_j|^p\right)^{1/p} \qquad \text{et} \qquad \|x\|_\infty = \max_{1\leqslant j\leqslant n} |x_j|.$$

La donnée d'un couple $(E,\|\cdot\|)$ s'appelle un espace vectoriel normé. Rappelons que :

Proposition 1.1. *Deux normes $\|\cdot\|_1$ et $\|\cdot\|_2$ sur un ev E de dimension finie sont toujours équivalentes c'est à dire qu'il existe deux constantes C_1 et C_2 strictement positives telles que :*

$$C_1\|x\|_1 \leqslant \|x\|_2 \leqslant C_2\|x\|_1 \qquad \textit{pour tout } x \in E.$$

Soit $a \in E$ et $(x_n)_n$ une suite dans E. On dira que la suite $(x_n)_n$ converge vers a dans E lorsque :

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

On remarque que, suivant la proposition 1.1, la notion de convergence ne dépend pas du choix de la norme. Ce résultat est faux dans un ev de dimension infinie : Toutes les normes ne sont plus équivalentes et donc la notion de convergence dépend du choix de la norme.

1.2 Applications continues

Définition 1.1 (Continuité)**.** *Soient E et F deux ev normés et $f : E \to F$. On dit que f est continue en un point $a \in E$ lorsque, pour tout $\varepsilon > 0$, il existe $\eta > 0$ tel que :*

$$\text{Pour tout } x \in E \quad \Big(\|x-a\|_E < \eta\Big) \quad \Longrightarrow \quad \Big(\|f(x)-f(a)\|_F < \varepsilon\Big).$$

Proposition 1.2. *Toute fonction construite à partir de fonctions continues par combinaison linéaire, produit, quotient (avec un dénominateur qui ne s'annule pas) ou produit de composition est encore continue.*

Exercice 1.1. Soit $f : \mathbb{R}^2 \to \mathbb{R}^2$ définie pour tout $x = (x_1,x_2)$ par :

$$f(x) = \begin{pmatrix} \dfrac{1}{2}\dfrac{x_1x_2^2}{x_1^2+x_2^2} \\ -\dfrac{1}{3}\dfrac{x_1^2x_2}{x_1^2+x_2^2} \end{pmatrix} \quad \text{si } x \neq (0,0) \qquad \text{et} \qquad f(0,0) = (0,0).$$

Vérifier que f est continue en $(0,0)$.

<u>Solution :</u> Pour tout $x = (x_1,x_2) \neq (0,0)$, on a :

$$\|f(x,y) - f(0,0)\|_\infty = \max\Big\{\Big|\frac{1}{2}\frac{x_1x_2^2}{x_1^2+x_2^2}\Big|, \Big|\frac{1}{3}\frac{x_1^2x_2}{x_1^2+x_2^2}\Big|\Big\},$$

et, notant $\bar{x}_1 = x_1/\|x\|_2$ et $\bar{x}_2 = x_2/\|x\|_2$:

$$\Big|\frac{x_1x_2^2}{x_1^2+x_2^2}\Big| \leqslant \|x\|_2|\bar{x}_1\bar{x}_2^2| \leqslant \|x\|_2 \quad \text{et} \quad \Big|\frac{x_1^2x_2}{x_1^2+x_2^2}\Big| \leqslant \|x\|_2|\bar{x}_1^2\bar{x}_2| \leqslant \|x\|_2,$$

car $|\bar{x}_1| < 1$ et $|\bar{x}_2| < 1$. On en déduit que :

$$\|f(x,y) - f(0,0)\|_\infty \leqslant \frac{1}{2}\|x\|_2,$$

et donc que f est continue en $(0,0)$.

Parmi les applications continues, on distinguera en particulier :

Définition 1.2 (Applications lipschitziennes)**.** *Soient E et F deux ev normés. Un application $f : E \to F$ est dite lipschitzienne de constante de Lipschitz $k \geqslant 0$ lorsque :*

$$\text{Pour tout } x,y \in E, \quad \|f(x)-f(y)\|_F \leqslant k\|x-y\|_E.$$

1.3 Applications linéaires

Soit E et F deux ev de dimensions finies. Rappelons que :

Proposition 1.3. *Toute application linéaire $\ell : E \to F$ est lipschitzienne.*

On note $\mathscr{L}(E,F)$ l'ensemble des applications linéaires de E dans F. Si E est de dimension n et F de dimension m alors $\mathscr{L}(E,F)$ est de dimension $n \times m$. Soit $\{e_k, k = 1,\ldots,n\}$ une base de E et $\{f_j, j = 1,\ldots,m\}$ une base de F, alors pour tout $\ell \in \mathscr{L}(E,F)$ il existe des cœfficients $\alpha_{kj} \in \mathbb{R}$ ($k = 1,\ldots,n$ et $j = 1,\ldots,m$) tels que :

$$\ell(e_k) = \sum_{j=1}^{m} \alpha_{kj} f_j.$$

Si $x = x_1e_1 + \ldots + x_ne_n$ dans E et $\ell(x) = y_1f_1 + \ldots + y_nf_n$ dans F alors :

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1m} \\ \alpha_{21} & \alpha_{22} & \ldots & \alpha_{2m} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \ldots & \alpha_{nm} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

À partir de deux normes $\|\cdot\|_E$ et $\|\cdot\|_F$ sur E et F on peut définir une norme $\|\cdot\|_{\mathscr{L}(E,F)}$ (appelée norme subordonnée) sur $\mathscr{L}(E,F)$ de la façon suivante. Pour tout $\ell \in \mathscr{L}(E,F)$, posons :

$$\begin{aligned} \|\ell\|_{\mathscr{L}(E,F)} &= \sup_{\substack{x\in E \\ x\neq 0}} \frac{\|\ell(x)\|_F}{\|x\|_E} = \sup_{\substack{x\in E \\ \|x\|_E=1}} \|\ell(x)\|_F \\ &= \inf\big\{k > 0 \text{ tel que pour tout } x \in E, \quad \|\ell(x)\|_F \leqslant k\|x\|_E\big\}. \end{aligned}$$

La norme de ℓ est donc sa plus petite constante de Lipschitz. En tant qu'espace vectoriel de dimension finie, on pourrait considérer sur $\mathscr{L}(E,F)$ n'importe quelle norme $\|\cdot\|_p$ ($1 \leqslant p \leqslant \infty$) définie plus haut. L'avantage de la norme $\|\cdot\|_{\mathscr{L}(E,F)}$ se voit lorsque l'on considère un produit de composition. En effet, soit G un autre ev normé. Si $\ell_1 \in \mathscr{L}(E,F)$ et $\ell_2 \in \mathscr{L}(F,G)$, alors :

$$\|\ell_2 \circ \ell_1\|_{\mathscr{L}(E,G)} \leqslant \|\ell_1\|_{\mathscr{L}(E,F)}\|\ell_2\|_{\mathscr{L}(F,G)}.$$

En dimension infinie, la proposition 1.3 est fausse : il existe des applications linéaires qui ne sont pas continues. On ajoute alors la continuité dans la définition de l'espace $\mathscr{L}(E,F)$.

1.4 Applications bilinéaires

Soient E, F et G trois ev. Une application $b : E \times F \to G$ est bilinéaire lorsque pour tout $x \in E$ et pour tout $y \in F$, les applications $b(x,\cdot) : F \to G$ et $b(\cdot,y) : E \to G$ sont linéaires.

Proposition 1.4. *Soient E, F et G trois ev normés. Toute application bilinéaire $b : E \times F \to G$ est continue. De plus, il existe une constante $k \geqslant 0$ telle que, pour tout $(x,y) \in E \times F$:*

$$\|b(x,y)\|_G \leqslant k\|x\|_E\|y\|_F.$$

On note $\mathscr{L}_2(E \times F,G)$ l'ensemble des applications bilinéaires de $E \times F$ dans G et pour toute application b dans cet espace on définit la norme subordonnée :

$$\begin{aligned} \|b\|_{\mathscr{L}_2(E\times F,G)} &= \sup_{\substack{(x,y)\in E\times F \\ x,y\neq 0}} \frac{\|b(x,y)\|_G}{\|x\|_E\|y\|_F}, \\ &= \inf\big\{k > 0 \text{ tel que pour tout } (x,y) \in E \times F, \quad \|b(x,y)\|_G \leqslant k\|x\|_E\|y\|_F\big\}. \end{aligned}$$

Alors $\left(\mathscr{L}_2(E\times F,G),\|\cdot\|_{\mathscr{L}_2(E\times F,G)}\right)$ est un espace vectoriel normé.

Soit l'application :

$$\begin{array}{rcl} T:\mathscr{L}(E,\mathscr{L}(F,G)) & \to & \mathscr{L}_2(E\times F,G) \\ B & \mapsto & b, \end{array}$$

tels que pour tout $(x,y)\in E\times F$, $b(x,y)=B(x)(y)$.

Proposition 1.5. *L'application linéaire T est une isométrie, c'est à dire qu'elle est bijective et que, pour tout $B\in\mathscr{L}(E,\mathscr{L}(F,G))$, on a :*

$$\|B\|_{\mathscr{L}(E,\mathscr{L}(F,G))}=\|TB\|_{\mathscr{L}_2(E\times F,G)}.$$

Démonstration. L'application T^{-1} est donnée par :

$$\begin{array}{rcl} T^{-1}:\mathscr{L}_2(E\times F,G) & \to & \mathscr{L}(E,\mathscr{L}(F,G)) \\ b & \mapsto & B \end{array}$$

où, pour tout $x\in E$, $B(x)=b(x,\cdot)$. D'autre part, pour tout B dans $\mathscr{L}(E,\mathscr{L}(F,G))$:

$$\begin{aligned} \|B\|_{\mathscr{L}(E,\mathscr{L}(F,G))} &= \sup_{\substack{x\in E\\ x\neq 0}}\frac{\|B(x)\|_{\mathscr{L}(F,G)}}{\|x\|_E} \\ &= \sup_{\substack{x\in E\\ x\neq 0}}\frac{1}{\|x\|_E}\left(\sup_{\substack{y\in F\\ y\neq 0}}\frac{\|B(x)(y)\|_G}{\|y\|_F}\right) \\ &= \sup_{\substack{(x,y)\in E\times F\\ x,y\neq 0}}\frac{\|B(x)(y)\|_G}{\|x\|_E\|y\|_F}=\|TB\|_{\mathscr{L}_2(E\times F,G)}, \end{aligned}$$

ce qui conclut la démonstration. □

On a bien sûr les mêmes résultats avec l'application :

$$\begin{array}{rcl} \mathscr{L}(F,\mathscr{L}(E,G)) & \to & \mathscr{L}_2(E\times F,G) \\ B & \mapsto & b, \end{array}$$

où pour tout $(x,y)\in E\times F$, $b(x,y)=B(y)(x)$.

Dans la suite du cours, nous ferons l'identification :

$$\mathscr{L}(E,\mathscr{L}(F,G))\sim\mathscr{L}_2(E\times F,G).$$

Comme pour les applications linéaires, la Proposition 1.4 est fausse en dimension infinie. On ajoute donc encore la continuité dans la définition de $\mathscr{L}_2(E\times F,G)$.

1.5 Intégration d'une fonction à valeurs dans un evn

Soit $(E,\|\cdot\|_E)$ un espace vectoriel normé, soit $[a,b]$ un invervalle de $\mathbb{R}$ et $f:[a,b]\to E$ une fonction. Pour tout σ, réel strictement positif tel que $\sigma<(b-a)$, on peut définir une subdivision $a=x_0^\sigma<x_1^\sigma<\ldots<x_{n_\sigma}^\sigma=b$ de l'intervalle $[a,b]$ de pas σ c'est à dire telle que

$$\sigma=\sup_{k=1,\ldots,n_\sigma}|x_k^\sigma-x_{k-1}^\sigma|.$$

Pour toute famille $T_\sigma = \{t_1^\sigma, \ldots, t_{n_\sigma}^\sigma\}$ de réels tels que $t_k^\sigma \in [x_{k-1}^\sigma, x_k^\sigma]$ (pour tout $k = 1, \ldots, n_\sigma$) on définit la somme de Riemann

$$S_\sigma(f) = \sum_{k=1}^{n_\sigma} (x_k^\sigma - x_{k-1}^\sigma) f(t_k^\sigma).$$

Définition 1.3. *Une fonction $f : [a,b] \to E$ est dite intégrable au sens de Riemann lorsqu'il existe $I(f) \in E$ vérifiant la propriété suivante :*
Pour tout $\varepsilon > 0$, il existe $0 < \delta < (b-a)$ tel que pour toute subdivision de pas $\sigma < \delta$ et toute famille de réelles $T_\sigma = \{t_1^\sigma, \ldots, t_{n_\sigma}^\sigma\}$ comme ci-dessus, on a :

$$\|S_\sigma(f) - I(f)\|_E < \varepsilon.$$

On note alors $I(f) = \int_a^b f(t)\,\mathrm{d}t$.

On utilisera le plus souvent la propriété suivante :

Proposition 1.6. *Tout fonction continue sur $[a,b]$ est intégrable au sens de Riemann.*

Via l'introduction d'une base de E, l'intégrale d'une fonction à valeurs dans un ev peut se déduire de l'intégration de fonctions à valeurs réelles.

Proposition 1.7. *Soit $f : [a,b] \to E$ une fonction continue et soit $\{e_1, \ldots, e_m\}$ une base de E. La fonction f se décompose dans cette base sous la forme :*

$$f(t) = \sum_{j=1}^{m} f_j(t) e_j \quad \textit{pour tout } t \in [a,b],$$

où chaque fonction f_j est une fonction continue de $[a,b]$ dans $\mathbb{R}$. On a alors :

$$\int_a^b f(t)\,\mathrm{d}t = \sum_{j=1}^{m} \left(\int_a^b f_j(t)\,\mathrm{d}t \right) e_j.$$

Chapitre 2

Différentielle d'une fonction

2.1 Différentiabilité

Soient E et F deux ev normés, $\mathscr{U}$ un ouvert de E, a un point de $\mathscr{U}$, et $f : \mathscr{U} \to F$ une fonction. On note $\mathscr{O} = \mathscr{U} - \{a\} = \{h \in \mathscr{U} \text{ tels que } a + h \in \mathscr{U}\}$ (c'est un ouvert qui contient 0_E).

Définition 2.1 (Développement limité à l'ordre 1)**.** *On dit que f est différentiable en a s'il existe une application $\ell \in \mathscr{L}(E,F)$ et une fonction $\varepsilon : \mathscr{O} \to F$ telles que :*

$$\lim_{h \to 0_E} \|\varepsilon(h)\|_F = 0 \quad \textit{et pour tout } h \in \mathscr{O}, \quad f(a+h) = f(a) + \ell(h) + \|h\|_E \varepsilon(h).$$

Définition 2.2 (Taux d'accroissement)**.** *On dit que f est différentiable en a s'il existe une application $\ell \in \mathscr{L}(E,F)$ telle que :*

$$\lim_{\substack{h \to 0_E \\ h \in \mathscr{O} \setminus \{0_E\}}} \frac{\|f(a+h) - f(a) - \ell(h)\|_F}{\|h\|_E} = 0.$$

Proposition 2.1. *Les définitions 2.1 et 2.2 sont équivalentes. L'application ℓ est unique (et dépend du point a). On l'appelle la différentielle de f en a et on la note $Df(a)$.*

Les définitions 2.1 et 2.2 sont encore valables en dimension infinie (il ne faut pas oublier que la continuité de ℓ découle alors de la définition de $\mathscr{L}(E,F)$).

Démonstration. Il est évident que si f est différentiable au sens de la définition 2.1 alors elle est aussi différentiable au sens de la définition 2.2. Pour la réciproque, il suffit de poser :

$$\begin{aligned} \varepsilon(h) &= \frac{1}{\|h\|_E}\Big(f(a+h) - f(a) - \ell(h)\Big) && \text{pour } h \in \mathscr{O} \setminus \{0_E\}, \\ &= 0 && \text{lorsque } h = 0_E. \end{aligned}$$

Montrons que ℓ est unique en utilisant par exemple la définition 2.1. Supposons qu'il existe ℓ_1 et ℓ_2 et ε_1 et ε_2 telles que, pour $k = 1{,}2$:

$$\lim_{h \to 0_E} \|\varepsilon_k(h)\|_F = 0 \quad \text{et pour tout } h \in \mathscr{O}, \quad f(a+h) = f(a) + \ell_k(h) + \|h\|_E \varepsilon_k(h).$$

On note $\bar{\ell} = \ell_2 - \ell_1$ et $\bar{\varepsilon} = \varepsilon_2 - \varepsilon_1$. En faisant la différence des deux égalités ci-dessus, on obtient que :

$$\bar{\ell}(h) + \|h\|_E \bar{\varepsilon}(h) = 0 \qquad \text{pour tout } h \in \mathscr{O}.$$

Soit $h \in E$. Alors pour $t > 0$ assez petit, th est dans $\mathscr{O}$. On a donc :

$$\bar{\ell}(th) + \|th\|_E \bar{\varepsilon}(th) = 0 \qquad \text{pour tout } t > 0 \text{ petit.}$$

On divise ensuite cette égalité par t et l'on fait tendre t vers 0 pour obtenir que $\bar{\ell}(h) = 0$, c'est à dire $\ell_1(h) = \ell_2(h)$. □

Remarque 2.1. ***1.*** *Si f est différentiable en a, on peut définir $Df(a)$ comme un élément de $\mathscr{L}(E,F)$. Si f est différentiable sur un ouvert $\mathscr{U}$ de E, on peut définir l'application :*

$$\begin{array}{rcl} Df : \mathscr{U} & \to & \mathscr{L}(E,F) \\ a & \mapsto & Df(a) \end{array}$$

Il faut faire très attention ici et ne pas confondre l'application Df (qui n'est à priori pas linéaire) avec l'application linéaire $Df(a)$.

2. *La différentielle est la "meilleure" approximation de la fonction $h \mapsto f(a+h) - f(a)$ par une application linéaire.*

3. *Considérons le cas $E = \mathbb{R}$. Soit donc $f : \mathbb{R} \to F$ différentiable en $a \in \mathbb{R}$. Selon la définition il existe $Df(a) \in \mathscr{L}(\mathbb{R},F)$ et $\varepsilon : \mathbb{R} \to F$ tels que :*

$$\lim_{t \to 0} \|\varepsilon(t)\|_F = 0 \qquad \text{et pour tout } t \in \mathbb{R}, \quad f(a+t) = f(a) + Df(a)(t) + |t|\varepsilon(t).$$

Mais $\mathscr{L}(\mathbb{R},F) = \{t \mapsto t\alpha, \alpha \in F\}$. Dans ce cas, on note $f'(a) \in F$ le vecteur de F tel que $Df(a)(t) = tf'(a)$. Le vecteur $f'(a)$ s'appelle la dérivée de f en a.

On vérifie facilement avec la définition 2.1 que :

Proposition 2.2. *Toute fonction differentiable en un point est continue en ce point.*

Terminons cette section avec une dernière définition.

Définition 2.3. *Soit f une fonction à valeurs dans F et différentiable sur un ouvert $\mathscr{U}$ de E. Si la différentielle :*

$$Df : \mathscr{U} \to \mathscr{L}(E,F),$$

est continue, on dit que la fonction f est continûment différentiable sur $\mathscr{U}$ ou de classe $\mathscr{C}^1$ sur $\mathscr{U}$.

Quelques exemples

1. Toute application constante est continûment différentiable, de différentielle nulle (la réciproque sera examinée plus loin).

2. Si $f \in \mathscr{L}(E,F)$ alors f est continûment différentiable et pour tout $a \in E$, $Df(a) = f$.

3. Si $f \in \mathscr{L}_2(E \times F,G)$ (E,F et G sont trois ev normés), alors f est continûment différentiable et pour tout $a = (a_1,a_2) \in E \times F$ et tout $h = (h_1,h_2) \in E \times F$:

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

La démonstration sera faite en TD.

2.2 Opérations sur les différentielles

Proposition 2.3 (Somme)**.** *Soient f_1 et f_2 deux applications de E dans F différentiables en un point a (respectivement de classe $\mathscr{C}^1$ sur un ouvert $\mathscr{U}$ de E) et α et β deux réels. Alors l'application $f = \alpha f_1 + \beta f_2$ est différentiable en a (respectivement de classe $\mathscr{C}^1$ sur $\mathscr{U}$) et*

$$Df(a) = \alpha Df_1(a) + \beta Df_2(a).$$

Démonstration. Laissée en exercice. □

Proposition 2.4 (Produit de composition)**.** *Soient E, F et G trois ev normés. Soit $\mathscr{U}$ un ouvert de E et $\mathscr{V}$ un ouvert de F et $f : \mathscr{U} \to F$ et $g : \mathscr{V} \to G$. Soit a un point de $\mathscr{U}$ tel que $b = f(a) \in \mathscr{V}$. Si f est différentiable en a (respectivement de classe $\mathscr{C}^1$ sur $\mathscr{U}$) et g est différentiable en b (respectivement de classe $\mathscr{C}^1$ sur $\mathscr{V}$) alors $g \circ f$ est différentiable en a (respectivement de classe $\mathscr{C}^1$ sur $\mathscr{U}$) et :*

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

Démonstration. Il existe un voisinage $\mathscr{U}_1$ de 0_E dans E tel que $\mathscr{U}_1 + \{a\} \subset \mathscr{U}$ et une fonction $\varepsilon_1 : \mathscr{U}_1 \to F$ tels que $\lim_{h \to 0_E} \varepsilon_1(h) = 0_F$ et pour tout h dans $\mathscr{U}_1$:

$$f(a+h) = f(a) + Df(a)h + \|h\|_E \varepsilon_1(h).$$

De même, il existe un voisinage $\mathscr{U}_2$ de 0_F dans F tel que $\mathscr{U}_2 + \{b\} \subset \mathscr{V}$ et une fonction $\varepsilon_2 : \mathscr{U}_2 \to G$ tels que $\lim_{k \to 0_F} \varepsilon_2(k) = 0_G$ et pour tout k dans $\mathscr{U}_2$:

$$g(b+k) = g(b) + Dg(b)k + \|k\|_F \varepsilon_2(k). \tag{2.1}$$

Soit $\mathscr{U}_3$ un voisinage de 0_E dans E tel que $\mathscr{U}_3 \subset \mathscr{U}_2$ et pour tout h dans $\mathscr{U}_3$:

$$Df(a)h + \|h\|_E \varepsilon_1(h) \in \mathscr{U}_2.$$

On a alors, en choisissant $k = Df(a)h + \|h\|_E \varepsilon_1(h)$ dans (2.1) :

$$\begin{aligned} g \circ f(a+h) =& g\big[b + Df(a)h + \|h\|_E \varepsilon_1(h)\big] \\ =& g \circ f(a) + Dg(b)\big[Df(a)h\big] + \|h\|_E \varepsilon_3(h) \end{aligned}$$

avec

$$\varepsilon_3(h) = Dg(b)\varepsilon_1(h) + \big[Df(a)h/\|h\|_E + \varepsilon_1(h)\big]\varepsilon_2\big(Df(a)h + \|h\|_E \varepsilon_1(h)\big),$$

si $h \neq 0_E$ et $\varepsilon_3(h) = 0_G$ sinon, et on vérifie facilement que $\lim_{h \to 0_E} \varepsilon_3(h) = 0_G$. La continuité de la différentielle résulte de résultats classiques sur la continuité des fonctions composées. □

En utilisant la différentiabilité des applications bilinéaires avec la proposition ci-dessus, on montre que :

1. Soit $f_1 : E \to \mathbb{R}$ et $f_2 : E \to \mathbb{R}$ deux applications différentiables en un point $a \in E$. Alors $f = f_1 f_2$ est différentiable en a et :

$$Df(a)v = f_1(a)Df_2(a)v + f_2(a)Df_1(a)v \qquad \text{pour tout } v \in E.$$

2. Soit $f_1 : E \to F$ et $f_2 : E \to \mathbb{R}$ deux applications différentiables en $a \in E$. Alors $f_1 f_2$ est différentiable en a et :

$$Df(a)v = f_1(a)(Df_2(a)v) + (Df_1(a)v)f_2(a) \qquad \text{pour tout } v \in E.$$

3. Si F est muni d'un produit scalaire $\langle\cdot,\cdot\rangle$ et si $f_1 : E \to F$ et $f_2 : E \to F$ sont deux fonctions différentiables en $a \in E$, alors $f = \langle f_1, f_2\rangle$ est différentiable en a et :

$$Df(a)v = \langle Df_1(a)v, f_2(a)\rangle + \langle Df_2(a)v, f_1(a)\rangle \qquad \text{pour tout } v \in E.$$

On a montré en TD que :

Proposition 2.5. *On note $\mathscr{L}(E) = \mathscr{L}(E,E)$ et $\mathscr{U}$ est l'ouvert de $\mathscr{L}(E)$ constitué des applications inversibles. Alors l'application*

$$\begin{array}{rcl} T : \mathscr{U} & \to & \mathscr{L}(E) \\ \ell & \mapsto & \ell^{-1}, \end{array}$$

est différentiable en tout point ℓ_0 de $\mathscr{U}$ et :

$$DT(\ell_0)\ell = -\ell_0^{-1} \circ \ell \circ \ell_0^{-1} \qquad \textit{pour tout } \ell \in \mathscr{L}(E).$$

En combinant ce résultat avec la Proposition 2.4 ou en considérant tout simplement la fonction $x \longmapsto 1/x$ de $\mathbb{R} \setminus \{0\}$ dans $\mathbb{R}$, on prouve :

Proposition 2.6 (Inverse)**.** *Soit $f : E \to \mathbb{R}$ une fonction différentiable en un point a et telle que $f(a) \neq 0$ (respectivement de classe $\mathscr{C}^1$ sur un ouvert $\mathscr{U}$ de E sur lequel elle ne s'annule pas). Alors la fonction $g = 1/f$ est définie sur un voisinage de a (respectivement sur $\mathscr{U}$), elle est différentiable en a (respectivement de classe $\mathscr{C}^1$ sur $\mathscr{U}$) et :*

$$Dg(a) = -\frac{1}{f(a)^2} Df(a).$$

2.3 Dérivées directionnelles, dérivées partielles

Commençons par la notion de dérivée directionnelle : Soit $f : \mathscr{U} \to F$ où $\mathscr{U}$ est un ouvert de E. Soit $a \in \mathscr{U}$ et $h \in E$, $h \neq 0_E$. Notons

$$\mathscr{U}(a,h) = \{t \in \mathbb{R} \,:\, a + th \in \mathscr{U}\},$$

qui est un ouvert de $\mathbb{R}$ contenant 0. Enfin définissons :

$$\begin{array}{rcl} g : \mathscr{U}(a,h) & \to & F \\ t & \mapsto & f(a+th). \end{array} \tag{2.2}$$

Définition 2.4 (Dérivée directionnelle)**.** *Si g est dérivable en 0, alors le vecteur $g'(0)$ de F s'appelle la dérivée directionnelle de f en a dans la direction h.*

Proposition 2.7. *Si f est différentiable en a alors f admet des dérivées directionnelles dans toutes les directions. De plus, pour tout $h \in E$, on a :*

$$g'(0) = Df(a)h,$$

où g est définie par (2.2) et $g'(0)$ dans le point 3 de la remarque 2.1.

Démonstration. D'après la définition de la différentielle de f en a, on a pour tout $h \in E$:

$$f(a+h) = f(a) + Df(a)h + \|h\|_E \varepsilon(h),$$

où ε est définie sur un voisinage de 0_E, est à valeurs dans F et vérifie $\varepsilon(h) \to 0_F$ lorsque $h \to 0_E$. Ceci conduit, en remplaçant h par th avec $t \in \mathscr{U}(a,h)$ à l'égalité :

$$g(t) = g(0) + tDf(a)h + |t|\bar{\varepsilon}(t),$$

avec $\bar{\varepsilon}(t) = \|h\|_E \varepsilon(th)$. Ceci prouve que g est dérivable en 0 et que $g'(0) = Df(a)h$. □

La réciproque de cette proposition est fausse. Une fonction peut admettre des dérivées directionnelles dans toutes les directions en un point et pas même être continue en ce point. Considérons en effet la fonction suivante :

$$f(x,y) = \frac{x^6}{(x^2-y)^2 + x^8} \quad \text{si } (x,y) \neq (0,0) \tag{2.3a}$$

$$f(x,y) = 0 \quad \text{si } (x,y) = (0,0). \tag{2.3b}$$

On peut montrer que cette fonction admet des dérivées directionnelles dans toutes les directions au point (0,0) mais qu'elle n'est pas continue en ce point (cf. figure 2.1).

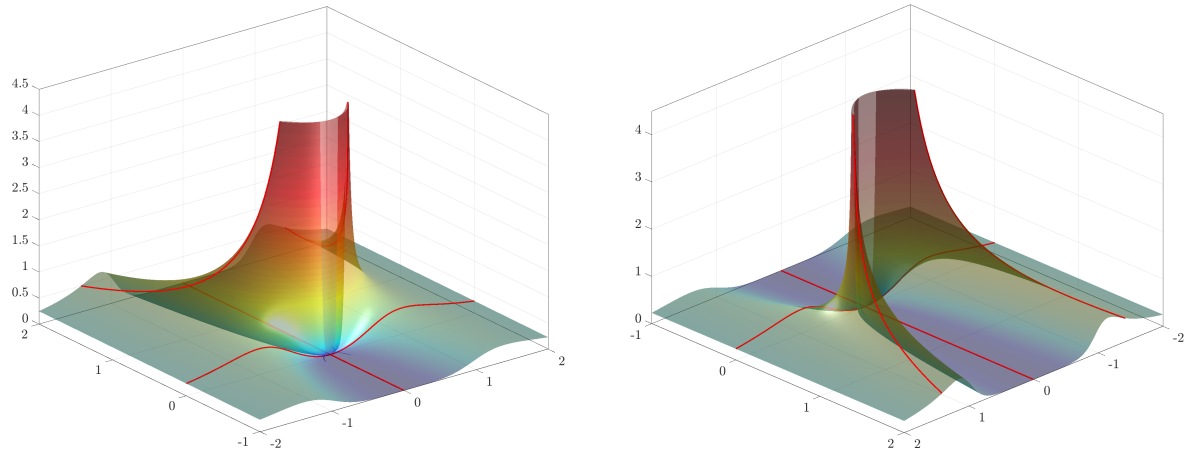


FIG. 2.1 – *Graphe de la fonction f définie par* (2.3) *sous deux angles différents (la fonction est tronquée à une certaine hauteur car elle prend des valeurs infinies au voisinage de* (0,0)*). Les images des droites $x = 0$, $y = 0$ et de la parabole $y = x^2$ sont tracées en rouge. Les limites de la fonction à l'origine suivant les droites sont clairement 0. En revanche la limite est $+\infty$ le long de la parabole.*

On considère toujours une fonction $f : \mathscr{U} \to F$ où $\mathscr{U}$ est un ouvert de E et on introduit une base $\{\beta_1, \ldots, \beta_p\}$ de F. L'application f peut alors s'écrire dans cette base :

$$\text{Pour tout } x \in \mathscr{U}, \qquad f(x) = f_1(x)\beta_1 + \ldots + f_p(x)\beta_p,$$

où pour tout $1 \leqslant j \leqslant p$:

$$\begin{array}{rcl} f_j : \mathscr{U} & \to & \mathbb{R} \\ x & \mapsto & f_j(x). \end{array}$$

Proposition 2.8. *La fonction f est différentiable en $a \in \mathscr{U}$ si et seulement si chaque function f_j $(1 \leqslant j \leqslant p)$ est différentiable en a. Dans ce cas :*

$$Df(a)h = \sum_{j=1}^{p} (Df_j(a)h)\beta_j.$$

Démonstration. Supposons que chaque foncion f_j est différentiable en a. Alors il existe des fonctions ε_j définies sur un voisinage de 0_E et à valeurs dans $\mathbb{R}$ telles que $|\varepsilon_j(h)| \to 0$ quand $h \to 0_E$ et pour tout h dans un voisinage de 0_E :

$$f_j(a+h) = f_j(a) + Df_j(a)h + \|h\|_E \varepsilon_j(h).$$

Ceci entraîne que :

$$\sum_{j=1}^{p} f_j(a+h)\beta_j = \sum_{j=1}^{p} f_j(a)\beta_j + \sum_{j=1}^{p} (Df_j(a)h)\beta_j + \|h\|_E \sum_{j=1}^{p} \varepsilon_j(h)\beta_j,$$

ce qui est la définition de la différentiabilité de f en a. Réciproquement, si f est différentiable en a, alors pour tout h dans un voisinage de 0_E :

$$f(a+h) = f(a) + Df(a)h + \|h\|_E \varepsilon(h),$$

où ε est défini sur un voisinage de 0_E, est à valeurs dans F et vérifie $\varepsilon(h) \to 0_F$ quand $h \to 0_E$. Il suffit ensuite de décomposer cette égalité sur la base des β_j. □

Selon cette proposition, l'étude de la différentiabilité de f (qui est à valeurs dans F) se réduit à l'étude de la différentiabilité de ses composantes f_j (qui sont à valeurs dans $\mathbb{R}$).

Introduisons un base $\{\alpha_1, \ldots, \alpha_n\}$ de E. On peut maintenant écrire, pour tout x dans $\mathscr{U}$ que :

$$f(x) = f(x_1\alpha_1 + \ldots + x_n\alpha_n) = f_1(x)\beta_1 + \ldots + f_p(x)\beta_p.$$

Définition 2.5 (Dérivées partielles)**.** *Les dérivées des fonctions f_j $(1 \leqslant j \leqslant p)$ dans les directions α_k $(1 \leqslant k \leqslant n)$ sont appelées les dérivées partielles de f. On note :*

$$Df_j(a)(\alpha_k) = \frac{\partial f_j}{\partial x_k}(a).$$

En combinant les propositions 2.7, 2.8 et la définition 2.5, on obtient :

Proposition 2.9. *Si f est différentiable en a alors toutes ses dérivées partielles en a existent et pour tout $h = h_1\alpha_1 + \ldots + h_n\alpha_n \in E$:*

$$\begin{aligned} Df(a)h &= \sum_{j=1}^{p} Df_j(a)(h_1\alpha_1 + \ldots + h_n\alpha_n)\beta_j \\ &= \sum_{j=1}^{p} \left(\sum_{k=1}^{n} Df_j(a)(\alpha_k)h_k \right) \beta_j = \sum_{j=1}^{p} \left(\sum_{k=1}^{n} \frac{\partial f_j}{\partial x_k}(a)h_k \right) \beta_j. \end{aligned}$$

Définition 2.6. *Dans les bases* $\{\alpha_1, \ldots, \alpha_n\}$ *et* $\{\beta_1, \ldots, \beta_p\}$*, la matrice qui représente l'application linéaire* $Df(a)$ *est :*

$$M(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \ldots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & & & \vdots \\ \frac{\partial f_p}{\partial x_1}(a) & \frac{\partial f_p}{\partial x_2}(a) & \ldots & \frac{\partial f_p}{\partial x_n}(a) \end{pmatrix}.$$

Cette matrice s'appelle la matrice Jacobienne de f*. Si* $h = h_1\alpha_1 + \ldots + h_n\alpha_n$ *et si on note* $Df(a)h = y_1\beta_1 + \ldots + y_p\beta_p$ *alors :*

$$\begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix} = M(a) \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}.$$

Comme on le fait pour les applications linéaires, on identifie parfois $Df(a)$ *avec sa matrice Jacobienne et l'on note abusivement (mais c'est très pratique)* $Df(a) = M(a)$*. Notons deux cas particuliers :*

1. *Lorsque* $E = \mathbb{R}^n$ *et* $F = \mathbb{R}$*, la matrice Jacobienne est une matrice ligne que l'on identifie cette fois avec un vecteur de* $\mathbb{R}^n$*. On note ce vecteur* $\nabla f(a)$ *(gradient de* f *en* a*). On a donc :*

$$\nabla f(a) = \left(\frac{\partial f_1}{\partial x_1}(a), \frac{\partial f_1}{\partial x_2}(a), \ldots, \frac{\partial f_1}{\partial x_n}(a) \right),$$

   et pour tout $h \in \mathbb{R}^n$ *:*

$$Df(a)h = \nabla f(a) \cdot h.$$

2. *Lorsque* $E = \mathbb{R}$ *et* $F = \mathbb{R}^n$*, la matrice jacobienne est une matrice colonne. On note alors :*

$$Df(a) = \begin{pmatrix} f_1'(a) \\ f_2'(a) \\ \vdots \\ f_n'(a) \end{pmatrix} = f'(a),$$

   et pour tout $h \in \mathbb{R}$ *:*

$$Df(a)h = hf'(a) = h \begin{pmatrix} f_1'(a) \\ f_2'(a) \\ \vdots \\ f_n'(a) \end{pmatrix}.$$

Les propriétés suivantes sont évidentes :

Proposition 2.10. *Si* $\dim E = \dim F$ *alors la matrice Jacobienne de* $Df(a)$ *est une matrice carrée et* $Df(a)$ *est un isomorphisme si est seulement si le déterminant de sa matrice Jacobienne est non nul.*

Si $f : E \to F$ *est différentiable en* a *et* $g : F \to G$ *est différentiable en* $f(a)$ *alors on a montré (proposition 2.4) que* $g \circ f$ *était différentiable en* a *et que :*

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

La matrice Jacobienne de $g \circ f$ *en* a *est obtenue en faisant le produit matricielle de la matrice Jacobienne de* g *en* $f(a)$ *avec la matrice Jacobienne de* f *en* a*.*

Comme déjà observé plus haut, la réciproque est fausse (l'existence de toutes les dérivées partielles n'entraîne pas la différentiabilité de la fonction).

En revanche, on peut énoncer :

Théorème 2.1. *Soit $\mathscr{U}$ un ouvert d'un espace vectoriel normé $(E,\|\cdot\|_E)$ et $f : \mathscr{U} \to F$ où $(F,\|\cdot\|_F)$ est un autre ev normé. On suppose que les dérivées partielles de f existent en tout point de $\mathscr{U}$ et qu'elles sont continues en un point $a \in \mathscr{U}$. Alors f est différentiable au point a.*

Démonstration. D'après la proposition 2.8, on peut se restreindre au cas où $F = \mathbb{R}$ et via l'introduction d'une base $\{\alpha_1, \ldots, \alpha_n\}$ de E, on peut identifier E avec $\mathbb{R}^n$. Soit $\mathscr{U} = \mathscr{U} - \{a\}$ et $h \in \mathscr{U}$. Notons, pour tout $x = (x_1, \ldots, x_n)$ dans E :

$$\ell(x) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_k}(a) x_k.$$

On cherche à estimer la quantité :

$$\begin{aligned} r(h) &= f(a+h) - f(a) - \ell(h) \\ &= f(a_1 + h_1, \ldots, a_n + h_n) - f(a_1, \ldots, a_n) - \sum_{j=1}^{n} \frac{\partial f}{\partial x_k}(a) h_k. \end{aligned}$$

Notons, pour tout $k = 1, \ldots, n$:

$$\begin{aligned} d_k(h) &= f(a_1, \ldots, a_{k-1}, a_k + h_k, a_{k+1} + h_{k+1}, \ldots, a_n + h_n) - \\ &\qquad\qquad f(a_1, \ldots, a_{k-1}, a_k, a_{k+1} + h_{k+1}, \ldots, a_n + h_n) \\ &= h_k \int_0^1 \frac{\partial f}{\partial x_k}(a_1, \ldots, a_{k-1}, a_k + t h_k, a_{k+1} + h_{k+1}, \ldots, a_n + h_n)\, \mathrm{d}t. \end{aligned}$$

On note ensuite que :

$$\begin{aligned} r(h) &= \sum_{k=1}^{n} d_k(h) - \frac{\partial f}{\partial x_k}(a) h_k \\ &= \sum_{k=1}^{n} h_k \int_0^1 \left[\frac{\partial f}{\partial x_k}(a_1, \ldots, a_{k-1}, a_k + t h_k, a_{k+1} + h_{k+1}, \ldots, a_n + h_n) - \frac{\partial f}{\partial x_k}(a) \right] \mathrm{d}t \end{aligned}$$

Par continuité des dérivées partielles en a, pour tout $k = 1, \ldots, n$ et tout $\varepsilon > 0$ il existe $\eta_k > 0$ tel que si $\|h\|_1 \leqslant \eta_k$ alors :

$$\left| \frac{\partial f}{\partial x_k}(a+h) - \frac{\partial f}{\partial x_k}(a) \right| < \varepsilon.$$

Notons $\eta = \min\{\eta_k, k = 1, \ldots, n\}$. Alors pour $\|h\|_1 < \eta$, on obtient que :

$$r(h) < \varepsilon \sum_{k=1}^{n} |h_k| = \varepsilon \|h\|_1.$$

On a montré que pour tout $h \in \mathscr{U}$:

$$f(a+h) = f(a) + \ell(h) + r(h),$$

avec $r(h) = o(\|h\|_1)$. Ceci prouve que f est différentiable en a et que sa différentielle est ℓ. $\square$

Corollaire 2.1. *Si les dérivées partielles de f existent et sont continues sur $\mathscr{U}$ alors f est de classe $\mathscr{C}^1$ sur $\mathscr{U}$.*

Exemple 2.1. La fonction

$$\begin{array}{rcl} f : \mathbb{R}^3 & \longrightarrow & \mathbb{R}^2 \\ (x_1,x_2,x_3) & \longmapsto & \begin{pmatrix} \sin(x_1^2 e^{x_2}) x_3^3 \\ \cos(x_3^3 + x_2^4) e^{3x_1} \end{pmatrix} \end{array}$$

est de classe $\mathscr{C}^1$ sur $\mathbb{R}^3$.

2.4 Représentations graphiques

On considère le cas d'une fonction $f : \mathscr{U} \to \mathbb{R}$ où $\mathscr{U}$ est un ouvert de $\mathbb{R}^2$. Le graphe de f est une surface S_f de $\mathbb{R}^3$ d'équation $z = f(x,y)$ ou plus précisément :

$$S_f = \big\{(x,y,z) \in \mathbb{R}^3 \,:\, z = f(x,y),\ (x,y) \in \mathscr{U}\big\}.$$

Si f est différentiable en un point $(a,b) \in \mathbb{R}^2$, alors, au voisinage de ce point :

$$f(x,y) = f(a,b) + Df(a,b)((x-a),(y-b)) + \|(x-a,y-b)\|\varepsilon((x-a,y-b)),$$

la fonction ε étant définie sur un voisinage de (0,0) et tendant vers 0 en (0,0). La différentielle de f s'écrit alors :

$$Df(a,b)((x-a),(y-b)) = \nabla f(a,b) \cdot ((x-a),(y-b)) = \frac{\partial f}{\partial x}(a,b)(x-a) + \frac{\partial f}{\partial y}(a,b)(y-b).$$

Le plan tangent à la surface S_f au point $(a,b,f(a,b))$ est donc le plan P d'équation :

$$P \,:\, z = f(a,b) + \frac{\partial f}{\partial x}(a,b)(x-a) + \frac{\partial f}{\partial y}(a,b)(y-b).$$

Exemple 2.2. La fonction $f : (x,y) \mapsto 2 - x^2 - y^2$ est différentiable sur $\mathbb{R}^2$. Un simple calcul conduit à l'expression $\nabla f(x,y) = (-2x, -2y)$ et l'équation de l'hyperplan P tangent au graphe de f au point (1,2) est :

$$P \,:\, z = f(1,2) + \nabla f(1,2) \cdot (x-1, y-2) = -3 - 2(x-1) - 4(y-2) = 7 - 2x - 4y.$$

La surface S_f et le plan P sont représentés sur la figure 2.2.

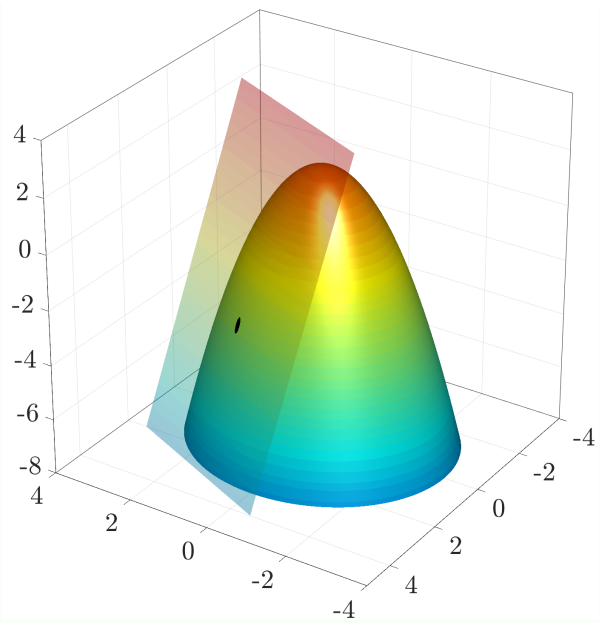


FIG. 2.2 – *La surface S_f et le plan P tangent au point* (1,2).

On peut aussi chercher à représenter, comme sur la figure 2.3, les lignes de niveau de f, c'est à dire, pour tout $\lambda \in \mathbb{R}$, les ensembles

$$\mathscr{C}_\lambda = \{(x,y) \in \mathbb{R}^2 \,:\, f(x,y) = \lambda\},$$

ainsi que le champ de vecteurs du gradient, c'est à dire la fonction

$$\begin{array}{rcl} \mathscr{U} & \longrightarrow & \mathbb{R}^2 \\ (x,y) & \longmapsto & \nabla f(x,y). \end{array}$$

Les lignes de niveau de la fonction définie en (2.3) (fonction admettant des dérivées directionnelles dans toutes les directions en (0,0) mais qui n'est pas continue en ce point) sont tracées sur la figure 2.4.

Soit (a,b) un point de $\mathscr{U}$ et $\gamma : [-1,1] \to \mathscr{U}$ une paramétrisation de classe $\mathscr{C}^1$ d'une courbe Γ telle que $\gamma(0) = (a,b)$ et $\|\gamma'(0)\| = 1$. La fonction :

$$\begin{array}{rcl} H : [-1,1] & \longrightarrow & \mathscr{U} \\ t & \longmapsto & f \circ \gamma(t), \end{array}$$

est alors de classe $\mathscr{C}^1$ et au voisinage de $t = 0$ elle admet le développement limité à l'ordre 1 :

$$H(t) = H(0) + tH'(0) + |t|\eta(t), \tag{2.4a}$$

où η est définie au voisinage de $t = 0$ et tend vers 0 quand t tend vers 0. La formule de dérivation d'une fonction composée conduit à l'expression :

$$H'(0) = \nabla f(a,b) \cdot \gamma'(0) = \|\nabla f(a,b)\| \cos(\theta), \tag{2.4b}$$

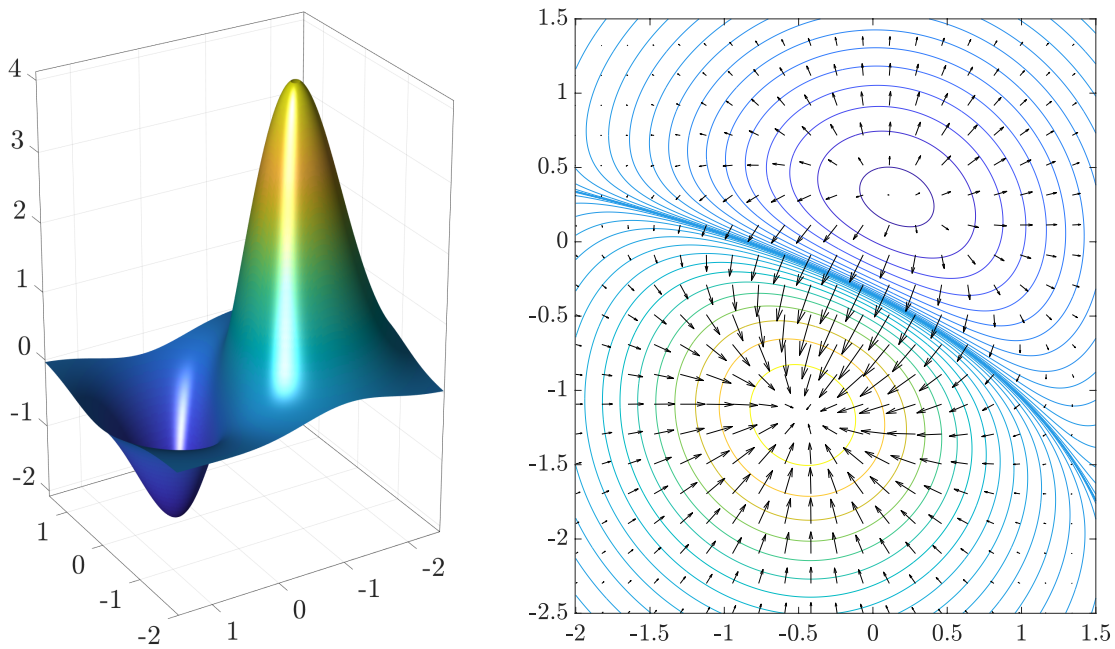


FIG. 2.3 – *Graphe d'une fonction (à gauche) et de ses lignes de niveau et son champ de gradients (à droite). Les vecteurs gradients sont orthogonaux aux lignes de niveau.*

avec θ l'angle entre les deux vecteurs $\nabla f(a,b)$ et $\gamma'(0)$ (cf. Fig. 2.5). Examinons deux cas particuliers :

1. Supposons qu'au voisinage de $t = 0$, γ soit une paramétrisation de la ligne de niveau $\mathscr{C}_{f(a,b)}$. Dans ce cas, la fonction H est localement constante au voisinage de $t = 0$ et donc $H'(0) = 0$. Ceci signifie que les vecteurs $\gamma'(0)$ et $\nabla f(a,b)$ sont orthogonaux. Mais le vecteur $\gamma'(0)$ est tangent en (a,b) à la courbe Γ est donc aussi à la courbe de niveau $\mathscr{C}_{f(a,b)}$. Il s'en suit que :

   > Le vecteur gradient $\nabla f(a,b)$ est orthogonal à la courbe de niveau qui passe par le point (a,b).

   Ce résultat est illustré sur la figure 2.3.
2. Reprenons les deux égalités (2.4) pour obtenir qu'au voisinage de $t = 0$:

$$\frac{H(t) - H(0)}{t} = \|\nabla f(a,b)\| \cos(\theta) + \operatorname{sign}(t)\eta(t).$$

   Cherchant à maximiser le taux d'accroissement, il faut alors choisir $\theta = 0$ c'est à dire que les vecteurs $\nabla f(a,b)$ et $\gamma'(0)$ sont colinéaires et orientés dans le même sens. On en déduit :

   > Le vecteur gradient indique la direction dans laquelle la fonction croît le plus rapidement.

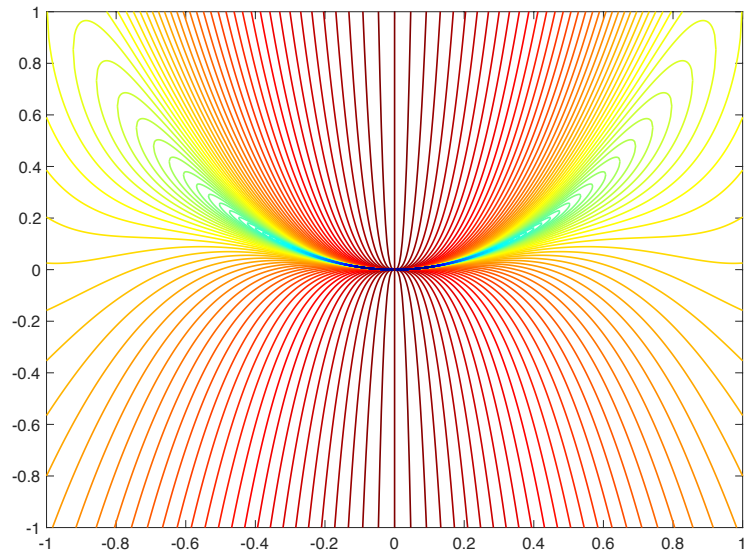


FIG. 2.4 – *Quelques lignes de niveau de la fonction f définie en* (2.4).

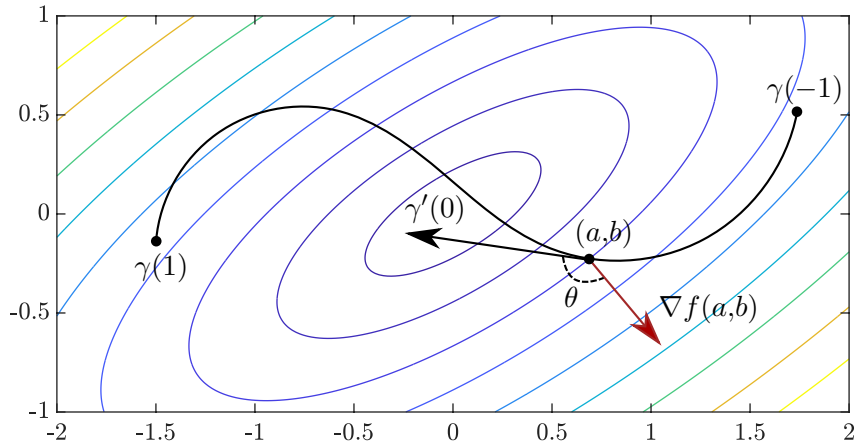


FIG. 2.5 – *Le vecteur gradient en un point est orthogonal à la ligne de niveau de la fonction qui passe par ce point. Le vecteur gradient indique la direction (et le sens) dans laquelle la fonction croît le plus rapidement.*

Chapitre 3

Théorème de la moyenne

3.1 Rappels sur les fonctions de $\mathbb{R}$ dans $\mathbb{R}$

Théorème 3.1 (Égalité des accroissements finis)**.** *Soit $f : [a,b] \to \mathbb{R}$ (où $a < b$) une fonction continue sur $[a,b]$ et dérivable sur $]a,b[$, alors il $c \in]a,b[$ tel que :*

$$f(b) - f(a) = f'(c)(b-a).$$

On en déduit:

Corollaire 3.1 (Inegalité des accroissements finis)**.** *Sous les mêmes hypothèses que le théorème ci-dessus. Si de plus il existe une contante $k \geqslant 0$ telle que, pour tout $x \in]a,b[$:*

$$|f'(x)| \leqslant k,$$

alors, pour tout couple $(x,y) \in [a,b] \times [a,b]$:

$$|f(x) - f(y)| \leqslant k|x-y|.$$

3.2 Fonctions de E dans F

Commençons par énoncer des résultats qui se déduisent directement de la section précédente.

Définition 3.1. *Soient a et b deux points d'un ev E. On définit le segment $[a,b]$ (respectivement $]a,b[$) par :*

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

On a montré en TD que :

Théorème 3.2. *Soient $(E,\|\cdot\|_E)$ et $(F,\|\cdot\|_F)$ deux ev normés munis de normes euclidiennes. Soit $f : \mathscr{U} \to F$ où $\mathscr{U}$ est un ouvert de E contenant un segment $[a,b]$. On suppose que f est différentiable sur $\mathscr{U}$. Alors il existe un point $c \in]a,b[$ tel que :*

$$\|f(b) - f(a)\|_F \leqslant \|Df(c)\|_{\mathscr{L}(E,F)} \|b-a\|_E.$$

Nous allons maintenant établir une version un peu plus faible de ce résultat mais valable pour toute norme.

Théorème 3.3 (de la moyenne)**.** *Soit* $[a,b]$ *un intervalle de* $\mathbb{R}$, $(F,\|\cdot\|_F)$ *un ev normé et* $f : [a,b] \to F$, $g : [a,b] \to \mathbb{R}$ *deux applications dérivables sur* $[a,b]$. *On suppose que pour tout* $t \in [a,b]$ *:*

$$\|f'(t)\|_F \leqslant g'(t). \tag{3.1}$$

On a alors :

$$\|f(b) - f(a)\|_F \leqslant g(b) - g(a).$$

Démonstration. Soit $\varepsilon > 0$. Posons I_ε l'ensemble des réels s dans $[a,b]$ tels que, pour tout $t \in [a,s]$:

$$\|f(t) - f(a)\|_F \leqslant g(t) - g(a) + \varepsilon(t - a).$$

Remarquons que

1. $a \in I_\varepsilon$;
2. si $\theta \in I_\varepsilon$ alors $[a,\theta] \subset I_\varepsilon$;
3. I_ε est fermé. En effet, notons φ_ε la fonction définie pour tout $s \in [a,b]$ par

$$\varphi_\varepsilon(s) = \|f(s) - f(a)\|_F - \big(g(s) - g(a) + \varepsilon(s - a)\big),$$

et considérons $(s_n)_n$ une suite de I_ε qui converge vers s^* dans $\mathbb{R}$. Alors, par continuité de l'application φ_ε :

$$\lim_{n\to+\infty} \varphi_\varepsilon(s_n) = \varphi(s^*),$$

et donc $\varphi_\varepsilon(s^*) \leqslant 0$. D'autre part, pour tout $t \in [a,s^*[$, il existe un rang n_0 tel que $t < s_{n_0}$. Comme $s_{n_0} \in I_\varepsilon$, alors $[a,s_{n_0}] \in I_\varepsilon$, donc $\varphi_\varepsilon(t) \leqslant 0$ et $t \in I_\varepsilon$.

Notons $\theta^* = \sup I_\varepsilon$. Alors, de 1, 2 et 3, on déduit que $I_\varepsilon = [a,\theta^*]$. Supposons que $\theta^* < b$. Comme $\theta^* \in I_\varepsilon$, on a :

$$\|f(\theta^*) - f(a)\|_F \leqslant g(\theta^*) - g(a) + \varepsilon(s - a).$$

Comme f et g sont dérivables en θ^*, il existe $\delta > 0$ (que l'on peut choisir tel que $[\theta^*,\theta^* + \delta] \subset [a,b]$) tel que, pour tout $s \in [0,\delta]$:

$$\|f(\theta^* + s) - f(\theta^*) - sf'(\theta^*)\|_F < \frac{\varepsilon}{2}s,$$
$$|g(\theta^* + s) - g(\theta^*) - sg'(\theta^*)| < \frac{\varepsilon}{2}s.$$

On en déduit que pour tout $s \in [0,\delta]$:

$$\|f(\theta^* + s) - f(\theta^*)\|_F \leqslant s\|f'(\theta^*)\|_F + \frac{\varepsilon}{2}s,$$
$$sg'(\theta^*) \leqslant g(\theta^* + s) - g(\theta^*) + \frac{\varepsilon}{2}s.$$

En utilisant (3.1), il vient dans un premier temps :

$$\|f(\theta^* + s) - f(\theta^*)\|_F \leqslant g(\theta^* + s) - g(\theta^*) + \varepsilon s,$$

et ensuite :

$$\begin{aligned}\|f(\theta^* + s) - f(a)\|_F &\leqslant \|f(\theta^* + s) - f(\theta^*)\|_F + \|f(\theta^*) - f(a)\|_F \\ &\leqslant g(\theta^* + s) - g(a) + \varepsilon(s + \theta^* - a).\end{aligned}$$

Ceci entraîne $[\theta^*,\delta[\subset I_\varepsilon$ et est en contradiction avec la définition de θ^*. On a donc $\theta^* = b$ et on a montré que pour tout $\varepsilon > 0$, on avait :

$$\|f(b) - f(a)\|_F \leqslant g(b) - g(a) + \varepsilon(b - a).$$

En faisant tendre ε vers 0, on conclut la démonstration du théorème. □

Définition 3.2. *On dit qu'un ensemble $\mathscr{C}$ de E est convexe si pour tout couple de points $(a,b) \in \mathscr{C} \times \mathscr{C}$, le segment $[a,b]$ est inclus dans $\mathscr{C}$.*

On déduit immédiatement du théorème précédent le

Corollaire 3.2. *Soit $\mathscr{C}$ un ouvert convexe de E et $f : \mathscr{C} \to F$ une fonction différentiable. On suppose qu'il existe une constant $k \geqslant 0$ telle que pour tout $c \in \mathscr{C}$:*

$$\|Df(c)\|_{\mathscr{L}(E,F)} \leqslant k.$$

Alors pour tout couple $(a,b) \in \mathscr{C} \times \mathscr{C}$:

$$\|f(b) - f(a)\|_F \leqslant k\|b - a\|_E.$$

(La fonction est donc Lipschitzienne sur $\mathscr{C}$).

Démonstration. Par convexité, le segment $[a,b]$ est inclus dans $\mathscr{C}$. On considère la fonction $F : t \in [0,1] \mapsto f(a + t(b - a))$. Cette fonction est différentiable (dérivable) et

$$F'(t) = Df(a + t(b - a))(b - a) \qquad \text{pour tout } t \in [0,1].$$

On a alors, pour tout $t \in [0,1]$:

$$\|F'(t)\|_F \leqslant \|Df(a + t(b - a))\|_{\mathscr{L}(E,F)}\|b - a\|_E \leqslant k\|b - a\|_E.$$

On applique ensuite le théorème 3.3 avec la fonction $g : t \in [0,1] \mapsto tk\|b - a\|_E$. □

Une autre conséquence est :

Corollaire 3.3. *Soit $\mathscr{U}$ un ouvert connexe de E et $f : \mathscr{U} \to F$ une fonction différentiable telle que $Df(x) = 0$ pour tout $x \in \mathscr{U}$. Alors f est constante sur $\mathscr{U}$.*

Démonstration. Soit a un point de $\mathscr{U}$. On pose $\mathscr{E} = \{x \in \mathscr{U} \ : \ f(x) = f(a)\}$. Alors $\mathscr{E}$ est non vide (il contient a) et fermé. Soit $x \in \mathscr{E}$. Alors il existe une boule ouverte centrée en x et incluse dans $\mathscr{U}$. On applique le Corollaire 3.2 sur cette boule (qui est convexe), ce qui prouve que la boule est incluse dans $\mathscr{E}$. L'ensemble $\mathscr{E}$ est donc aussi ouvert. On en déduit que $\mathscr{U} = \mathscr{E}$. □

Donnons une dernière application du théorème de la moyenne :

Proposition 3.1. *Soit $f : [a,b] \to E$ une fonction continue sur $[a,b]$. Alors :*

$$\left\| \int_a^b f(s)\,\mathrm{d}s \right\|_E \leqslant \int_a^b \|f(s)\|_E\,\mathrm{d}s.$$

Démonstration. Pour tout t dans $[a,b]$, posons :

$$F(t) = \int_a^t f(s)\,\mathrm{d}s \qquad \text{et} \qquad g(t) = \int_a^t \|f(s)\|_E\,\mathrm{d}s.$$

En introduisant une base $\{e_1,\ldots,e_n\}$ de E et en écrivant que $f(s) = \sum_{k=1}^n f_k(s)e_k$ avec $f_k : [a,b] \to \mathbb{R}$, on obtient que :

$$F(t) = \sum_{k=1}^n \left(\int_a^t f_k(s)\,\mathrm{d}s\right) e_k.$$

Les fonctions f_k étant continues (comme f), la fonction F est différentiable (dérivable) sur $[a,b]$ et :

$$F'(t) = \sum_{k=1}^n f_k(t)e_k = f(t) \qquad \text{pour tout } t \in]a,b[.$$

D'autre part, la fonction $t \mapsto \|f(t)\|_E$ est continue sur $[a,b]$ (comme composée de fonctions continues) donc g est dérivable sur $[a,b]$ et $g'(t) = \|f(t)\|_E$. On applique alors le théorème 3.3 et on obtient l'inégalité annoncée. □

Le théorème de la moyenne et les corollaires 3.2 et 3.3 sont encore valables en dimension infinie.

3.3 Remarques

La proposition 3.1 se déduit en fait directement de la définition d'une intégrale comme limite d'une somme de Riemann et de l'inégalité triangulaire. En effet, soit σ un réel strictement positif tel que $\sigma < (b-a)$ et $a = x_0^\sigma < x_1^\sigma < \ldots < x_{n_\sigma}^\sigma = b$, une subdivision de l'intervalle $[a,b]$ de pas σ c'est à dire telle que $\sigma = \inf_{k=1,\ldots,n_\sigma} |x_k^\sigma - x_{k-1}^\sigma|$. Soient $t_1^\sigma,\ldots,t_{n_\sigma}^\sigma$ des réels tels que $t_k^\sigma \in [x_{k-1}^\sigma, x_k^\sigma]$ pour tout $k = 1,\ldots,n_\sigma$. D'après l'inégalité triangulaire :

$$\left\|\sum_{k=1}^{n_\sigma}(x_k^\sigma - x_{k-1}^\sigma)f(t_k)\right\|_E \leqslant \sum_{k=1}^{n_\sigma}(x_k^\sigma - x_{k-1}^\sigma)\|f(t_k^\sigma)\|_E. \tag{3.2}$$

Par définition de l'intégrale de Riemann :

$$\int_a^b f(t)\mathrm{d}t = \lim_{\sigma\to 0}\sum_{k=1}^{n_\sigma}(x_k^\sigma - x_{k-1}^\sigma)f(t_k^\sigma) \quad \text{et} \quad \int_a^b \|f(t)\|_E\mathrm{d}t = \lim_{\sigma\to 0}\sum_{k=1}^{n_\sigma}(x_k^\sigma - x_{k-1}^\sigma)\|f(t_k^\sigma)\|_E.$$

On peut ensuite passer à la limite dans l'égalité (3.2) lorsque $\sigma \to 0$ pour obtenir que :

$$\left\|\int_a^b f(s)\,\mathrm{d}s\right\|_E \leqslant \int_a^b \|f(s)\|_E\,\mathrm{d}s.$$

Avec ce résultat, il est possible de démontrer une version un peu plus faible du Corollaire 3.2 dans laquelle on suppose que la fonction f n'est plus seulement différentiable mais de classe $\mathscr{C}^1$. Soient a et b deux points de $\mathscr{C}$. La fonction $F : t \in [0,1] \mapsto f(a + t(b-a))$ est alors elle aussi de classe $\mathscr{C}^1$ et on peut écrire que :

$$F(1) - F(0) = \int_0^1 F'(t)\,\mathrm{d}t = \int_0^1 Df(a + t(b-a))(b-a)\,\mathrm{d}t,$$

d'où l'on déduit que :

$$\begin{aligned}\|f(b)-f(a)\|_E &\leqslant \left\| \int_0^1 Df(a+t(b-a))(b-a)\,\mathrm{d}t \right\|_E \\ &\leqslant \int_0^1 \|Df(a+t(b-a))(b-a)\|_E\,\mathrm{d}t \leqslant k\|b-a\|_E.\end{aligned}$$

Chapitre 4

Théorèmes d'inversion

4.1 Théorème d'inversion locale

Soient $(E,\|\cdot\|_E)$ et $(F,\|\cdot\|_F)$ deux ev normés de même dimension. Soit $\mathscr{U}$ un ouvert de E et $\mathscr{V}$ un ouvert de F et $H : \mathscr{U} \to \mathscr{V}$ une application.

Définition 4.1. *On dit que H est un difféomorphisme de classe $\mathscr{C}^1$ de $\mathscr{U}$ sur $\mathscr{V}$ si :*

1. *H est une application de classe $\mathscr{C}^1$ sur $\mathscr{U}$;*
2. *H est une bijection de $\mathscr{U}$ sur $\mathscr{V}$;*
3. *H^{-1} est une application de classe $\mathscr{C}^1$ sur $\mathscr{V}$.*

On notera que H est aussi un homéomorphisme de $\mathscr{U}$ sur $\mathscr{V}$ (application continue dont l'inverse est continue). En appliquant la formule de la différentielle d'un produit de composition à la fonction $H \circ H^{-1}$, on trouve que pour tout y dans $\mathscr{V}$:

$$DH(H^{-1}(y))DH^{-1}(y) = \mathrm{Id},$$

et donc que :

$$DH^{-1}(y) = \big[DH(H^{-1}(y))\big]^{-1}.$$

Théorème 4.1 (d'inversion locale)**.** *Soit $H : \mathscr{U} \to \mathscr{V}$ une application de classe $\mathscr{C}^1$ et soit $a \in \mathscr{U}$. On suppose que $DH(a)$ est inversible. Alors il existe un ouvert $\widetilde{\mathscr{U}}$ dans $\mathscr{U}$ contenant a et un ouvert $\widetilde{\mathscr{V}}$ dans $\mathscr{V}$ contenant $H(a)$ tels que la restriction de H à $\widetilde{\mathscr{U}}$ soit un difféomorphisme de classe $\mathscr{C}^1$ de $\widetilde{\mathscr{U}}$ sur $\widetilde{\mathscr{V}}$.*

La démonstration repose sur le :

Lemme 4.1 (Point fixe de Banach)**.** *Soit $\mathscr{F}$ un ensemble fermé d'un espace de Banach F (i.e. un ev normé complet) et $\phi : \mathscr{F} \to \mathscr{F}$ une fonction contractante c'est à dire une fonction lipschitzienne avec une constante de lipschitz $k < 1$. Alors il existe un unique $x^* \in \mathscr{F}$ tel que $\phi(x^*) = x^*$.*

Démonstration. Soit $x_0 \in \mathscr{F}$. Posons pour tout $n \geqslant 0$, $x_{n+1} = \phi(x_n)$. On vérifie facilement par récurrence que pour tout $n \geqslant 0$:

$$\|x_{n+1} - x_n\|_F \leqslant k^n \|x_1 - x_0\|_F.$$

On a ensuite, pour tout $m > n \geqslant 0$:

$$\begin{aligned}\|x_m - x_n\|_F &\leqslant \|x_m - x_{m-1}\|_F + \|x_{m-1} - x_{m-2}\|_F + \cdots + \|x_{n+1} - x_n\|_F \\ &\leqslant k^n\big(k^{m-n-1} + \cdots + 1\big)\|x_1 - x_0\|_F \\ &\leqslant \frac{k^n}{1-k}\|x_1 - x_0\|_F.\end{aligned}$$

Il s'en suit que $(x_n)_n$ est une suite de Cauchy donc convergente puisque F est complet. D'autre part, la limite, notée x^* est dans $\mathscr{F}$ car $\mathscr{F}$ est fermé. De plus, pour tout $n \geqslant 0$:

$$\begin{aligned}\|\phi(x^*) - x^*\|_F &\leqslant \|\phi(x^*) - \phi(x_n)\|_F + \|\phi(x_n) - x_n\|_F + \|x_n - x^*\|_F \\ &\leqslant k\|x^* - x_{n+1}\|_F + \|x_{n+1} - x_n\|_F + \|x_n - x^*\|_F,\end{aligned}$$

et tous les membres de droite tendent vers 0 quand n tend vers $+\infty$. Ceci prouve que $\phi(x^*) = x^*$. L'unicité du point fixe est facile à montrer et la démonstration est laissée en exercice. □

Démonstration du théorème d'inversion locale. Pour tout y dans $\mathscr{V}$ et x dans $\mathscr{U}$, posons :

$$\phi_y(x) = \big[DH(a)\big]^{-1}(y - H(x)) + x.$$

Résoudre l'équation $H(x) = y$ est équivalent à résoudre le problème de point fixe $\phi_y(x) = x$. Nous allons montrer que la fonction ϕ_y vérifie un peu plus que les seules hypothèses du Lemma 4.1 (nous en aurons besoin plus loin dans la démonstration).

Lemme 4.2. *Pour tout $\varepsilon > 0$ il existe $\delta_\varepsilon > 0$ et $\delta'_\varepsilon > 0$ tels que pour tout y dans la boule fermée $\overline{B}_F(H(a),\delta'_\varepsilon)$, le fonction ϕ_y envoie la boule fermée $\overline{B}_E(a,\delta_\varepsilon)$ dans la boule fermée $\overline{B}_E(a,2\varepsilon\delta_\varepsilon)$ et la fonction*

$$\phi_y : \overline{B}_E(a,\delta_\varepsilon) \longrightarrow \overline{B}_E(a,2\varepsilon\delta_\varepsilon),$$

est lipschitzienne de constante de lipschitz ε.

Démonstration de lemme. Pour tout y dans $\mathscr{V}$ et x dans $\mathscr{U}$, on peut écrire que :

$$\begin{aligned}\phi_y(x) - a &= \big[DH(a)\big]^{-1}(y - H(x)) + x - a, \\ &= \underbrace{\big[DH(a)\big]^{-1}(y - H(a))}_{\text{constante}} + \varphi(x),\end{aligned}$$

où l'on a posé $\varphi(x) = \big[DH(a)\big]^{-1}(H(a) - H(x)) + x - a$. On en déduit que :

$$\|\phi_y(x) - a\|_E \leqslant \|\big[DH(a)\big]^{-1}(y - H(a))\|_E + \|\varphi(x)\|_E. \tag{4.1a}$$

Or la fonction $\varphi : \mathscr{U} \to E$ est de classe $\mathscr{C}^1$ (comme H) et $\varphi(a) = 0$ et $D\varphi(a) = 0$.

Soit $\varepsilon > 0$. Par continuité de la différentielle de φ, il existe $\delta_\varepsilon > 0$ tel que si $\|x - a\|_E < \delta_\varepsilon$ alors $\|D\varphi(x)\|_{\mathscr{L}(E)} < \varepsilon$. En appliquant le Corollaire 3.2 sur la boule fermée $\overline{B}_E(0,\delta_\varepsilon)$, on obtient que :

$$\|\varphi(x)\|_E \leqslant \varepsilon\|x - a\|_E \qquad \text{pour tout } x \in \overline{B}_E(0,\delta_\varepsilon). \tag{4.1b}$$

Nous n'avons encore rien imposé sur y. Choisissons y assez proche de $H(a)$ et plus précisément tel que :

$$\|y - H(a)\|_F \leqslant \delta'_\varepsilon \qquad \text{avec} \qquad \delta'_\varepsilon = \frac{\varepsilon\delta_\varepsilon}{\left\|\left[DH(a)\right]^{-1}\right\|_{\mathscr{L}(F,E)}}.$$

Pour un tel choix de y, les estimations (4.1) entraînent que la fonction ϕ_y envoie la la boule fermée $\overline{B}_E(0,\delta_\varepsilon)$ dans la boule fermée $\overline{B}_E(0,2\varepsilon\delta_\varepsilon)$. Comme d'autre part, $D\phi_y(x) = D\varphi(x)$, en appliquant de nouveau le Corollaire 3.2 sur $\overline{B}_E(0,\delta_\varepsilon)$, on montre que l'application ϕ_y est lispchitzienne de constante de lipschitz ε, ce qui conclut la démonstration du lemme. □

On choisit maintenant $\varepsilon = 1/2$ dans le lemme 4.2 ci-dessus et on applique le Lemme 4.1 (point fixe de Banach) pour obtenir l'existence d'un unique x dans $\overline{B}_E(0,\delta_\varepsilon)$ tel que $\phi_y(x) = x$, c'est à dire tel que $H(x) = y$ et ceci est vrai pour tout y dans $\overline{B}_F(H(a),\delta'_\varepsilon)$. L'application H^{-1}, notée K peut donc être définie sur cette boule. On peut alors poser $\mathscr{V}' = B_F(H(a),\delta'_\varepsilon)$ (la boule ouverte) et $\mathscr{U}' = H^{-1}(\mathscr{V}') \cap \mathscr{U}$ (un ouvert de E car H est continue). Comme H est une bijection continue de $\mathscr{U}'$ sur $\mathscr{V}'$, sa bijection réciproque K est également continue. On sait d'autre part que $x \in \mathscr{U}' \mapsto DH(x) \in \mathscr{L}(E,F)$ est continue, que $DH(a)$ est inversible et que l'ensemble des applications inversibles dans $\mathscr{L}(E,F)$ est un ouvert. On peut donc restreindre $\mathscr{U}'$ (et on note $\widetilde{\mathscr{U}}$ cette restriction) de telle sorte que pour tout x in $\widetilde{\mathscr{U}}$, $DH(x)$ soit inversible. On note enfin $\widetilde{\mathscr{V}} = K^{-1}(\widetilde{\mathscr{U}})$ (un ouvert de F car K est continue).

Gardons à l'esprit que pour tout $y \in \widetilde{\mathscr{V}}$, $K(y)$ est l'unique point fixe de ϕ_y dans $\widetilde{\mathscr{U}}$. En particulier, $K(H(a)) = a$. Nous allons maintenant vérifier que K est différentiable en $y = H(a)$ est que :

$$DK(H(a)) = \left[DH(a)\right]^{-1}.$$

Pour tout y dans $\widetilde{\mathscr{V}}$, on a :

$$\begin{aligned}\|K(y) - K(H(a)) - \left[DH(a)\right]^{-1}(y - H(a))\|_E &= \|K(y) - a - \left[DH(a)\right]^{-1}(y - H(a))\|_E \\ &= \|K(y) - \phi_y(a)\|_E.\end{aligned}$$

Soit $0 < \varepsilon < 1/2$. Selon le lemme 4.2, si y est dans $B_F(H(a),\delta'_\varepsilon)$ alors ϕ_y est une contraction de constante de lipschitz ε de la boule fermée $\overline{B}_E(a,\delta_\varepsilon)$ dans elle même. Définissons une suite $(x_n)_{n\geqslant 0}$ dans $\overline{B}_E(a,\delta_\varepsilon)$ en posant $x_0 = a$ et pour tout $n \geqslant 0$, $x_{n+1} = \phi_y(x_n)$. La suite $(x_n)_n$ converge vers l'unique point fixe de ϕ_y dans la boule, c'est à dire vers $K(y)$. En reprenant les calculs ci-dessus, on obtient donc que :

$$\begin{aligned}\|K(y) - K(H(a)) - \left[DH(a)\right]^{-1}(y - H(a))\|_E &= \|K(y) - \phi_y(a)\|_E \\ &= \lim_{n\to+\infty} \|\phi_y(x_n) - \phi_y(a)\|_E.\end{aligned}$$

Or, pour tout $n \geqslant 0$:

$$\|\phi_y(x_n) - \phi_y(a)\|_E \leqslant \varepsilon\|x_n - a\|_E.$$

On fait ensuite tendre n vers $+\infty$ pour obtenir dans un premier temps que :

$$\|K(y) - K(H(a)) - \left[DH(a)\right]^{-1}(y - H(a))\|_E \leqslant \varepsilon\|K(y) - K(H(a))\|_E.$$

Or

$$\begin{aligned}&\|K(y) - K(H(a))\|_E \leqslant \\ &\qquad \|K(y) - K(H(a)) - \left[DH(a)\right]^{-1}(y - H(a))\|_E + \|\left[DH(a)\right]^{-1}(y - H(a))\|_E,\end{aligned}$$

et donc finalement :

$$\begin{aligned}\|K(y) - K(H(a)) - [DH(a)]^{-1}(y - H(a))\|_E &\leqslant \frac{\varepsilon\|[DH(a)]^{-1}\|_{\mathscr{L}(F,E)}}{1-\varepsilon}\|y - H(a)\|_F \\ &\leqslant 2\varepsilon\|[DH(a)]^{-1}\|_{\mathscr{L}(F,E)}\|y - H(a)\|_F.\end{aligned}$$

Ceci prouve que K est différentiable en $y = H(a)$ et que sa différentielle est $DK(H(a)) = [DH(a)]^{-1}$, ce qui s'écrit encore, en posant $b = H(a)$:

$$DK(b) = [DH(K^{-1}(b))]^{-1}.$$

Mais ce résultat peut être généralisé à tous les points de l'ouvert $\widetilde{\mathscr{V}}$. En effet $DH(x)$ est inversible en tout point x de $\widetilde{\mathscr{U}}$ (par définition de cet ouvert). On peut donc faire jouer le rôle du point a à tout point x de $\widetilde{\mathscr{U}}$ ce qui entraîne que :

$$DK(y) = [DH(K^{-1}(y))]^{-1} \qquad \text{pour tout } y \in \widetilde{\mathscr{V}}.$$

Enfin, on sait que $y \in \widetilde{\mathscr{V}} \mapsto K^{-1}(y)$ est continue, que $x \in \widetilde{\mathscr{U}} \mapsto DH(x) \in \mathscr{L}(E,F)$ est continue et que $\ell \in \mathrm{GL}(E,F) \mapsto \ell^{-1} \in \mathscr{L}(F,E)$ est continue. Par composition, la différentielle de K est donc elle aussi continue sur $\widetilde{\mathscr{U}}$. □

Le théorème affirme que H est une difféomorphisme <u>local</u>. Même si $DH(x)$ est inversible en tout point de $\mathscr{U}$, cela ne prouve pas que H est un difféomorphisme de $\mathscr{U}$ sur $H(\mathscr{U})$ comme le montre l'exemple suivant :

Exemple 4.1. Posons $E = F = \mathbb{R}^2$ et $\mathscr{U} = \{(r,\theta) \in \mathbb{R}^2 \,:\, r > 0\}$ et

$$\phi(r,\theta) = (r\cos(\theta), r\sin(\theta)).$$

La matrice Jacobienne de ϕ est :

$$D\phi(r,\theta) = \begin{pmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{pmatrix}.$$

Alors $\det(D\phi(r,\theta)) = r \neq 0$ sur $\mathscr{U}$ (donc ϕ est un difféomorphisme local en tout point de $\mathscr{U}$) mais ϕ n'est pas injective sur $\mathscr{U}$ car $\phi(r,\theta + 2\pi) = \phi(r,\theta)$.

Nous pouvons énoncer un résultat d'inversion globale :

Théorème 4.2 (d'inversion globale)**.** *Soit $H : \mathscr{U} \to F$ une fonction de classe $\mathscr{C}^1$ telle que $DH(x)$ est inversible pour tout $x \in \mathscr{U}$. Si de plus H est injective, alors $H(\mathscr{U})$ est ouvert et H est une difféormophisme de $\mathscr{U}$ sur $H(\mathscr{U})$.*

À ce propos, on pourra aussi regarder l'exercice 9 de la feuille 1 de TD.

4.2 Théorème des fonctions implicites

On considère maintenant que $E = E_1 \times E_2$ (E_1 et E_2 des ev normés). Soit $\mathscr{U}$ un ouvert de E et une fonction $H : \mathscr{U} \to F$. Soit (a_1,a_2) un point de E (avec $a_1 \in E_1$ et $a_2 \in E_2$).

Le théorème suivant affirme que sous certaines hypothèses, l'ensemble des solutions de l'équation $H(x) = H(a)$ dans $\mathscr{U}$ est le graphe d'une certaine fonction de classe $\mathscr{C}^1$.

Théorème 4.3 (des fonctions implicites)**.** *On suppose que :*

1. *H est de classe $\mathscr{C}^1$ sur $\mathscr{U}$;*
2. *L'application de $\mathscr{L}(E_2,F)$ définie par :*

$$h_2 \in E_2 \longmapsto DH(a_1,a_2)(0,h_2),$$

est inversible (remarquer que ceci nécessite que E_2 et F aient la même dimension).

Alors il existe un ouvert $\mathscr{U}_1$ de E_1 contenant a_1, un ouvert $\mathscr{U}_2$ de E_2 contenant a_2 tels que :

1. *$\mathscr{U}_1 \times \mathscr{U}_2 \subset \mathscr{U}$;*
2. *Il existe une application $K : \mathscr{U}_1 \to \mathscr{U}_2$ de classe $\mathscr{C}^1$ telle que :*

$$\big\{(x_1,x_2) \in \mathscr{U}_1 \times \mathscr{U}_2 \,:\, H(x_1,x_2) = H(a_1,a_2)\big\} = \big\{(x_1,K(x_1)) \,:\, x_1 \in \mathscr{U}_1\big\}.$$

Exemple 4.2. La fonction

$$\begin{array}{rcl} H : \mathbb{R}^2 & \to & \mathbb{R} \\ (x_1,x_2) & \mapsto & x_1^2 + x_2^2, \end{array}$$

est de classe $\mathscr{C}^1$ sur $\mathbb{R}^2$. On calcule facilement que pour tout $(x_1,x_2) \in \mathbb{R}^2$:

$$DH(x_1,x_2) = (2x_1,2x_2).$$

Posons $(a_1,a_2) = (0,1)$. Alors, pour tout $h_2 \in \mathbb{R}$:

$$DH(a_1,a_2)(0,h_2) = 2h_2,$$

et donc l'application $h_2 \mapsto DH(a_1,a_2)(0,h_2)$ est bien inversible, son inverse étant $h_2 \mapsto h_2/2$. Les hypothèses du théorème des fonctions implicites sont vérifiées et donc : Il existe des ouverts $\mathscr{U}_1$ et $\mathscr{U}_2$ dans $\mathbb{R}$ tels que $(0,1) \in \mathscr{U}_1 \times \mathscr{U}_2$ et une application $K : \mathscr{U}_1 \to \mathscr{U}_2$ telle que :

$$\{(x_1,x_2) \in \mathbb{R}^2 \,:\, H(x_1,x_2) = H(a_1,a_2)\} = \{(x_1,K(x_1)) \,:\, x_1 \in \mathscr{U}_1\}.$$

Ici, on peut résoudre explicitement l'équation $x_1^2 + x_2^2 = 1^2$ pour obtenir $x_2^2 = \pm\sqrt{1-x_1^2}$. Comme on travaille au voisinage du point $(0,1)$, on trouve que $K(x_1) = \sqrt{1-x_1^2}$ avec $\mathscr{U}_1 =]-1,1[$ et $\mathscr{U}_2 =]0,+\infty[$ (par exemple). Que se passe-t-il au voisinage du point $(0,-1)$? et au voisinage du point $(1,0)$?

Démonstration. Définissons la fonction $G : \mathscr{U} \to E_1 \times F$ en posant :

$$G(x_1,x_2) = (x_1,H(x_1,x_2)) \qquad \text{pour tout } x \in \mathscr{U}. \tag{4.2}$$

Nous allons appliquer le théorème d'inversion locale à cette fonction, représentée sur la figure 4.1. Elle est de classe $\mathscr{C}^1$ et sa différentielle au point (a_1,a_2) est donnée pour tout $(h_1,h_2) \in E_1 \times E_2$ par :

$$DG(a_1,a_2)(h_1,h_2) = (h_1,DH(a_1,a_2)(h_1,h_2)).$$

L'application linéaire $DG(a_1,a_2) : E_1 \times E_2 \to E_1 \times F$ est inversible. En effet, pour tout $(w_1,w) \in E_1 \times F$, l'équation:

$$DG(a_1,a_2)(h_1,h_2) = (w_1,w),$$

conduit à :

$$w_1 = h_1 \qquad \text{et} \qquad DH(a_1,a_2)(h_1,h_2) = w.$$

Mais on peut réécrire :

$$DH(a_1,a_2)(h_1,h_2) = DH(a_1,a_2)(h_1,0) + DH(a_1,a_2)(0,h_2).$$

On doit donc résoudre :

$$DH(a_1,a_2)(0,h_2) = w - DH(a_1,a_2)(w_1,0),$$

et comme $h_2 \mapsto DH(a_1,a_2)(0,h_2)$ est inversible par hypothèse, la différentielle de G en (a_1,a_2) est bien inversible. Les hypothèses du théorème d'inversion locale sont vérifiées. On en déduit qu'il existe un ouvert $\widetilde{\mathscr{U}}$ de E contenant (a_1,a_2) et un ouvert $\widetilde{\mathscr{V}}$ de $E_1 \times F$ contenant $(a_1,H(a_1,a_2))$ tels que la restriction de G à $\widetilde{\mathscr{U}}$ soit un difféomorphisme de classe $\mathscr{C}^1$. Notons

$$\widetilde{K} : \widetilde{\mathscr{V}} \to \widetilde{\mathscr{U}},$$

le difféormorphisme inverse. La forme particulière (4.2) de G entraîne que $\widetilde{K}$ peut s'écrire :

$$\widetilde{K}(y_1,y_2) = (y_1,\widetilde{K}_2(y_1,y_2)) \qquad \text{pour tout } (y_1,y_2) \in \widetilde{\mathscr{V}}.$$

Quite à diminuer $\widetilde{\mathscr{U}}$ (et donc aussi $\widetilde{\mathscr{V}}$ en posant $\widetilde{\mathscr{V}} = \widetilde{K}^{-1}(\widetilde{\mathscr{U}})$ qui est bien un ouvert) on peut

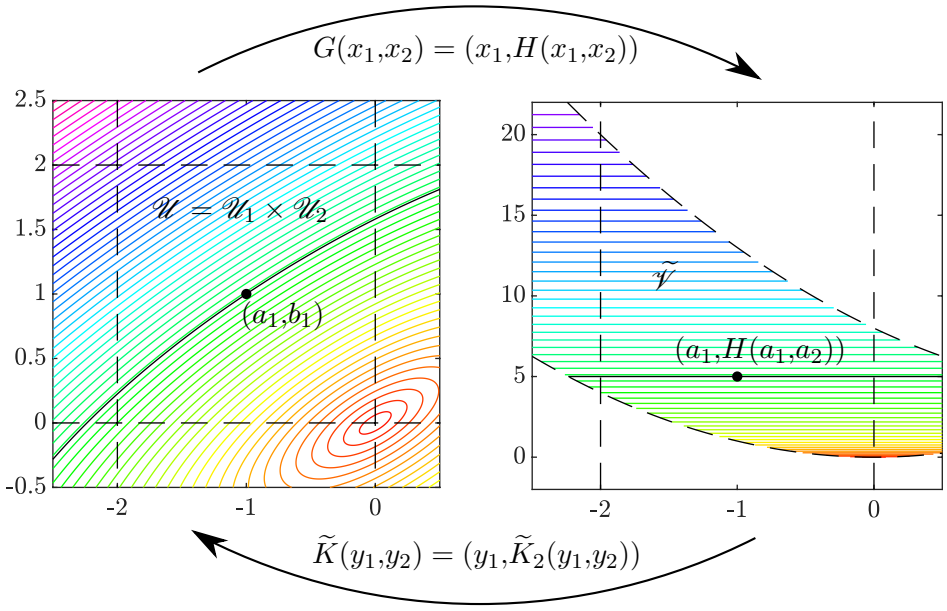


FIG. 4.1 – *La fonction G envoie les lignes de niveau de la fonction H sur des droites.*

supposer que cet ouvert est de la forme $\widetilde{\mathscr{U}_1} \times \mathscr{U}_2$ où $\widetilde{\mathscr{U}_1}$ est un ouvert de E_1 contenant a_1 et $\mathscr{U}_2$ est un ouvert de E_2 contenant a_2. Enfin posons:

$$\mathscr{U}_1 = \left\{x_1 \in \widetilde{\mathscr{U}_1} \,:\, (x_1,H(a_1,a_2)) \in \widetilde{\mathscr{V}}\right\}.$$

C'est un ouvert de E_1 contenant a_1 (c'est l'image réciproque d'un ouvert par l'application continue $x_1 \in \widetilde{\mathscr{U}_1} \mapsto (x_1,H(a_1,a_2))$) et tel que :

$$\mathscr{U}_1 \times \mathscr{U}_2 \subset \widetilde{\mathscr{U}_1} \times \mathscr{U}_2 = \widetilde{\mathscr{U}} \subset \mathscr{U}.$$

Pour (x_1,x_2) dans $\mathscr{U}_1 \times \mathscr{U}_2$, on a alors :

$$\begin{aligned} H(x_1,x_2) = H(a_1,a_2) &\Leftrightarrow G(x_1,x_2) = (x_1,H(a_1,a_2)) \\ &\Leftrightarrow (x_1,x_2) = \widetilde{K}(x_1,H(a_1,a_2)) \\ &\Leftrightarrow x_2 = \widetilde{K}_2(x_1,H(a_1,a_2)). \end{aligned}$$

Il ne reste plus qu'à définir la fonction $K : \mathscr{U}_1 \to \mathscr{U}_2$ en posant $K(x_1) = \widetilde{K}_2(x_1,H(a_1,a_2))$. Comme $\widetilde{K}$ est de classe $\mathscr{C}^1$, il en est de même pour $\widetilde{K}_2$ et donc aussi pour K. □

Comme toujours, via l'introduction de bases pour E_1, E_2 et F, on peut supposer que $E_1 = \mathbb{R}^n$ et $E_2 = F = \mathbb{R}^m$. Tout $x = (x_1,x_2) \in E_1 \times E_2$ peut s'écrire $x_1 = (x_1^1,x_1^2,\ldots,x_1^n)$ et $x_2 = (x_2^1,x_2^2,\ldots,x_2^p)$ et tout $y \in F$, $y = (y_1,\ldots,y_p)$. La matrice Jacobienne de H se décompose par blocs comme :

$$DH(x_1,x_2) = (D_1H(x_1,x_2),D_2H(x_1,x_2)),$$

avec :

$$D_1H(x_1,x_2) = \begin{pmatrix} \frac{\partial H_1}{\partial x_1^1}(x_1,x_2) & \ldots & \frac{\partial H_1}{\partial x_1^n}(x_1,x_2) \\ \vdots & & \vdots \\ \frac{\partial H_p}{\partial x_1^1}(x_1,x_2) & \ldots & \frac{\partial H_p}{\partial x_1^n}(x_1,x_2) \end{pmatrix}$$

et

$$D_2H(x_1,x_2) = \begin{pmatrix} \frac{\partial H_1}{\partial x_2^1}(x_1,x_2) & \ldots & \frac{\partial H_1}{\partial x_2^p}(x_1,x_2) \\ \vdots & & \vdots \\ \frac{\partial H_p}{\partial x_2^1}(x_1,x_2) & \ldots & \frac{\partial H_p}{\partial x_2^p}(x_1,x_2) \end{pmatrix}.$$

On a alors, pour tout $(h_1,h_2) \in E_1 \times E_2$,

$$D_1H(x_1,x_2)h_1 = DH(x_1,x_2)(h_1,0) \quad \text{et} \quad D_2H(x_1,x_2)h_2 = DH(x_1,x_2)(0,h_2).$$

Les hypothèses du théorème des fonctions implicites se traduisent par la condition :

$$\det D_2H(a_1,a_2) \neq 0.$$

D'autre part, pour tout $x_1 \in \mathscr{U}_1$:

$$H(x_1,K(x_1)) = H(a_1,a_2).$$

En différentiant cette relation au point a_1 de $\mathscr{U}_1$, on obtient que pour tout $h_1 \in E_1$:

$$DH(a_1,K(a_1))(h_1,DK(a_1)h_1) = 0.$$

Mais $K(a_1) = a_2$ et la relation ci-dessus peut encore s'écrire :

$$D_1H(a_1,a_2)h_1 + D_2H(a_1,a_2) \circ DK(a_1)h_1 = 0.$$

Comme par hypothèse, l'application $D_2H(a_1,a_2)$ est inversible dans $\mathscr{L}(E_2,F)$, on obtient :

Corollaire 4.1. *Sous les hypothèses du théorème des fonctions implicites, l'expression de la différentielle de l'application K au point a_1 est :*

$$DK(a_1) = -D_2H(a_1,a_2)^{-1} \circ D_1H(a_1,a_2),$$

dans $\mathscr{L}(E_1,E_2)$.

Exemple 4.3. Reconsidérons l'exemple de la fonction

$$\begin{array}{rcl} H : \mathbb{R}^2 & \to & \mathbb{R} \\ (x_1,x_2) & \mapsto & x_1^2 + x_2^2, \end{array}$$

avec $(a_1,a_2) = (0,1)$. Dans ce cas, on a vu que pour tout $(h_1,h_2) \in \mathbb{R}\times\mathbb{R}$ on avait $D_2H(a_1,a_2)h_2 = 2a_2h_2 = 2h_2$ et $D_1H(a_1,a_2)h_1 = 2a_1h_1 = 0$. Avec le corollaire, on trouve donc que $K'(0) = 0$.

Pour terminer ce chapitre, remarquons que le théorème d'inversion locale et le théorème des fonctions implicites s'énoncent de la même façon lorsque les espaces sont de dimension infinie.

Chapitre 5

Différentielles d'ordre supérieur

5.1 Différentielle seconde

Soient $(E,\|\cdot\|_E)$ et $(F,\|\cdot\|_F)$ des ev normés de dimension finie, $\mathscr{U}$ un ouvert de E et $H : \mathscr{U} \to F$ une application de classe $\mathscr{C}^1$. La différentielle de H est une application :

$$a \in \mathscr{U} \mapsto DH(a) \in \mathscr{L}(E,F),$$

et cette application est continue par définition des applications de classe $\mathscr{C}^1$.

Définition 5.1. *La fonction H est de classe $\mathscr{C}^2$ lorsque la fonction $DH : \mathscr{U} \to \mathscr{L}(E,F)$ est de classe $\mathscr{C}^1$.*

En tout point $a \in \mathscr{U}$, l'application $D(DH)(a)$ est donc dans $\mathscr{L}(E,\mathscr{L}(E,F))$. Nous avons montré dans le chapitre 1 que cet espace était isométrique à $\mathscr{L}_2(E \times E,F)$ que nous noterons plus simplement $\mathscr{L}_2(E,F)$.

Définition 5.2. *Soit $H : \mathscr{U} \to F$ une application de classe $\mathscr{C}^2$. Alors pour tout $a \in \mathscr{U}$, l'application $D(DH)(a)$ sera considérée comme un élément de $\mathscr{L}_2(E,F)$ et sera notée $D^2H(a)$.*

Exemple 5.1. Soit $H : \mathbb{R}^2 \to \mathbb{R}$ définie par $H(x_1,x_2) = 2x_1^2x_2$ pour tout $(x_1,x_2) \in \mathbb{R}^2$. Alors H est de classe $\mathscr{C}^2$ et l'expression de la différentielle est :

$$\begin{array}{rcl} DH(x_1,x_2) : \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (h_1,h_2) & \longmapsto & (4x_1x_2 \quad 2x_1^2)\begin{pmatrix} h_1 \\ h_2 \end{pmatrix}. \end{array}$$

Pour tout (x_1,x_2) dans $\mathbb{R}^2$, l'expression de la différentielle seconde est :

$$\begin{array}{rcl} D^2H(x_1,x_2) : \mathbb{R}^2 \times \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ ((h_1,h_2),(k_1,k_2)) & \longmapsto & (k_1 \quad k_2)\begin{pmatrix} 4x_2 & 4x_1 \\ 4x_1 & 0 \end{pmatrix}\begin{pmatrix} h_1 \\ h_2 \end{pmatrix}. \end{array}$$

Le caractère $\mathscr{C}^2$ d'une fonction se déduit directement de la régularité de ses dérivées partielles. Soient $\{e_1,\ldots,e_m\}$ une base de E, $\{f_1,\ldots,f_n\}$ une base de F et soit une fonction $H : \mathscr{U} \to F$ ($\mathscr{U}$

un ouvert de E). Dans ces deux bases, H s'écrit :

$$H(x) = \sum_{j=1}^{n} H_j(x_1e_1 + \ldots + x_me_m)f_j \qquad \text{pour tout } x \in \mathscr{U}.$$

Proposition 5.1. *La fonction H est de classe $\mathscr{C}^2$ si et seulement si, pour tous indices $1 \leqslant i,j \leqslant m$ et $1 \leqslant k \leqslant n$:*

1. *Les dérivées partielles $\frac{\partial H_k}{\partial x_i}(x)$ existent pour tout $x \in \mathscr{U}$;*
2. *Les dérivées partielles $\frac{\partial}{\partial x_j}\left(\frac{\partial H_k}{\partial x_i}\right)(x)$ existent pour tout $x \in \mathscr{U}$;*
3. *Les applications $x \in \mathscr{U} \mapsto \frac{\partial}{\partial x_j}\left(\frac{\partial H_k}{\partial x_i}\right)(x)$ sont continues.*

En termes de coordonnées, la différentielle seconde de H en un point $a \in \mathscr{U}$ dans les directions $h = \sum_{i=1}^m h_ie_i$ et $h' = \sum_{i=1}^m h'_ie_i$ est donnée par :

$$D^2H(a)(h,h') = \sum_{k=1}^{n}\left(\sum_{i,j=1}^{m}\frac{\partial}{\partial x_j}\left(\frac{\partial H_k}{\partial x_i}\right)(a)h_ih'_j\right)f_k.$$

Démonstration. Appliquer le Corollaire 2.1 à la fonction DH sur $\mathscr{U}$. □

Théorème 5.1 (Lemme de Schwarz)**.** *Soit $H : \mathscr{U} \to F$ une application de classe $\mathscr{C}^2$. Alors pour tout $a \in \mathscr{U}$, l'application $D^2H(a) \in \mathscr{L}_2(E,F)$ est symétrique c'est à dire :*

$$D^2H(a)(h,h') = D^2H(a)(h',h) \qquad \textit{pour tout } h,h' \in E.$$

En coordonnées, ceci se traduit par :

$$\frac{\partial}{\partial x_j}\left(\frac{\partial H_k}{\partial x_i}\right)(a) = \frac{\partial}{\partial x_i}\left(\frac{\partial H_k}{\partial x_j}\right)(a) \qquad \textit{pour tout } 1 \leqslant i,j \leqslant m \textit{ et } 1 \leqslant k \leqslant n.$$

Démonstration. Il suffit de considérer le cas $E = \mathbb{R}^2$ et $F = \mathbb{R}$ (on fait le même raisonnement pour chaque fonction H_k à valeurs dans $\mathbb{R}$ et pour chaque couple de coordonnées (x_i,x_j)). Pour simplifier les notations, nous supposerons également que $a = (0,0)$.

Soit $\delta > 0$ tel que $[-\delta,\delta] \times [-\delta,\delta] \subset \mathscr{U}$. Pour tout réel u tel que $|u| < \delta$, on pose :

$$\Delta(u) = H(u,u) - H(u,0) - H(0,u) + H(0,0).$$

On peut alors écrire que :

$$H(u,u) - H(u,0) = u\int_0^1 \frac{\partial H}{\partial x_2}(u,tu)\,\mathrm{d}t, \tag{5.1a}$$

$$H(0,u) - H(0,0) = u\int_0^1 \frac{\partial H}{\partial x_2}(0,tu)\,\mathrm{d}t, \tag{5.1b}$$

et donc :

$$\Delta(u) = u\int_0^1 \frac{\partial H}{\partial x_2}(u,tu) - \frac{\partial H}{\partial x_2}(0,tu)\,\mathrm{d}t = u^2\int_0^1\int_0^1 \frac{\partial}{\partial x_1}\left(\frac{\partial H}{\partial x_2}\right)(su,tu)\,\mathrm{d}s\mathrm{d}t.$$

On en déduit que :

$$\frac{\Delta(u)}{u^2} - \frac{\partial}{\partial x_1}\left(\frac{\partial H}{\partial x_2}\right)(0,0) = \int_0^1\int_0^1 \left[\frac{\partial}{\partial x_1}\left(\frac{\partial H}{\partial x_2}\right)(su,tu) - \frac{\partial}{\partial x_1}\left(\frac{\partial H}{\partial x_2}\right)(0,0)\right] \mathrm{d}s\mathrm{d}t.$$

La continuité de la fonction $x \mapsto \frac{\partial}{\partial x_1}\left(\frac{\partial H}{\partial x_2}\right)(x)$ en $(0,0)$ entraîne ensuite que pour tout $\varepsilon > 0$, on peut choisir $\delta > 0$ assez petit de telle sorte que pour tout $(x_1,x_2) \in]-\delta,\delta[\times]-\delta,\delta[$:

$$\left|\frac{\partial}{\partial x_1}\left(\frac{\partial H}{\partial x_2}\right)(x_1,x_2) - \frac{\partial}{\partial x_1}\left(\frac{\partial H}{\partial x_2}\right)(0,0)\right| < \varepsilon.$$

On en déduit dans un premier temps que :

$$\left|\frac{\Delta(u)}{u^2} - \frac{\partial}{\partial x_1}\left(\frac{\partial H}{\partial x_2}\right)(0,0)\right| < \varepsilon,$$

puis que :

$$\lim_{\substack{u\to 0\\ u\neq 0}} \frac{\Delta(u)}{u^2} = \frac{\partial}{\partial x_1}\left(\frac{\partial H}{\partial x_2}\right)(0,0).$$

À la place des égalités (5.1), on peut tout aussi bien écrire aussi que :

$$H(u,u) - H(0,u) = u\int_0^1 \frac{\partial H}{\partial x_1}(tu,u)\,\mathrm{d}t,$$
$$H(u,0) - H(0,0) = u\int_0^1 \frac{\partial H}{\partial x_1}(tu,0)\,\mathrm{d}t,$$

ce qui conduit à l'expression :

$$\Delta(u) = u^2\int_0^1\int_0^1 \frac{\partial}{\partial x_2}\left(\frac{\partial H}{\partial x_1}\right)(su,tu)\,\mathrm{d}s\mathrm{d}t.$$

Les mêmes arguments que ceux utilisés précédemment, nous amènent enfin à la conclusion :

$$\lim_{\substack{u\to 0\\ u\neq 0}} \frac{\Delta(u)}{u^2} = \frac{\partial}{\partial x_2}\left(\frac{\partial H}{\partial x_1}\right)(0,0),$$

et le théorème est démontré. □

On utilise quelquefois ce théorème pour montrer qu'une application n'est pas de classe $\mathscr{C}^2$.

Définition 5.3 (Matrice Hessienne)**.** *Soit $H : \mathscr{U} \to \mathbb{R}$. On suppose que H est de classe $\mathscr{C}^2$ sur $\mathscr{U}$ et que a est un point de $\mathscr{U}$. On introduit une base $\{e_1,\ldots,e_m\}$ de E. Alors la matrice :*

$$\begin{pmatrix} \frac{\partial^2 H}{\partial x_1^2}(a) & \frac{\partial^2 H}{\partial x_2\partial x_1}(a) & \ldots & \frac{\partial^2 H}{\partial x_m\partial x_1}(a) \\ \frac{\partial^2 H}{\partial x_1\partial x_2}(a) & \frac{\partial^2 H}{\partial^2 x_2}(a) & \ldots & \frac{\partial^2 H}{\partial x_m\partial x_2}(a) \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 H}{\partial x_1\partial x_m}(a) & \frac{\partial^2 H}{\partial x_2\partial x_m}(a) & \ldots & \frac{\partial^2 H}{\partial^2 x_m}(a) \end{pmatrix},$$

est symétrique et s'appelle la matrice Hessienne de H au point a.

Proposition 5.2. *Les propositions 2.3 (somme de fonctions de classe $\mathscr{C}^1$), 2.4 (produit de composition de fonctions de classe $\mathscr{C}^1$) et 2.6 (inverse d'une fonction de classe $\mathscr{C}^1$) s'énoncent de la même façon pour des fonctions de classe $\mathscr{C}^2$.*

Démonstration. Laissée en exercice. □

Exemple 5.2.

1. Toute application $\ell \in \mathscr{L}(E,F)$ est de classe $\mathscr{C}^2$ et $D^2\ell(a) = 0$ pour tout $a \in E$. En effet, $D\ell(a) = \ell$ pour tout $a \in E$ donc $a \in E \mapsto D\ell(a) \in \mathscr{L}(E,F)$ est une application constante. Sa différentielle est nulle.
2. Toute application $b \in \mathscr{L}_2(E \times F,G)$ (E, F et G des ev normés) est de classe $\mathscr{C}^2$ et pour tout $(a,a') \in E \times F$:

$$D^2b(a,a')((h_1,h_1'),(h_2,h_2')) = b(h_1,h_2') + b(h_2,h_1') \qquad \text{pour tout } (h_k,h_k') \in E \times F, k = 1,2.$$

En effet, on sait que $Db(a,a')(h,h') = b(a,h')+b(h,a')$ pour tout $(a,a') \in E\times F$ et $(h,h') \in E\times F$ et donc l'application :

$$(a,a') \mapsto Db(a,a') \in \mathscr{L}(E \times F,G),$$

est linéaire. Sa différentielle en tout point est donc elle même.

5.2 Formule de Taylor

Soit $H : \mathscr{U} \to F$ une application où $\mathscr{U}$ est un ouvert de E. Soient $a \in \mathscr{U}$ et $h \in E$ tels que le segment $[a,a+h]$ soit inclus dans $\mathscr{U}$.

Théorème 5.2 (Formule de Taylor avec reste intégrale)**.** *On suppose que H est de classe $\mathscr{C}^2$ sur $\mathscr{U}$. Alors :*

$$H(a+h) = H(a) + DH(a)h + \int_0^1 (1-t)D^2H(a+th)(h,h)\,\mathrm{d}t. \tag{5.2}$$

Démonstration. Il suffit de montrer la formule précédente pour chacune des composantes de H (en introduisant une base de F). On peut donc considérer que $F = \mathbb{R}$. Posons $g(t) = H(a+th)$ pour tout $t \in [0,1]$. L'application g est alors de classe $\mathscr{C}^2$ et :

$$g'(t) = DH(a+th)h, \qquad g''(t) = D^2H(a+th)(h,h).$$

La formule de Taylor (avec reste intégral) pour la fonction g s'écrit :

$$g(1) = g(0) + g'(0) + \int_0^1 (1-t)g''(t)\mathrm{d}t,$$

ce qui donne bien la formule du théorème. □

Plutôt que la formule (5.2), on utilisera le plus souvent :

Corollaire 5.1 (formule de Taylor–Young)**.** *Soit $H : \mathscr{U} \to F$ une application de classe $\mathscr{C}^2$ et soit $a \in \mathscr{U}$. Alors il existe une application ε définie sur un voisinage de 0_E et à valeurs dans*

F telle que, pour tout h assez petit :

$$H(a+h) = H(a) + DH(a)h + \frac{1}{2}D^2H(a)(h,h) + \|h\|_E^2\varepsilon(h).$$

Démonstration. Pour h assez petit, le segment $[a,a+h]$ est inclus dans $\mathscr{U}$ et on peut appliquer le théorème pour obtenir que :

$$H(a+h) = H(a) + DH(a)h + \frac{1}{2}D^2H(a)(h,h) + r(h),$$

avec

$$r(h) = \int_0^1 (1-t)D^2H(a+th)(h,h)\,\mathrm{d}t - \frac{1}{2}D^2H(a)(h,h).$$

Comme

$$\int_0^1 (1-t)\,\mathrm{d}t = \frac{1}{2},$$

on déduit que :

$$r(h) = \int_0^1 (1-t)\big[D^2H(a+th)(h,h) - D^2H(a)(h,h)\big]\,\mathrm{d}t.$$

Par continuité de l'application $x \mapsto D^2H(x) \in \mathscr{L}_2(E,F)$, pour tout $\varepsilon > 0$ il existe $\delta > 0$ tel que si $\|x\|_E < \delta$ alors $\|D^2H(a+x) - D^2H(a)\|_{\mathscr{L}_2(E,F)} < \varepsilon$. Ceci signifie (par définition de la norme dans $\mathscr{L}_2(E,F)$) que :

$$\|D^2H(a+x)(h,h) - D^2H(a)(h,h)\|_F \leqslant \varepsilon\|h\|_E^2 \qquad \text{pour tout } h \in E.$$

Il s'en suit que pour tout h tel que $\|h\|_E < \delta$, on a $\|r(h)\|_F \leqslant \varepsilon\|h\|_E^2$, ce qui est le résultat annoncé. □

Ce résultat sera très utile pour la recherche d'extrema locaux. Pour le moment, considérons l'exemple suivant :

Exemple 5.3. Soit $H : \mathbb{R}^2 \to \mathbb{R}$ définie par $H(x,y) = \sin(\pi xy^2)$. Alors H est de classe $\mathscr{C}^2$ sur $\mathbb{R}^2$. Le gradient de H au point (x,y) est :

$$\nabla H(x,y) = \pi y\big(y\cos(\pi xy^2) \quad 2x\cos(\pi xy^2)\big),$$

On trouve également l'expression de la matrice Hessienne :

$$\mathcal{H}(x,y) = \pi\begin{pmatrix} -\pi y^4\sin(\pi xy^2) & 2\pi y[\cos(\pi xy^2) - xy^2\sin(\pi xy^2)] \\ 2y[\cos(\pi xy^2) - \pi xy^2\sin(\pi xy^2)] & 2x[\cos(\pi xy^2) - 2\pi xy^2\sin(\pi xy^2)] \end{pmatrix}.$$

La différentielle seconde au point (x,y) est l'application bilinéaire :

$$\begin{array}{rcl} D^2H(x,y) : \mathbb{R}^2 \times \mathbb{R}^2 & \to & \mathbb{R} \\ ((h,h'),(k,k')) & \mapsto & (h \;\; h')\mathcal{H}(x,y)\begin{pmatrix} k \\ k' \end{pmatrix}. \end{array}$$

On en déduit que, au voisinage de tout point (a,b) de $\mathbb{R}^2$ (d.l. à l'ordre 2) :

$$H(x,y) = H(a,b) + \nabla H(a,b) \cdot (x-a,y-b) + \frac{1}{2}(x-a,y-b)\mathcal{H}(a,b)\begin{pmatrix} x-a \\ y-b \end{pmatrix} + \|(x-a,y-b)\|_2^2 \varepsilon(x-a,y-b).$$

En particulier, la fonction P (polynôme de degré 2) définie pour tout $(x,y) \in \mathbb{R}^2$ par :

$$P(x,y) = H(a,b) + \nabla H(a,b) \cdot (x-a,y-b) + \frac{1}{2}(x-a,y-b)\mathcal{H}(a,b)\begin{pmatrix} x-a \\ y-b \end{pmatrix},$$

est une "très bonne" approximation de la fonction H au voisinage du point (a,b). Lorsque $a = 1/2$ et $b = 1/2$, le graphe de la fonction H et celui du polynôme P sont tracés sur la figure 5.1.

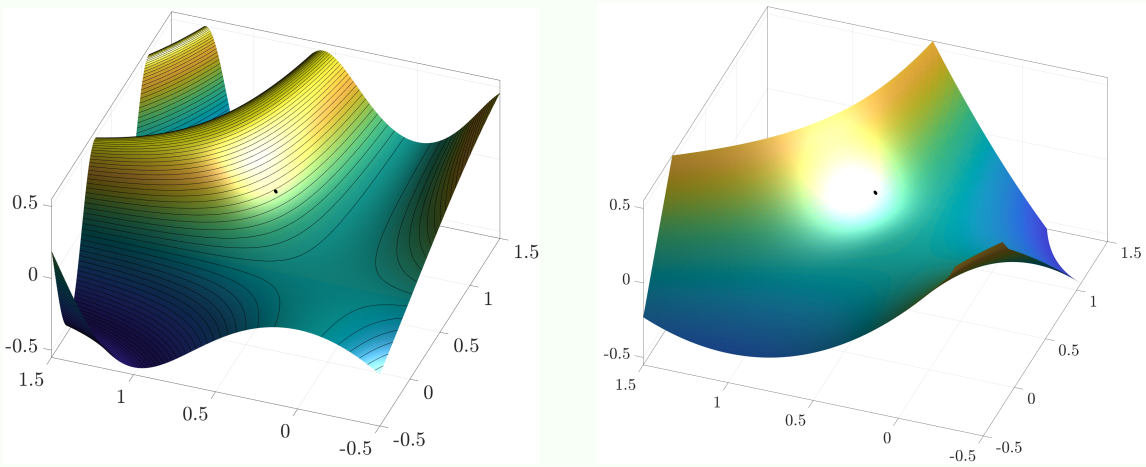


FIG. 5.1 – *Graphe de la fonction H (à gauche) et du polynôme P (à droite) obtenue à partir du développement limité à l'ordre 2 au voisinage du point* $(1/2,1/2)$ *(marqué en noir sur les dessins).*

Les différentielles d'ordre supérieur à 2 ne sont plus au programme mais on peut facilement comprendre que si l'application :

$$D^2H : \mathscr{U} \to \mathscr{L}_2(E,F),$$

est différentiable au point a, sa différentielle sera la différentielle d'ordre 3 de H, notée $D^3H(a)$. Cette différentielle sera dans $\mathscr{L}(E,\mathscr{L}_2(E,F))$ que l'on va identifier avec $\mathscr{L}_3(E,F)$, c'est à dire l'espace des applications trilinéaires de E dans F.

Tous les résultats de ce chapitre (en dehors de ceux concernant la matrice Hessienne) sont encore vérifiés en dimension infinie.

Chapitre 6

Points critiques et extrema

On considère la situation suivante : E est une ev normé de dimension finie, $\mathscr{U}$ un ouvert de E et $H : \mathscr{U} \to \mathbb{R}$ une fonction de classe $\mathscr{C}^1$. On se propose de décrire l'allure de la fonction H au voisinage d'un point a de $\mathscr{U}$.

6.1 Points réguliers et points critiques

Définition 6.1. *On dit que a est un point régulier de H si $DH(a) \neq 0$. Un point qui n'est pas régulier est appelé point critique.*

On introduit une base $\{e_1, \ldots, e_m\}$ de E. On peut alors calculer le gradient de H au point a :

$$\nabla H(a) = \left(\frac{\partial H}{\partial x_1}(a), \frac{\partial H}{\partial x_2}(a), \ldots, \frac{\partial H}{\partial x_m}(a)\right).$$

D'après la définition, a est un point critique si et seulement si :

$$\frac{\partial H}{\partial x_1}(a) = \frac{\partial H}{\partial x_2}(a) = \ldots = \frac{\partial H}{\partial x_m}(a) = 0.$$

Supposons que a soit un point régulier. Alors le noyau de l'application linéaire $DH(a)$ est de dimension $m-1$.

Remarque 6.1. *Rappelons le théorème du rang pour une application $\ell \in \mathscr{L}(E,F)$:*

$$\dim \operatorname{Im} \ell + \dim \ker \ell = \dim E.$$

On peut alors choisir la base $\{e_1, \ldots, e_n\}$ de telle sorte que $\{e_2, \ldots, e_m\}$ soit une base de $\ker DH(a)$ et qu'on ait $DH(a)e_1 = 1$. En termes de dérivées partielles, nous aurons donc :

$$\frac{\partial H}{\partial x_1}(a) = 1 \quad \text{et} \quad \frac{\partial H}{\partial x_2}(a) = \ldots = \frac{\partial H}{\partial x_m}(a) = 0.$$

Ceci conduit au d.l. à l'ordre 1 :

$$H(a+h) = H(a) + h_1 + \|h\|_E \varepsilon(h),$$

avec $h = \sum_{j=1}^m h_j e_j$.

On suppose maintenant que a est un point critique de H et que H est une fonction de classe $\mathscr{C}^2$. Un d.l. à l'ordre 2 nous donne, pour h dans un voisinage de 0_E :

$$H(a+h) = H(a) + \frac{1}{2}D^2H(a)(h,h) + \|h\|_E^2\varepsilon(h),$$

avec $\varepsilon : E \to \mathbb{R}$ telle que $\varepsilon(h) \to 0$ quand $h \to 0_E$. Rappelons que $D^2H(a)$ est une forme bilinéaire symétrique d'après le lemme de Schwarz. Le comportement de H au voisinage de a est donc donné par la forme quadratique :

$$h \in E \mapsto \frac{1}{2}D^2H(a)(h,h) \in \mathbb{R}.$$

6.2 Formes quadratiques

Soit $B \in \mathscr{L}_2(E,\mathbb{R})$. Lorsque B est symétrique, on peut définir la forme quadratique Q associée par :

$$Q(x) = B(x,x) \qquad \text{pour tout } x \in E.$$

On peut retrouver l'expression de B à partir de Q grâce aux l'identités de polarisation :

$$B(x,y) = \frac{1}{2}\big[Q(x+y) - Q(x) - Q(y)\big] = \frac{1}{4}\big[Q(x+y) - Q(x-y)\big].$$

Introduisons une base $\{e_1,\ldots,e_m\}$ de E. Alors tout $x \in E$ s'écrit $x = \sum_{j=1}^m x_je_j$ et on a :

$$B(x,x) = \sum_{i,j=1}^m B(e_i,e_j)x_ix_j.$$

D'un point de vue matricielle, on peut introduire la matrice symétrique :

$$\mathcal{B} = \begin{pmatrix} B(e_1,e_1) & B(e_1,e_2) & \ldots & B(e_1,e_m) \\ B(e_2,e_1) & B(e_2,e_2) & \ldots & B(e_2,e_m) \\ \vdots & \vdots & & \vdots \\ B(e_m,e_1) & B(e_m,e_2) & \ldots & B(e_m,e_m) \end{pmatrix},$$

et on a alors :

$$B(x,y) = (x_1,\ldots,x_m)\mathcal{B}\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}.$$

Rappelons que les valeurs propres d'une matrice symétrique sont toutes réelles. Nous admettrons :

Proposition 6.1 (Réduction de Gauss et signature)**.** *Pour toute forme quadratique Q sur E, il existe une base $\{e_1,\ldots,e_m\}$ de E et deux entiers naturels p^+ et p^- tels que $p^+ + p^- \leqslant m$ et :*

$$Q\left(\sum_{j=1}^m x_je_j\right) = \sum_{i=1}^{p^+} x_i^2 - \sum_{i=1+p^+}^{p^++p^-} x_i^2.$$

Le couple (p^+,p^-) *s'appelle la signature de la forme quadratique. Si* B *est la forme bilinéaire associée à* Q *et* $\mathcal{B}$ *la matrice de* B*, alors il existe une matrice inversible* P *telle que :*

$$
{}^tP\mathcal{B}P = \begin{bmatrix}
1 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\
0 & \ddots & \ddots & & & & & & \vdots \\
\vdots & \ddots & 1 & \ddots & & & & & \vdots \\
\vdots & & \ddots & -1 & \ddots & & & & \vdots \\
\vdots & & & \ddots & \ddots & \ddots & & & \vdots \\
\vdots & & & & \ddots & -1 & \ddots & & \vdots \\
\vdots & & & & & \ddots & 0 & \ddots & \vdots \\
\vdots & & & & & & \ddots & 0 & \vdots \\
0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 & 0
\end{bmatrix},
$$

avec bien entendu p^+ *cœfficients* $+1$ *et* p^- *cœfficients* -1 *sur la diagonale.*

Plutôt que de faire une démonstration dans le cas général, traitons un exemple.

Exemple 6.1. Soit $Q : \mathbb{R}^3 \to \mathbb{R}$ une forme quadratique définie dans la base canonique $\{e_1,e_2,e_3\}$ par :

$$Q(x) = x_1^2 - x_2^2 + \frac{5}{4}x_3^2 + 2x_1x_2 - x_1x_3 + 3x_2x_3.$$

La matrice de la forme bilinéaire B associée est :

$$\mathcal{B} = \begin{pmatrix} 1 & 1 & -1/2 \\ 1 & -1 & 3/2 \\ -1/2 & 3/2 & 5/4 \end{pmatrix}.$$

On a alors :

$$\begin{aligned} Q(x) &= \left(x_1 + x_2 - \frac{1}{2}x_3\right)^2 - 2x_2^2 + x_3^2 + 4x_2x_3 \\ &= \left(x_1 + x_2 - \frac{1}{2}x_3\right)^2 + (x_3 + 2x_2)^2 - 6x_2^2 \end{aligned}$$

Le changement de variables :

$$x_1' = x_1 + x_2 - \frac{1}{2}x_3, \qquad x_2' = 2x_2 + x_3 \quad \text{and} \quad x_3' = \sqrt{6}x_2,$$

conduit à l'introduction de la nouvelle base :

$$e_1' = e_1, \qquad e_2' = \frac{1}{2}e_1 + e_3 \quad \text{and} \quad e_3' = -\frac{\sqrt{2}}{\sqrt{3}}e_1 + \frac{1}{\sqrt{6}}e_2 - \frac{\sqrt{2}}{\sqrt{3}}e_3,$$

dans laquelle l'expression de Q est :

$$Q(x') = (x_1')^2 + (x_2')^2 - (x_3')^2.$$

Notons P la matrice de passage de la base $\{e_1,e_2,e_3\}$ à la base $\{e'_1,e'_2,e'_3\}$, c'est à dire :

$$P = \begin{pmatrix} 1 & 1/2 & -\sqrt{2}/\sqrt{3} \\ 0 & 0 & 1/\sqrt{6} \\ 0 & 1 & -\sqrt{2}/\sqrt{3} \end{pmatrix}.$$

On a alors bien :

$${}^tP\mathcal{B}P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Définition 6.2. *Soit Q une forme quadratique sur un evn E de dimension m. Soient p^+ et p^- comme dans la proposition 6.1. On dira que:*

- *Q est positive si $p^- = 0$;*
- *Q est négative si $p^+ = 0$;*
- *Q est indéfinie si $p^+ \neq 0$ et $p^- \neq 0$;*
- *Q est non dégénérée si $p^+ + p^- = m$. Dans ce cas on peut distinguer :*
  - *Q est définie positive si $p^+ = m$;*
  - *Q est définie négative si $p^- = m$;*

Ces propriétés se traduisent sur la forme bilinéaire B associée et sa matrice $\mathcal{B}$ comme suit :

- Q est positive si $B(x,x) \geqslant 0$ pour tout $x \in E$. Toutes les valeurs propres de $\mathcal{B}$ sont positives ou nulles.
- Q est négative si $B(x,x) \leqslant 0$ pour tout $x \in E$. Toutes les valeurs propres de $\mathcal{B}$ sont négatives ou nulles.
- Q est indéfinie s'il existe $x \in E$ et $y \in E$ tels que $Q(x) > 0$ et $Q(y) < 0$. La matrice $\mathcal{B}$ a au moins une valeur propre strictement positive et une valeur propre strictement négative.
- Q est non dégénérée si $x = 0_E$ est le seul élément de E tel que $B(x,y) = 0$ pour tout $y \in E$. La matrice $\mathcal{B}$ n'a aucune valeur propre nulle. On distingue alors :
  - Q est définie positive si $Q(x) > 0$ pour tout $x \neq 0_E$. Toutes les valeurs propres de $\mathcal{B}$ sont strictement positives.
  - Q est définie négative si $Q(x) < 0$ pour tout $x \neq 0_E$. Toutes les valeurs propres de $\mathcal{B}$ sont strictement négatives.

6.3 Extrema

Définition 6.3. *Un point a d'une partie ouverte $\mathscr{U}$ de E est un maximum (respectivement minimum) local d'une fonction $H : \mathscr{U} \to \mathbb{R}$ s'il existe un ouvert $\mathscr{V}$ contenant a tel que :*

$$H(x) \leqslant H(a) \quad (\textit{respectivement } \geqslant) \qquad \textit{pour tout } x \in \mathscr{V}.$$

On parle de max (respect. min) local strict lorsque les inégalités sont strictes. Un extremum est un maximum ou un minimum.

La proposition suivante nous dit que les maxima et les minima locaux sont à chercher parmi les points critiques de la fonction.

Proposition 6.2. *Soient $H : \mathscr{U} \to \mathbb{R}$ une fonction différentiable et a un point de $\mathscr{U}$. Si a est un extremum local alors a est un point critique de H. Lorsque H est de classe $\mathscr{C}^2$, la forme quadratique associée à la forme bilinéaire symétrique $D^2H(a)$ est positive si a est un minimum local et négative si a est un maximum local.*

Démonstration. Soit $h \in E$, $h \neq 0_E$. Posons $\mathscr{V}(h) = \{t \in \mathbb{R} \,:\, a + th \in \mathscr{U}\}$ (un ouvert de $\mathbb{R}$ qui contient 0) et pour tout t dans $\mathscr{V}(h)$:

$$g_h(t) = H(a + th).$$

La fonction $g_h : \mathscr{V}(h) \to \mathbb{R}$ est dérivable sur $\mathscr{V}(h)$, $g_h'(0) = DH(a)h$ et g_h admet un extremum local en $t = 0$. Donc $DH(a)h = 0$. Ceci étant vérifié pour tout $h \in E$, on en déduit que $DH(a) = 0$.

Si H est de classe $\mathscr{C}^2$, alors g_h est également de classe $\mathscr{C}^2$ sur $\mathscr{V}(h)$ et $g_h''(0) = D^2H(a)(h,h)$. Si a est un maximum local de H alors 0 est un maximum local de g_h et donc $g_h''(0) \leqslant 0$. Encore une fois, ceci étant vrai pour tout $h \in E$, on en déduit que la forme quadratique associée à $D^2H(a)$ est négative. On procède de la même façon lorsque a est un minimum local. □

Il faut en fait un peu plus que la simple positivité (ou la négativité) de la forme quadratique associée à la Hessienne pour assurer que le point est bien un extremum local.

Proposition 6.3. *Soit $H : \mathscr{U} \to \mathbb{R}$ une fonction de classe $\mathscr{C}^2$ et $a \in \mathscr{U}$ un point critique de H. Si la forme quadratique associée à la forme bilinéaire $D^2H(a)$ est définie positive (respect. négative), alors a est est un minimum (respect. maximum) local strict de H.*

Démonstration. Notons $Q(v) = D^2H(a)(v,v)$ pour tout $v \in E$ et supposons que Q soit définie positive. Posons :

$$\Delta = \min_{\substack{v \in E \\ \|v\|_E = 1}} Q(v).$$

Alors $\Delta > 0$ car la sphère unité est compacte et $Q(v) > 0$ pour tout v sur cette sphère. D'autre part, pour tout $v \neq 0_E$:

$$Q(v) = D^2H(a)(v,v) = \|v\|_E^2 D^2H(a)\left(\frac{v}{\|v\|_E}, \frac{v}{\|v\|_E}\right) = \|v\|_E^2 Q\left(\frac{v}{\|v\|_E}\right) \geqslant \Delta \|v\|_E^2.$$

Un développement de Taylor–Young à l'ordre 2 au point a nous donne :

$$H(a + v) = H(a) + \frac{1}{2} D^2H(a)(v,v) + \|v\|_E^2 \varepsilon(v) \geqslant H(a) + \|v\|_E^2 \left(\frac{1}{2}\Delta + \varepsilon(v)\right),$$

avec $\varepsilon(v) \to 0_E$ quand $\|v\|_E \to 0$. On en déduit que $H(a + v) > H(a)$ pour tout v assez petit. □

Remarque 6.2. *Si la forme quadratique est seulement positive (mais dégénérée) alors il faut une étude plus fine pour pouvoir conclure. Considérons les exemples suivants :*

$$H_1(x,y) = x^2 + y^4, \qquad H_2(x,y) = x^2, \qquad H_3(x,y) = x^2 - y^4.$$

Le point (0,0) *est un point critique pour chaque fonction H_k et pour tout $(h,h') \in \mathbb{R}^2$,*

$$D^2H_k(0,0)((h,h'),(h,h')) = 2h^2.$$

C'est une forme positive mais dégénérée. Or (0,0) *est un min local strict pour H_1, un min local (non strict) pour H_2 et n'est pas un extrema pour H_3.*

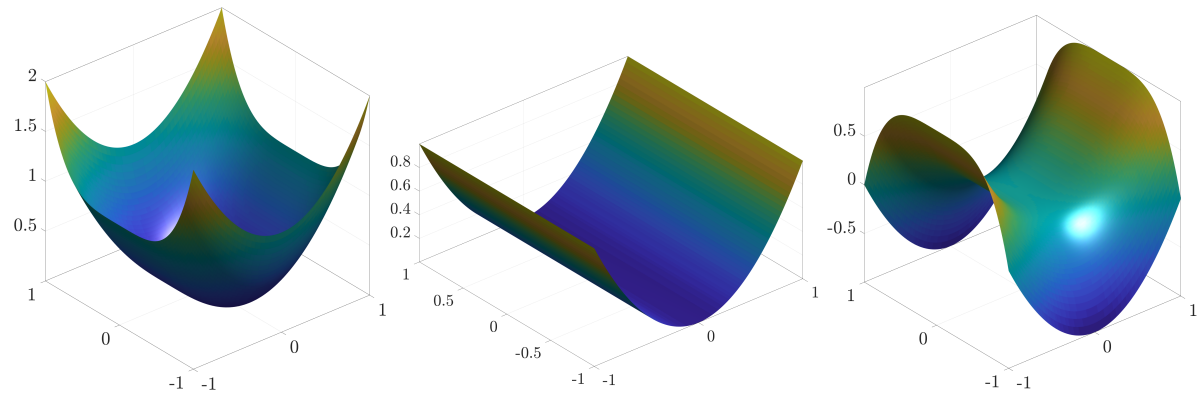


FIG. 6.1 – *Graphes des fonctions H_1, H_2 et H_3. Ces trois fonctions ont le même développement de Taylor–Young à l'ordre 2 en* (0,0).

6.4 Extrema liés

Soit $\mathscr{U}$ un ouvert de $\mathbb{R}^m$ et Soit $F : \mathscr{U} \to \mathbb{R}^k$ une fonction de classe $\mathscr{C}^2$ (avec $k < m$). On note :

$$\mathcal{M} = \{x \in \mathscr{U} \ : \ F(x) = 0\}. \tag{6.1}$$

Étant donnée $H : \mathscr{U} \to \mathbb{R}$ une fonction de classe $\mathscr{C}^2$, on cherche les extrema locaux de $x \mapsto H(x)$ lorsque x est astreint à rester dans l'ensemble $\mathcal{M}$.

Définition 6.4. *Un point $a \in \mathcal{M}$ est un minimum local de H sur $\mathcal{M}$ s'il existe un ouvert $\mathscr{V} \subset \mathscr{U}$ contenant a tel que :*

$$H(x) \geqslant H(a) \qquad \text{pour tout } x \in \mathscr{V} \cap \mathcal{M}.$$

On définit de même les notions de maximum local (et max local strict, min local strict avec des inégalités strictes).

6.4.1 Points réguliers

Proposition 6.4. *Soit a un point de $\mathcal{M}$. Les propriétés suivantes sont équivalentes :*

1. *La différentielle $DF(a)$ est surjective;*
2. *Les vecteurs $\nabla F_j(a)$ $(j = 1, \ldots, k)$ forment une famille libre;*
3. *La dimension de $\ker DF(a)$ est $m - k$.*

Démonstration. On sait que :

$$m = \dim \ker DF(a) + \dim \operatorname{Im} DF(a),$$

donc les formulations 1 et 3 sont équivalentes.

L'équivalence des points 2 et 1 découle de l'égalité :

$$\operatorname{Im} DF(a)^{\perp} = \Big\{ y = (y_1, \ldots, y_k) \in \mathbb{R}^k \ : \ \sum_{j=1}^{k} y_j \nabla F_j(a) = 0 \Big\}. \tag{6.2}$$

Par définition :

$$\operatorname{Im} DF(a)^{\perp} = \big\{ y \in \mathbb{R}^k \ : \ \langle y, DF(a)x \rangle_{\mathbb{R}^k} = 0 \quad \forall\, x \in \mathbb{R}^m \big\},$$

et l'on remarque que :

$$\langle y,DF(a)x\rangle_{\mathbb{R}^k} = \langle {}^tDF(a)y,x\rangle_{\mathbb{R}^m} = \Big\langle \sum_{j=1}^{k} y_j \nabla F_j(a),x \Big\rangle_{\mathbb{R}^m}.$$

L'égalité (6.2) s'en déduit. □

Définition 6.5. *Un point a de $\mathcal{M}$ qui vérifie l'une des propriété de la Proposition 6.4 est appelé un point régulier de $\mathcal{M}$. Dans ce cas, on note :*

$$\begin{aligned} T_a\mathcal{M} = \ker DF(a) &= \{u \in \mathbb{R}^2 \,:\, \nabla F_j(a)\cdot u = 0, \forall\, j = 1,\dots,k\}, \\ &= \langle \nabla F_1(a),\dots,\nabla F_k(a)\rangle^{\perp} = \cap_{j=1}^{k} \langle \nabla F_j(a)\rangle^{\perp}, \end{aligned}$$

l'espace tangent à $\mathcal{M}$ au point a (c'est donc un sous ev de $\mathbb{R}^m$ de dimension $m-k$).

Dans la suite, on ne s'intéressera qu'aux points réguliers de $\mathcal{M}$.

La premier point de la proposition 6.4 assure que l'on peut appliquer le théorème des fonctions implicites à la fonction F au voisinage de a. Un point régulier est donc un point au voisinage duquel $\mathcal{M}$ est le graphe d'une fonction $\mathscr{C}^1$ de $\mathbb{R}^{m-k}$ dans $\mathbb{R}^k$.

Proposition 6.5. *L'espace tangent à $\mathcal{M}$ en a peut aussi être défini comme :*

$$T_a\mathcal{M} = \big\{\gamma'(0) \,:\, \gamma :]-1,1[\to \mathcal{M} \text{ de classe } \mathscr{C}^1 \text{ telle que } \gamma(0) = a\big\}.$$

Démonstration. Soit $\gamma : t \in]-1,1[\mapsto \gamma(t) \in \mathbb{R}^k$ une courbe paramétrée $\mathscr{C}^1$ telle que $\gamma(t) \subset \mathcal{M}$ pour tout $t \in]-1,1[$. Alors $F(\gamma(t)) = 0$ pour tout t et en dérivant :

$$DF(\gamma(t))\gamma'(t) = 0,$$

c'est à dire en particulier pour $t = 0$, $\nabla F_j(a)\cdot\gamma'(0) = 0$ pour tout $j = 1,\dots,k$ et donc $\gamma'(0)$ est bien dans $T_a\mathcal{M}$ (tel que défini dans la définition 6.5).

Montrons réciproquement que tout vecteur de $T_a\mathcal{M}$ (tel que défini dans la définition 6.5) peut être obtenu comme la dérivée en $t = 0$ d'un courbe paramétrée. Via un changement de base, on peut toujours supposer que l'espace $\mathbb{R}^m$ peut se décomposer en un produit cartésien d'espaces $\mathbb{R}^{m-k}\times\mathbb{R}^k$ de telle sorte que l'application :

$$\begin{array}{rcl} D_2F(a) : \mathbb{R}^k & \longrightarrow & \mathbb{R}^k \\ h_2 & \longmapsto & DF(a)(0,h_2), \end{array}$$

soit un isomorphisme. On écrit le point a sous la forme $a = (a_1,a_2) \in \mathbb{R}^{m-k}\times\mathbb{R}^k$. Le théorème des fonctions implicites nous assure de l'existence d'un ouvert $\mathscr{U}$ de $\mathbb{R}^{m-k}$ contenant a_1, d'un ouvert $\mathscr{V}$ de $\mathbb{R}^k$ contenant a_2 et d'un difféormorphisme $\psi : \mathscr{U} \to \mathscr{V}$ de classe $\mathscr{C}^1$ tels que $\psi(a_1) = a_2$ et :

$$\big\{x \in \mathscr{U}\times\mathscr{V} \,:\, F(x) = 0\big\} = \big\{(x_1,\psi(x_1)) \,:\, x_1 \in \mathscr{U}\big\}.$$

L'application

$$\begin{array}{rcl} \phi : \mathscr{U} & \longrightarrow & \mathbb{R}^m \\ x_1 & \longmapsto & (x_1,\psi(x_1)), \end{array}$$

fournit une paramétrisation locale (au voisinage a) de $\mathcal{M}$. Elle est de classe $\mathscr{C}^1$ (comme ψ) et en tout point x_1 de $\mathscr{U}$:

$$D\phi(x_1) = (\mathrm{Id}, D\psi(x_1)) : \mathbb{R}^{m-k} \longrightarrow T_a\mathcal{M},$$

est au moins de rang $m-k$ (donc en fait exactement de rang $m-k$) et est donc un isomorphisme. Soit u un vecteur de $T_a\mathcal{M}$ et soit $v = [D\phi(a)]^{-1}u$. Pour tout t dans un voisinage de 0, on peut alors définir :

$$\gamma(t) = (a_1 + tv, \psi(a_1 + tv)),$$

et on vérifie aisément que $\gamma(0) = a$ et $\gamma'(0) = u$. □

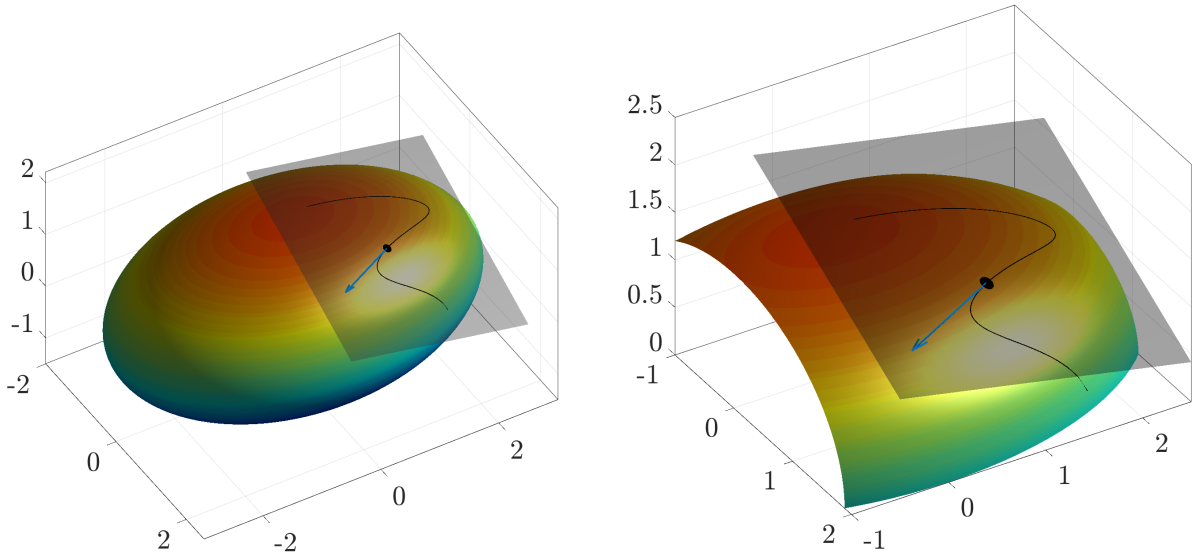


FIG. 6.2 – . *L'ellipsoïde $\mathcal{M}$ est défini par l'équation $F(x,y,z) = 0$ avec $F(x,y,z) = x^2/4 + y^2/9 + 4z^2/9 - 1$. On considère le point $a = (1,1,\sqrt{23}/4)$ (marqué en noir sur le dessin). Le plan tangent en a est l'espace affine $a + T_a\mathcal{M}$. L'espace tangent $T_a\mathcal{M}$ est constitué par tous les vecteurs $\gamma'(0)$ (en bleu sur le dessin) où $\gamma :]-1,1[\to \mathbb{R}^3$ est une courbe paramétrée tracée sur $\mathcal{M}$ telle que $\gamma(0) = a$. L'ellipsoïde est une sous–variété différentiable de classe $\mathscr{C}^\infty$ de dim 2 dans $\mathbb{R}^3$.*

Définition 6.6. *Un ensemble $\mathcal{M}$ défini par* (6.1) *dont tous les points sont réguliers est appelé une sous–variété de $\mathbb{R}^m$ de dimension $m-k$ de classe $\mathscr{C}^1$.*

Les sous–variétés (et plus généralement les variétés différentiables) seront étudiées l'année prochaine dans le cours "Géométrie différentielle".

6.4.2 Points critiques de H sur $\mathcal{M}$

Soit a un extremum local de H sur $\mathcal{M}$. On suppose que a est un point régulier de $\mathcal{M}$. Alors, pour toute courbe paramétrée $\gamma :]-1,1[\to \mathcal{M}$ de classe $\mathscr{C}^1$ telle que $\gamma(0) = a$ on a $g'_\gamma(0) = 0$ avec $g_\gamma(t) = H(\gamma(t))$. On doit donc avoir $\nabla H(\gamma(t))\gamma'(t) = 0$ pour toute courbe γ et cette condition se traduit par

$$\nabla H(a)u = 0 \quad \forall\, u \in T_a\mathcal{M}.$$

Ce calcul nous suggère de définir les points critiques comme suit :

Définition 6.7. *Soit a un point régulier de $\mathcal{M}$. Alors a est un point critique de H sur $\mathcal{M}$ si :*

$$T_a\mathcal{M} \subset \ker DH(a). \tag{6.3}$$

Tout extremum a de H sur $\mathcal{M}$ (qui est un point régulier de $\mathcal{M}$) est donc un point critique. Comme :

$$\ker DH(a) = \langle \nabla H(a) \rangle^{\perp},$$

l'inclusion (6.3) se réécrit :

$$\langle \nabla F_1(a), \ldots, \nabla F_k(a) \rangle^{\perp} \subset \langle \nabla H(a) \rangle^{\perp},$$

ce qui est équivalent à :

$$\nabla H(a) \in \langle \nabla F_1(a), \ldots, \nabla F_k(a) \rangle.$$

Proposition 6.6 (Multiplicateurs de Lagrange)**.** *Soit a un point régulier de $\mathcal{M}$. Alors a est un point critique de H sur $\mathcal{M}$ si et seulement si il existe des constantes $\lambda_1, \ldots, \lambda_k$ telles que:*

$$\nabla H(a) = \sum_{j=1}^{k} \lambda_j \nabla F_j(a).$$

Exemple 6.2. On chercher à déterminer le triangle de plus grande aire dont le cercle circonscrit est le cercle unité.

Par symétrie, on peut supposer que l'un des sommets est le point (1,0). Notons (x_1,y_1) et (x_2,y_2) les coordonnées des deux autres sommets. Ces sommets étant astreints à rester sur le cercle unité, on pose :

$$\mathcal{M} = \big\{(x_1,x_2,y_1,y_2) \in \mathbb{R}^4 \ : F(x_1,x_2,y_1,y_2) = 0\big\},$$

où la fonction F est définie par :

$$\begin{array}{rcl} F : \mathbb{R}^4 & \longrightarrow & \mathbb{R}^2 \\ (x_1,x_2,y_1,y_2) & \longmapsto & \begin{pmatrix} x_1^2 + y_1^2 - 1 \\ x_2^2 + y_2^2 - 1 \end{pmatrix} \end{array}.$$

Cette fonction est très régulière ($\mathscr{C}^\infty$ en fait). On note ensuite $F_1(x_1,x_2,y_1,y_2) = x_1^2 + y_1^2 - 1$ et $F_2(x_1,x_2,y_1,y_2) = x_2^2 + y_2^2 - 1$. Pour tout $(x_1,x_2,y_1,y_2) \in \mathbb{R}^4$, un simple calcul conduit à :

$$\nabla F_1(x_1,x_2,y_1,y_2) = (2x_1,0,2y_1,0) \qquad \text{et} \qquad \nabla F_2(x_1,x_2,y_1,y_2) = (0,2x_2,0,2y_2).$$

Sur $\mathcal{M}$, aucun de ces deux vecteurs ne peut s'annuler et ils forment en tout point une famille libre. Tous les points de $\mathcal{M}$ sont donc réguliers (et $\mathcal{M}$ est donc une sous–variété $\mathscr{C}^\infty$ de $\mathbb{R}^4$ de dimension 2). L'aire du triangle de sommets (1,0), (x_1,y_1) et (x_2,y_2) est égale à :

$$\frac{1}{2}|(x_1 - 1)y_2 - (x_2 - 1)y_1| = \frac{1}{2}|H(x_1,x_2,y_1,y_2)|, \tag{6.4}$$

la fonction H étant définie sur $\mathbb{R}^4$ par :

$$\begin{array}{rcl} H : \mathbb{R}^4 & \longrightarrow & \mathbb{R} \\ (x_1,x_2,y_1,y_2) & \longmapsto & (x_1-1)y_2-(x_2-1)y_1. \end{array}$$

L'application H est continue sur le compact $\mathcal{M}$ donc elle est bornée et atteint ses bornes. L'application H admet donc au moins un min global sur $\mathcal{M}$ et un max global sur $\mathcal{M}$. Le gradient de H est:

$$\nabla H(x_1,x_2,y_1,y_2) = (y_2, -y_1, 1-x_2, x_1-1).$$

Un point (x_1,x_2,y_1,y_2) est donc un point critique de H sur $\mathcal{M}$ ssi il existe deux réels λ_1 et λ_2 (des multiplicateurs de Lagrange) tels que :

$$\nabla H(x_1,x_2,y_1,y_2) = \lambda_1 \nabla F_1(x_1,x_2,y_1,y_2) + \lambda_2 \nabla F_2(x_1,x_2,y_1,y_2),$$

c'est à dire tels que :

$$\begin{array}{ll} y_2 = 2\lambda_1 x_1 & \qquad x_2 = -2\lambda_1 y_1 + 1 \\ y_1 = -2\lambda_2 x_2 & \qquad x_1 = 2\lambda_2 y_2 + 1. \end{array}$$

On en déduit, avec les deux équations de la première ligne, que $x_2^2+y_2^2 = 4\lambda_1^2(x_1^2+y_1^2)+1-4\lambda_1 y_1$. Les deux équations de la deuxième ligne conduisent à $x_1^2+y_1^2 = 4\lambda_2^2(x_2^2+y_2^2)+1+4\lambda_2 y_2$. Comme $x_1^2+y_1^2 = 1$ et $x_2^2+y_2^2 = 1$, on trouve que :

$$\lambda_2(\lambda_2+y_2) = 0 \quad \text{et} \quad \lambda_1(\lambda_1-y_1) = 0.$$

Considérons plusieurs cas :

- Si $(\lambda_1,\lambda_2) = (0,0)$ alors $(x_1,y_1) = (x_2,y_2) = (1,0)$: les trois sommets du triangle sont confondus.
- Si $\lambda_2 = 0$ et $\lambda_1 \neq 0$. Alors $y_1 = 0$, $x_1 = 1$ puis $x_2 = 1$ et $y_2 = 0$: De nouveau les trois sommets du triangle sont confondus. De même si $\lambda_1 = 0$ et $\lambda_2 \neq 0$.
- Supposons donc que $\lambda_1 \neq 0$ et $\lambda_2 \neq 0$. Dans ce cas $\lambda_1 = y_1$ et $\lambda_2 = -y_2$. Les deux équations de gauche du système nous donnent :

  $$y_2(4x_1x_2-1) = 0,$$

  c'est à dire (puisque $y_2 = -\lambda_2 \neq 0$) $4x_1x_2 = 1$. Les deux équations de droite du système conduisent à :

  $$x_1 = 1-2y_2^2 = -1+2x_2^2 \quad \text{et} \quad x_2 = 1-2y_1^2 = -1+2x_1^2. \tag{6.5}$$

  En multipliant termes à termes on obtient:

  $$x_1x_2 = (2x_1^2-1)(2x_2^2-1) = 4x_1^2x_2^2-2x_1^2-2x_2^2+1,$$

  c'est à dire :

  $$x_1^2+x_2^2 = \frac{1}{2}.$$

On obtient finalement que $(x_1 - x_2)^2 = x_1^2 + x_2^2 - 2x_1x_2 = 0$ et donc $x_1 = x_2$, ce qui conduit à $x_1^2 = x_2^2 = 1/4$ puis avec (6.5) à $x_1 = x_2 = -1/2$. On retourne à la première équation du système d'où l'on tire (en rappelant que $\lambda_1 = y_1$ et $\lambda_2 = -y_2$) que $y_2 = -y_1$. On trouve finalement, de nouveau avec les équations (6.5) que les points critiques sont $(-1/2,\sqrt{3}/2)$ et $(-1/2, -\sqrt{3}/2)$.

On peut maintenant résumer les résultats dans un tableau :

Points critiques (x_1,x_2,y_1,y_2)	λ_1,λ_2	$H(x_1,x_2,y_1,y_2)$
$(1,0,1,0)$	$\lambda_1 = 0$ ou $\lambda_2 = 0$	0
$(-1/2, -1/2,\sqrt{3}/2, -\sqrt{3}/2)$	$\lambda_1 = \lambda_2 = \sqrt{3}/2$	$3\sqrt{3}/2$
$(-1/2, -1/2, -\sqrt{3}/2,\sqrt{3}/2)$	$\lambda_1 = \lambda_2 = \sqrt{3}/2$	$-3\sqrt{3}/2$

Avec la formule (6.4), on en déduit que l'aire maximale pour un triangle dont le cercle circonscrit est le cercle unité est donc $3\sqrt{3}/4$. Le triangle ayant cette aire maximale est le triangle équilatéral.

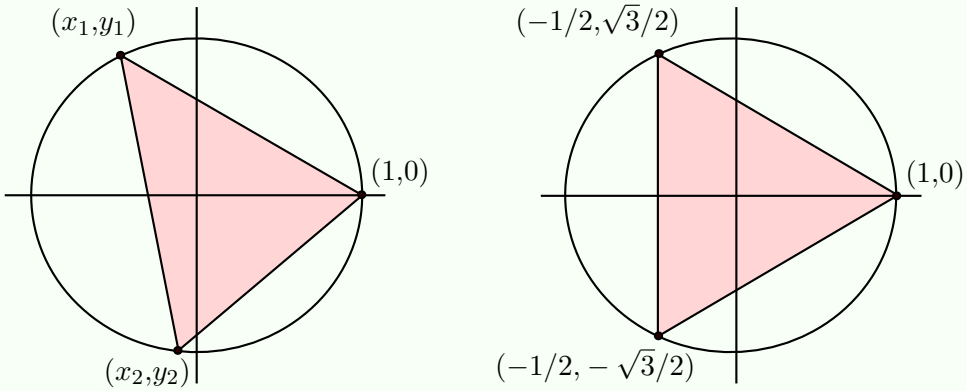


FIG. 6.3 – *Le triangle ayant une aire maximale est le triangle équilatéral.*

Chapitre 7

Équations différentielles

7.1 Définitions et notations

Soit $\mathscr{U}$ un ouvert de $\mathbb{R}\times\mathbb{R}^N$ et $f : \mathscr{U} \to \mathbb{R}^N$ une fonction continue. Dans ce cours, on s'intéresse aux équations différentielles ordinaires (notées en abrégé edo) du premier ordre, sous forme normale (ou résolue) :

$$x'(t) = f(t,x(t)). \tag{7.1}$$

Commençons par préciser la notion de solution pour ce type d'équation :

Définition 7.1. *Une solution de* (7.1) *est un couple* (φ,J) *où* J *est un intervalle de* $\mathbb{R}$ *et* $\varphi = (\varphi_1,\dots,\varphi_N)$ *est une fonction dérivable sur* J *à valeurs dans* $\mathbb{R}^N$ *telle que* $(t,\varphi(t)) \in \mathscr{U}$ *pour tout* $t \in J$ *et*

$$\varphi_i'(t) = f_i(t,\varphi(t)), \qquad \forall\, t \in J, \qquad \forall\, i = 1,\dots,N.$$

On remarque tout de suite que, f et φ étant deux fonctions continues, par composition $\varphi' = (\varphi_1',\dots,\varphi_N')$ est également continue sur J et φ est de classe $\mathscr{C}^1$ sur J.

Exemple 7.1. Dans le cas $N = 2$, notons $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ un vecteur de $\mathbb{R}^2$ et définissons

$$f(t,x) = M(t)x = \begin{pmatrix} a(t) & b(t) \\ c(t) & d(t) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

où $t \mapsto a(t), t \mapsto b(t), t \mapsto c(t)$ et $t \mapsto d(t)$ sont des fonctions réelles continues. Alors l'équation

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

est appelée équation linéaire du premier ordre.

Pour $N = 1$, lorsque $f(t,x) = a(t)x + b(t)x^\alpha$ où $t \mapsto a(t)$ et $t \mapsto b(t)$ sont des fonctions continues et $\alpha \in \mathbb{R} \setminus \{0,1\}$, l'edo (7.1) est appelée équation de Bernoulli.

Toujours lorsque $N = 1$, l'équation

$$x'(t) = a(t)x^2(t) + b(t)x(t) + c(t),$$

pour laquelle $f(t,x) = a(t)x^2 + b(t)x + c(t)$ où $t \mapsto a(t)$, $t \mapsto b(t)$ et $t \mapsto c(t)$ sont trois fonctions continues, est une équation de Riccati.

On introduit la notion de problème de Cauchy :

Définition 7.2 (Problème de Cauchy)**.** *Soit $(t_0,x_0) \in \mathscr{U}$. Résoudre le problème de Cauchy :*

$$\begin{cases} x'(t) = f(t,x) \\ x(t_0) = x_0, \end{cases} \tag{7.2}$$

consiste à déterminer un couple (φ,J) où J est un intervalle de $\mathbb{R}$ contenant t_0 et φ une fonction dérivable (en fait $\mathscr{C}^1$) de J dans E telle que $(t,\varphi(t)) \in \mathscr{U}$ pour tout $t \in J$, $\varphi'(t) = f(t,\varphi(t))$ pour tout $t \in J$ et $\varphi(t_0) = x_0$.

En intégrant l'edo du problème de Cauchy (7.2) entre t_0 et t et en tenant compte de la condition $x(t_0) = x_0$, on obtient que :

$$\varphi(t) = x_0 + \int_{t_0}^{t} f(s,\varphi(s))\,\mathrm{d}s, \tag{7.3}$$

où il faut comprendre l'intégrale comme :

$$\int_{t_0}^{t} f(s,\varphi(s))\,\mathrm{d}s = \begin{pmatrix} \int_{t_0}^{t} f_1(s,\varphi(s))\,\mathrm{d}s \\ \vdots \\ \int_{t_0}^{t} f_N(s,\varphi(s))\,\mathrm{d}s \end{pmatrix}.$$

Réciproquement, toute fonction φ vérifiant (7.3) est bien une solution $\mathscr{C}^1$ de (7.2). Nous utiliserons souvent l'équivalence entre les deux formulations (7.2) et (7.3) dans la suite du cours.

7.2 Réduction de l'ordre

Les formulations (7.1) et (7.2), bien que ne faisant intervenir que la dérivée première de $t \mapsto x(t)$, recouvrent en fait une large classe de problèmes. En effet, il est souvent possible de mettre sous la forme (7.1) des edo dans lesquelles apparaissent des dérivées à un ordre quelconque. Considérons pour simplifier que les fonctions $t \mapsto x(t)$ sont à valeurs dans $\mathbb{R}$ (i.e. $N = 1$). Soit $\mathscr{U}$ un ouvert de $\mathbb{R} \times \mathbb{R}^p$ avec $p \geqslant 1$ et $f : \mathscr{U} \to \mathbb{R}$ une fonction continue. En notant $t \mapsto x^{(k)}(t)$ la dérivée $k-$ème de $t \mapsto x(t)$, toute équation différentielle ordinaire du $p-$ème ordre associée à f qui s'écrit :

$$x^{(p)}(t) = f(t,x(t),x'(t),\ldots,x^{(p-1)}(t)), \tag{7.4}$$

peut se mettre sous la forme (7.1). Notons en effet $x_1(t) = x(t)$ et $x_{i+1}(t) = x^{(i)}(t)$ pour $i = 1,\ldots,p-1$ et introduisons

$$X(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_p(t) \end{pmatrix} \in \mathbb{R}^p \qquad \text{et} \qquad F(t,X) = \begin{pmatrix} x_2(t) \\ \vdots \\ x_p(t) \\ f(t,x_1(t),\ldots,x_p(t)) \end{pmatrix}.$$

L'edo (7.4) devient alors :

$$X'(t) = F(t,X(t)).$$

Le problème de Cauchy correspondant à l'équation ci-dessus s'obtient en ajoutant une condition du type $X(t_0) = X_0$ où $t_0 \in \mathbb{R}$ et $X_0 \in \mathbb{R}^p$. Ceci correspond à la donnée de $x(t_0),x'(t_0),\dots,x^{p-1}(t_0)$.

7.3 Équations différentielles autonomes

Une équation différentielle est dite autonome lorsque la fonction f qui apparait dans l'égalité (7.1) ne dépend pas de la variable t. Une équation différentielle est ainsi de la forme :

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

où f est une fonction d'un ouvert $\mathscr{U}$ de $\mathbb{R}^N$ dans $\mathbb{R}^N$. C'est un cas particulier important car toute équation différentielle (sous forme normale) peut être mise sous la forme d'une équation différentielle autonome. En effet, considérons un problème de Cauchy sous la forme la plus générale :

$$\begin{cases} x'(t) = f(t,x(t)) \\ x(t_0) = x_0, \end{cases}$$

où $f : \mathscr{U} \to \mathbb{R}^N$, $\mathscr{U} \subset \mathbb{R} \times \mathbb{R}^N$ et $(t_0,x_0) \in \mathscr{U}$. Ce problème est équivalent au problème suivant, dont l'edo est autonome :

$$\begin{cases} X'(t) = F(X(t)) \\ X(t_0) = X_0. \end{cases}$$

où la fonction $F : \mathscr{U} \to \mathbb{R} \times \mathbb{R}^N$ est définie pour tout $X = (\tau,x) \in \mathscr{U}$ par $F(X) = (1,f(\tau,x))$ et où $X_0 = (t_0,x_0)$.

7.4 Existence et unicité locale de solutions

Pour montrer l'existence et l'unicité de solutions locales en temps (c'est à dire sur un intervalle de temps petit contenant t_0) il faut supposer vérifiée l'hypothèse suivante :

Hypothèse 7.1. *La fonction f est continue sur $\mathscr{U}$ et $x \mapsto f(t,x)$ est lipschitzienne en x (uniformément en t), c'est à dire qu'il existe $k \geqslant 0$ telle que :*

$$\|f(t,x) - f(t,\tilde{x})\| \leqslant k\|x - \tilde{x}\| \qquad \textit{pour tout } (t,x),(t,\tilde{x}) \in \mathscr{U}. \tag{7.5}$$

Théorème 7.1 (de Cauchy–Lipschitz)**.** *Soit f une fonction vérifiant l'hypothèse 7.1 et soit (t_0,x_0) un point de $\mathscr{U}$. Alors pour tous $a > 0$ et $T > 0$ tels que :*

1. *Le cylindre $\mathscr{C} = \big\{(t,x) \in \mathbb{R} \times \mathbb{R}^N \,:\, \|x - x_0\| \leqslant a, |t - t_0| \leqslant T\big\}$ soit inclus dans $\mathscr{U}$;*
2. *$TM \leqslant a$ et $Tk < 1$ où $M = \max_{(t,x)\in\mathscr{C}} \|f(t,x)\|$;*

il existe une unique solution au problème de Cauchy (7.2) sur l'intervalle $[t_0 - T,t_0 + T]$.

Remarquez que lorsque le premier point est vérifié, il est toujours possible de réduire T suffisamment pour que le deuxième soit vérifié également. La démonstration repose sur le Lemme 4.1 (Point fixe de Banach).

Démonstration. Soit a, T et $\mathscr{C}$ comme dans les hypothèses. Notons

$$F = \big\{\varphi : [t_0 - T,t_0 + T] \to E \text{ continue}\big\}$$

une espace vectoriel que l'on munit de la norme de la convergence uniforme :

$$\|\varphi\|_\infty = \max_{t\in[t_0-T,t_0+T]} \|\varphi(t)\| \qquad \text{pour tout } \varphi \in F.$$

Dans l'espace F, on considère l'ensemble :

$$\Omega = \big\{\varphi : [t_0 - T,t_0 + T] \to \overline{B(x_0,a)}\big\}.$$

Alors F est un espace de Banach et Ω est fermé. Notons, pour tout $\varphi \in \Omega$:

$$\phi(\varphi)(t) = x_0 + \int_{t_0}^{t} f(s,\varphi(s))\,\mathrm{d}s \qquad \text{pour tout } t \in [t_0 - T,t_0 + T],$$

et vérifions que $\phi(\varphi) \in \Omega$. Comme déjà observé plus haut, $\phi(\varphi)$ est de classe $\mathscr{C}^1$ sur $[t_0 - T,t_0 + T]$. D'autre part, pour tout $t \in [t_0,t_0 + T]$:

$$\|\phi(\varphi(t)) - x_0\| \leqslant \int_{t_0}^{t} \|f(s,\varphi(s))\|\,\mathrm{d}s \leqslant TM \leqslant a.$$

On obtient la même inégalité pour $t \in [t_0 - T,t_0]$ et donc Ω est stable par ϕ. Vérifions maintenant que ϕ est contractante sur Ω. Pour tout $t \in [t_0,t_0 + T]$ et tous $\varphi,\tilde{\varphi} \in \Omega$:

$$\begin{aligned}\|\phi(\varphi)(t) - \phi(\tilde{\varphi})(t)\| &\leqslant \int_{t_0}^{t} \|f(\varphi(s),s) - f(\tilde{\varphi}(s),s)\|\,\mathrm{d}s \\ &\leqslant k\int_{t_0}^{t} \|\varphi(s) - \tilde{\varphi}(s)\|\,\mathrm{d}s \leqslant kT\|\varphi - \tilde{\varphi}\|_\infty.\end{aligned}$$

On obtient la même inégalité pour tout $t \in [t_0 - T,t_0]$ et on en déduit que :

$$\|\phi(\varphi) - \phi(\tilde{\varphi})\|_\infty \leqslant kT\|\varphi - \tilde{\varphi}\|_\infty,$$

ce qui prouve le caractère contractant de ϕ. Le théorème du point fixe de Banach (Lemme 4.1) nous assure donc de l'existence d'une unique fonction φ dans Ω telle que, pour tout $t \in [t_0 - T,t_0 + T]$:

$$\varphi(t) = x_0 + \int_{t_0}^{t} f(s,\varphi(s))\,\mathrm{d}s,$$

ce qui est la formulation intégrale du problème de Cauchy. Nous avons ainsi montré l'existence et l'unicité d'une solution au problème de Cauchy dans Ω. Vérifions qu'il n'existe pas de solution en dehors de Ω.

Considérons pour cela φ une solution quelconque du problème de Cauchy définie sur $[t_0-T,t_0+T]$ et notons:

$$T^* = \sup\big\{\theta \in [0,T] \,:\, \|\varphi(t) - x_0\| \leqslant a, \quad \forall\, t \in [t_0-\theta,t_0+\theta]\big\}.$$

L'ensemble est non vide (car il contient 0), majoré par T (donc T^* est bien défini) et on vérifie facilement qu'il contient T^*. Supposons $T^* < T$. On a alors :

$$\|\varphi(t_0+T^*) - x_0\| = \left\| \int_{t_0}^{t_0+T^*} f(s,\varphi(s))\,\mathrm{d}s \right\| \leqslant T^*M < TM \leqslant a,$$

et donc par continuité de φ, on a encore $\|\varphi(t)-x_0\| \leqslant a$ pour $t > t_0+T^*$ (et aussi pour $t < t_0-T^*$ pour les mêmes raisons) ce qui contredit la définition de T^*. Toute solution du problème de Cauchy définie sur $[t_0-T,t_0+T]$ est donc dans Ω. □

Lorsque la fonction f est seulement continue, l'existence d'une solution locale en temps au problème de Cauchy est encore assurée mais plus l'unicité. Signalons le résultat suivant (dont nous admettrons la démonstration) :

Théorème 7.2 (Ascoli–Peano)**.** *Soit $f \in \mathscr{C}(\mathscr{U})$. Alors pour tout $(t_0,x_0) \in \mathscr{U}$, il existe un voisinage de t_0 dans $\mathbb{R}$ sur lequel le problème de Cauchy* (7.2) *admet une solution (non nécessairement unique).*

Illustrons cette perte d'unicité par l'exemple suivant :

Exemple 7.2. Le problème de Cauchy

$$x'(t) = \sqrt{|x(t)|}, \qquad x(0) = 0,$$

où $f : (t,x) \mapsto \sqrt{|x|}$ est définie et continue sur $\mathscr{U} = \mathbb{R}\times\mathbb{R}$ et $(0,x(0)) \in \mathscr{U}$. La fonction f est continue mais ne satisfait pas l'hypothèse 7.1 au voisinage de $x = 0$. La solution identiquement nulle est solution du problème de Cauchy. On vérifie que pour tous réels α et β avec $\alpha \leqslant 0 \leqslant \beta$ la fonction φ définie par :

$$\varphi(t) = \begin{cases} \frac{1}{4}(t-\beta)^2 & \text{si } t > \beta \\ 0 & \text{si } t \in [\alpha,\beta] \\ -\frac{1}{4}(t-\alpha)^2 & \text{si } t < \alpha, \end{cases}$$

(cf. Fig. 7.1) est aussi solution du problème de Cauchy sur $\mathbb{R}$ (Le théorème d'Ascoli–Peano s'applique mais pas le théorème de Cauchy–Lipschitz). Remarquer que toutes les solutions sont bien $\mathscr{C}^1$ sur $\mathbb{R}$.

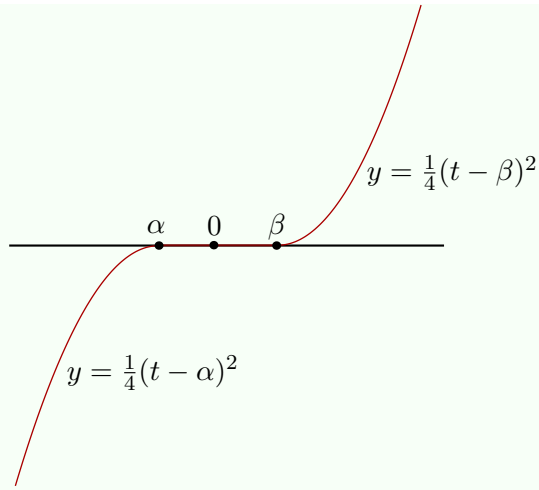


FIG. 7.1 – *Non unicité de la solution du problème de Cauchy.*

On utilisera le plus souvent le corollaire suivant dont les hypothèses sont plus simples à vérifier que celles du Théorème 7.1 :

Corollaire 7.1. *Soit f une fonction $\mathscr{C}^1$ sur $\mathscr{U}$. Alors pour tout $(t_0,x_0) \in \mathscr{U}$, il existe $T > 0$ tel que le problème de Cauchy (7.2) admette une unique solution sur $[t_0 - T,t_0 + T]$.*

Démonstration. Soit $\mathscr{B}$ une boule ouvert dans $\mathbb{R} \times \mathbb{R}^N$, centrée en (t_0,x_0) et telle que $\overline{\mathscr{B}}$ soit incluse dans $\mathscr{U}$. Par continuité de Df sur le compact $\overline{\mathscr{B}}$, on peut définir :

$$k = \max_{(t,x)\in\overline{\mathscr{B}}} \|Df(t,x)\|_{\mathscr{L}(\mathbb{R}\times\mathbb{R}^N,\mathbb{R}^N)}.$$

On applique ensuite le Corollaire 3.2 sur le convexe $\mathscr{B}$ pour obtenir que :

$$\|f(t,x) - f(\tilde{t},\tilde{x})\| \leqslant k\left(\|x - \tilde{x}\| + |t - \tilde{t}|\right) \qquad \text{pour tout } (t,x),(\tilde{t},\tilde{x}) \in \mathscr{B}.$$

En particulier, l'hypothèse 7.1 est vérifiée sur $\mathscr{B}$ et on peut appliquer le théorème de Cauchy–Lipschitz sur cet ouvert, ce qui nous donne le résultat. □

7.5 Unicité globale, solutions maximales

Le résultat suivant que deux solutions d'une même équation différentielle ne peuvent pas se croiser en un point où la fonction f est régulière.

Théorème 7.3 (Unicité globale)**.** *Soit $f \in \mathscr{C}^1(\mathscr{U})$ et soient (φ_1,J_1) et (φ_2,J_2) deux solutions de (7.1). S'il existe un point t_0 de $J_1 \cap J_2$ tel que $\varphi_1(t_0) = \varphi_2(t_0)$ alors $\varphi_1 = \varphi_2$ sur $J_1 \cap J_2$.*

Démonstration. Les ensembles J_1 et J_2 sont des intervalles. Il en est donc de même de $J = J_1 \cap J_2$ qui est en particulier connexe et non vide puisqu'il contient t_0. On note

$$I = \left\{t \in J \text{ tels que } \varphi_1(t) = \varphi_2(t)\right\}.$$

Montrons que I est ouvert et fermé dans J ce qui entraînera que $I = J$. Les fonctions φ_1 et φ_2 étant continues, on en déduit que $I = (\varphi_1 - \varphi_2)^{-1}(\{0\})$ est fermé. Soit $t_1 \in I$, notons $x_1 = \varphi_1(t_1) = \varphi_2(t_1)$. Alors $(t_1,x_1) \in \mathscr{U}$ et selon le Corollaire 7.1, le problème de Cauchy:

$$x'(t) = f(t,x(t)), \qquad x(t_1) = x_1,$$

admet une unique solution sur un petit intervalle $[t_1 - \alpha, t_1 + \alpha]$. On en déduit que $\varphi_1 = \varphi_2$ sur cet intervalle et que $]t_1 - \alpha, t_1 + \alpha[\subset I$ et donc que I est ouvert. □

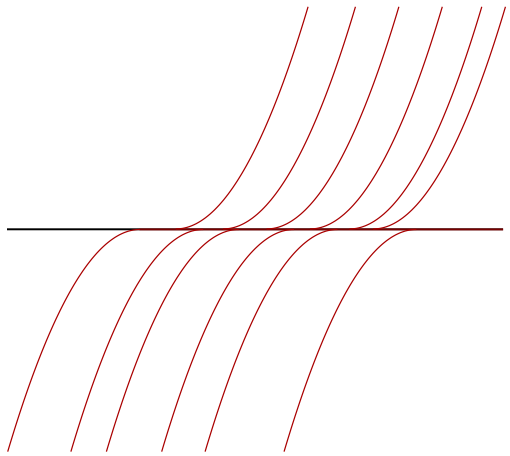


FIG. 7.2 – *Les solutions de l'edo $x'(t) = \sqrt{|x(t)|}$ de l'exemple 7.2 se croisent seulement sur l'axe des abscisses où f n'est plus différentiable.*

Définition 7.3 (Solution maximale)**.** *Soit (φ_1,J_1) et (φ_2,J_2) deux solutions de (7.1). On dit que (φ_2,J_2) prolonge (φ_1,J_1) si $J_1 \subset J_2$ et $\varphi_1 = \varphi_2$ sur J_1.*
Une solution (φ,J) de (7.1) est dite maximale si elle n'admet aucun prolongement.

Théorème 7.4 (Existence d'une solution maximale)**.** *Soit $f \in \mathscr{C}^1(\mathscr{U})$. Alors par tout point $(t_0,x_0) \in \mathscr{U}$ il passe une unique solution maximale au problème de Cauchy (7.2).*

Démonstration. Considérons l'ensemble $\mathcal{S}$ de tous les couples (φ,J) de solutions au problème de Cauchy (7.2). Si (φ_1,J_1) et (φ_2,J_2) sont deux tels couples alors $J_1 \cap J_2$ n'est pas vide car il contient t_0 et $\varphi_1 = \varphi_2$ sur $J_1 \cap J_2$ d'après le Théorème 7.3. Soit I la réunion de tous les intervalles J. Sur I on peut donc définir la fonction ψ par $\psi = \varphi$ sur J pour tout $(\varphi,J) \in \mathcal{S}$. Cette fonction est la solution maximale cherchée. □

La question à laquelle nous allons nous intéresser maintenant est celle du prolongement des solutions maximales. En effet, pourquoi une solution maximale définie sur un intervalle borné, ne peut–elle être prolongée sur un intervalle plus grand ?

Théorème 7.5. *On suppose que Ω est un ouvert de $\mathbb{R}^N$ et que $\mathscr{U} =]a,b[\times\Omega$ (avec $-\infty \leqslant a < b \leqslant +\infty$). Soient $f \in \mathscr{C}^1(\mathscr{U})$ et $(t_0,x_0) \in \mathscr{U}$. Si $(\varphi,]T_-,T_+[)$ est une solution maximale du problème de Cauchy (7.2), alors on a l'alternative suivante :*

- *Soit $T_+ = b$;*
- *Soit $T_+ < b$ et pour tout compact K de Ω il existe $t < T_+$ tel que $\varphi(t) \notin K$.*

On peut énoncer un résultat analogue pour T_-.

Démonstration. Supposons que $T_+ < b$ et qu'il existe un compact K dans Ω tel que $\varphi(t) \in K$ pour tout $t \in]t_0,T_+[$. Alors, comme $f \in \mathscr{C}(\mathscr{U})$, il existe $M > 0$ tel que $\|f(t,x)\| \leqslant M$ pour tout $(t,x) \in [t_0,T_+] \times K$. Soit $(t_n)_n$ une suite croissante tendant vers T_+ et telle que $t_0 < t_n < T_+$ pour tout n. En écrivant la solution φ sous forme intégrale, on obtient que :

$$\|\varphi(t_m) - \varphi(t_n)\| \leqslant \int_{t_n}^{t_m} \|f(s,\varphi(s))\| \,\mathrm{d}s \leqslant M|t_m - t_n|, \qquad \forall\, m > n.$$

La suite $(t_n)_n$ étant de Cauchy, il en est de même pour $(\varphi(t_n))_n$ qui est donc convergente. Notons $x_1 = \lim_{n\to\infty} \varphi(t_n)$. Alors $x_1 \in K \subset \Omega$ et on a donc $(T_+,x_1) \in \mathscr{U}$. La solution du problème de Cauchy

$$x'(t) = f(t,x(t)), \qquad x(T_+) = x_1,$$

admet selon le Corollaire 7.1 une solution locale qui, selon le Théorème 7.3, prolonge φ au delà de T_+. Ceci contredit la maximalité de $(\varphi,]T_-,T_+[)$. On procède de façon analogue pour T_-. □

Si Ω est borné, ce théorème se traduit par : Le point de $\mathbb{R} \times E$ de coordonées $(t,\varphi(t))$ tend vers un point de la frontière du cylindre $]a,b[\times\Omega$ quand $t \to T_+$ et $t \to T_-$.

Exemple 7.3. On souhaite étudier l'edo :

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

où la fonction $f : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ est définie pour tout $(t,x) \in \mathbb{R}\times\mathbb{R}$ par $f(t,x) = 2tx^2$. La fonction f est $\mathscr{C}^1$ (en fait $\mathscr{C}^\infty$) sur $\mathbb{R}\times\mathbb{R}$. Pour toute donnée de Cauchy $(t_0,x_0) \in \mathbb{R}\times\mathbb{R}$ il existe donc, selon le théorème 7.4, une unique solution maximale au problème de Cauchy.

$$\begin{cases} x'(t) = f(t,x) \\ x(t_0) = x_0. \end{cases}$$

D'après le théorème 7.5, chaque solution est soit définie sur $\mathbb{R}$ tout entier, soit sur un intervalle $]a,b[$ (avec possiblement $a = -\infty$ ou $b = +\infty$). Si $a \neq -\infty$ (respectivement $b \neq +\infty$), alors

$$\lim_{\substack{t\to a\\ t>a}} |x(t)| = +\infty \qquad \text{repectivement} \qquad \lim_{\substack{t\to b\\ t<b}} |x(t)| = +\infty. \tag{7.6}$$

Cherchons explicitement les solutions du problème de Cauchy. On remarque que la solution identiquement nulle (définie sur $\mathbb{R}$) est la solution cherchée pour t_0 quelconque et $x_0 = 0$ (on sait que cette solution est unique). On suppose que $x_0 \neq 0$. Alors on voisinage de t_0, la solution ne s'annule pas (par continuité) et satisfait :

$$\frac{x'(t)}{x^2(t)} = 2t.$$

On peut ensuite intégrer cette égalité entre t et t_0 pour obtenir :

$$\frac{1}{x_0} - \frac{1}{x(t)} = t^2 - t_0^2,$$

ce qui nous conduit à l'expression :

$$x(t) = \frac{1}{c_0 - t^2} \qquad \text{avec } c_0 = \frac{1}{x_0} + t_0^2.$$

- Si $c_0 = 0$ alors $x(t) = -1/t^2$ est la solution cherchée. Plus précisément, si $t_0 < 0$ la solution est $(\varphi, J) = \big(-1/t^2,]-\infty, 0[\big)$ et si $t_0 > 0$, la solution est $\big(-1/t^2,]0, +\infty[\big)$. Remarquer que les conditions (7.6) sont satisfaites.
- Si $c_0 > 0$ et $x_0 > 0$ (cette zone est la zone A sur la figure 7.3). Les solutions sont $\big(-1/(c_0 - t^2),]-\sqrt{c_0}, \sqrt{c_0}[\big)$ et les conditions (7.6) sont bien entendu satisfaites aux bornes de l'intervalle d'existence.
- Si $c_0 > 0$ et $x_0 < 0$ (cette zone est la zone C sur la figure 7.3, elle se décompose en deux zones disjointes suivant le signe de t_0). Les solutions sont $\big(-1/(c_0 - t^2),]-\infty, \sqrt{c_0}[\big)$ dans la partie gauche de la zone C et $\big(-1/(c_0 - t^2),]\sqrt{c_0}, +\infty[\big)$ dans la partie droite de la zone C.
- Enfin, si $c_0 < 0$ (zone B sur la figure 7.3), les solutions sont $\big(-1/(c_0 - t^2), \mathbb{R}\big)$. La solution "n'explose pas" et est donc définie sur $\mathbb{R}$ tout entier.

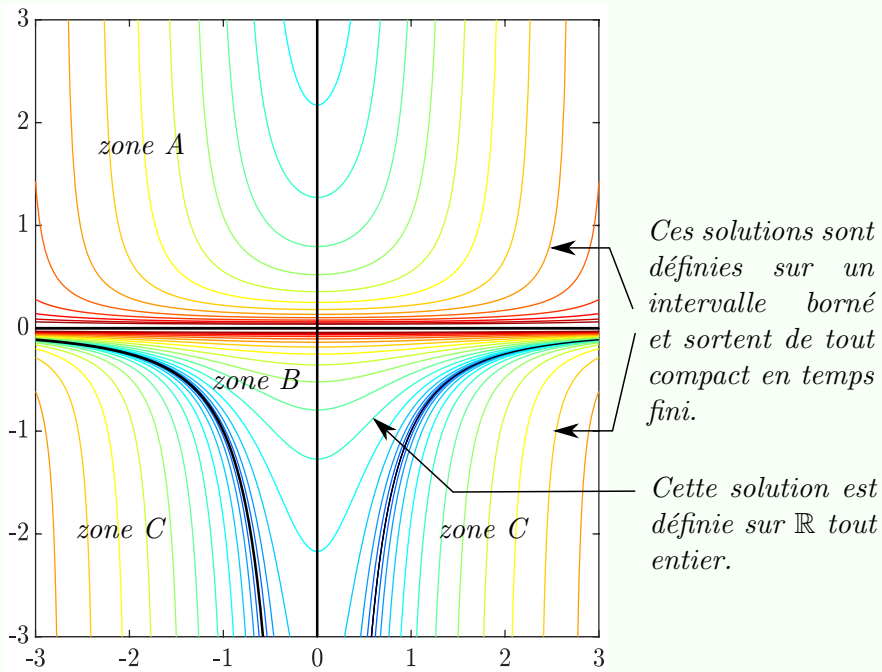


FIG. 7.3 – *Graphe des solutions de l'edo $x'(t) = 2tx^2(t)$. Celles qui ne sont pas définies sur $\mathbb{R}$ tout entier (dans la zone B) "explosent" en temps fini, conformément au résultat du théorème 7.5.*

Terminons ce paragraphe avec un outil incontournable en théorie des edo :

Lemme 7.1 (de Grönwall)**.** *Soient φ et ψ deux fonctions continues de $[a,b]$ dans $\mathbb{R}_+$ et $t_0 \in [a,b]$. On suppose qu'il existe une constante positive C telle que*

$$\varphi(t) \leqslant C + \left| \int_{t_0}^{t} \psi(s)\varphi(s)\,\mathrm{d}s \right|, \qquad \forall\, t \in [a,b].$$

On a alors :

$$\varphi(t) \leqslant C \exp\left(\left| \int_{t_0}^{t} \psi(s)\,\mathrm{d}s \right| \right), \qquad \forall\, t \in [a,b]. \tag{7.7}$$

Démonstration. Pour tout $t \geqslant t_0$, posons :

$$h(t) = \frac{C + \int_{t_0}^{t} \psi(s)\varphi(s)\,\mathrm{d}s}{\exp\left(\int_{t_0}^{t} \psi(s)\,\mathrm{d}s\right)}.$$

Alors h est dérivable et :

$$h'(t) = \psi(t)\frac{\varphi(t) - C - \int_{t_0}^{t} \psi(s)\varphi(s)\,\mathrm{d}s}{\exp\left(\int_{t_0}^{t} \psi(s)\,\mathrm{d}s\right)} \leqslant 0.$$

On en déduit que $h(t) \leqslant h(t_0) = C$ ce qui donne le résultat. Pour tout $t \leqslant t_0$, on pose :

$$h(t) = \frac{C + \int_{t}^{t_0} \psi(s)\varphi(s)\,\mathrm{d}s}{\exp\left(\int_{t}^{t_0} \psi(s)\,\mathrm{d}s\right)},$$

et on conclut de la même façon. □

7.6 Cas particulier des edo linéaires : Premières propriétés

Un cas particulier intéressant est celui où f est une application linéaire en x. Considérons I un intervalle de $\mathbb{R}$ et $N \times N$ fonctions réelles continues :

$$\begin{array}{rcl} m_{ij} : I & \longrightarrow & \mathbb{R} \\ t & \longmapsto & m_{ij}(t). \end{array} \qquad (1 \leqslant i,j \leqslant N).$$

Notons alors $M(t)$ la matrice carrée $N \times N$ dont les cœfficients sont les $m_{ij}(t)$. On notera $\mathscr{M}_N(\mathbb{R})$ l'espace vectoriel des matrices carrées $N \times N$ muni de la norme subordonnée :

$$\|M\|_{\mathscr{M}_N(\mathbb{R})} = \max_{\substack{x \in \mathbb{R}^N \\ \|x\|=1}} \|Mx\| \qquad \text{pour tout } M \in \mathscr{M}_N(\mathbb{R}).$$

On dira qu'une ode (7.1) est linéaire homogène si elle s'écrit :

$$x'(t) = M(t)x(t). \tag{7.8}$$

Théorème 7.6. *Si $m_{ij} \in \mathscr{C}(I)$ pour tout $i,j = 1,\ldots,N$, alors pour tout $t_0 \in I$ et $x_0 \in \mathbb{R}^N$, le problème de Cauchy :*

$$\begin{cases} x'(t) = M(t)x(t) \\ x(t_0) = x_0, \end{cases}$$

admet une unique solution maximale définie sur I tout entier.

Démonstration. Il est clair que la fonction

$$\begin{aligned} f : I \times \mathbb{R}^N &\longrightarrow \mathbb{R}^N \\ (t,x) &\longmapsto M(t)x \end{aligned}$$

est une fonction continue sur $I \times E$. Soit $I(t_0)$ un voisinage compact de t_0 dans I (si $t_0 \in \mathring{I}$, on peut choisir un intervalle de la forme $[t_0 - \delta, t_0 + \delta]$, $\delta > 0$, sinon, t_0 est une extremité de I et on peut choisir $[t_0, t_0 + \delta]$ par exemple). Les fonctions m_{ij} étant continues sur $I(t_0)$, la fonction

$$\begin{aligned} k : I &\longrightarrow \mathbb{R} \\ t &\longmapsto \|M(t)\|_{\mathscr{M}_N(\mathbb{R})} \end{aligned}$$

est elle aussi continue sur $I(t_0)$ (le vérifier à titre d'exercice). On peut alors considérer

$$\bar{k} = \max_{t \in I(t_0)} k(t),$$

et f vérifie l'hypothèse 7.1 avec la constante $\bar{k}$ sur $I(t_0) \times E$. Selon le Théorème 7.1, il existe une unique solution locale (φ, J) au problème de Cauchy considéré. D'autre part, en appliquant l'inégalité de Grönwall (7.7), on obtient l'estimation:

$$\|\varphi(t)\|_E \leqslant \|x_0\|_E e^{k|t-t_0|} \qquad \text{pour tout } t \in J.$$

La solution reste donc bornée sur tout intervalle borné et suivant le Théorème 7.5, elle peut donc être prolongée sur I tout entier. □

Les edo linéaires seront étudiées plus en détail dans le chapitre 7.10.

7.7 Continuité par rapport aux conditions initiales

Supposons qu'une fonction f définie sur un voisinage $\mathscr{U}$ d'un point (t_0, x_0) de $\mathbb{R} \times \mathbb{R}^N$ et à valeurs dans $\mathbb{R}^N$, satisfasse l'hypothèse 7.1. Alors, pour tout $(\tau, \xi) \in \mathscr{U}$, la solution φ du problème de Cauchy :

$$\begin{cases} x'(t) = f(t, x(t)), \\ x(\tau) = \xi, \end{cases}$$

peut être vue comme une fonction de la variable t dépendant du paramètre $(\tau, \xi) \in \mathscr{U}$. Pour mettre en évidence cette dépendance, on écrira cette solution $\varphi(t, \tau, \xi)$. Noter que l'unicité, assurée par le théorème de Cauchy–Lipschitz, est fondamentale : si le problème de Cauchy ci-dessus admettait plusieurs solutions, la valeur de $\varphi(t, \tau, \xi)$ ne serait pas définie de façon univoque et on ne pourrait alors pas parler de fonction.

En considérant donc la fonction φ définie sur un ouvert de $\mathbb{R} \times \mathbb{R} \times \mathbb{R}^N$, on peut se poser la question de la régularité de la fonction par rapport à l'ensemble de ses variables.

Les résultats de régularité s'énoncent sur un voisinage du graphe d'une solution connue. Plus précisément si ψ est une fonction continue définie sur un intervalle $]a,b[$ et à valeurs dans $\mathbb{R}^N$ et soit $\delta > 0$. On notera :

$$\mathscr{C}(\psi,\delta) = \big\{(t,x) \in]a,b[\times E \,:\, \|x - \psi(t)\| < \delta\big\},$$

le cylindre ouvert centré sur la graphe $\{(t,\psi(t)) \,:\, t \in]a,b[\}$ (cf. Fig. 7.4). On peut alors énoncer le résultat de continuité suivant :

Théorème 7.7. *Soit f vérifiant l'hypothèse 7.1 sur $\mathscr{U}$ et soit ψ une solution de* (7.1) *définie sur un intervalle $[a,b]$. Alors :*

1. *Il existe $\delta > 0$ tel que $\mathscr{C}(\psi,\delta) \subset \mathscr{U}$ et pour tout $(\tau,\xi) \in \mathscr{C}(\psi,\delta)$ il existe une unique solution φ à* (7.1) *définie sur $]a,b[$ et vérifiant $\varphi(\tau,\tau,\xi) = \xi$.*
2. *φ est continue sur $\mathscr{V} =]a,b[\times\mathscr{C}(\psi,\delta)$ (qui est donc un ouvert de $]a,b[\times]a,b[\times\mathbb{R}^N$).*

La démonstration n'est pas très compliquée mais longue et un peu technique. Son étude détaillée pourra être omise lors d'une première lecture.

Démonstration. Par définition d'une solution de (7.1), la graphe $\big\{(t,\psi(t)), t \in [a,b]\big\}$ est inclus dans l'ouvert $\mathscr{U}$. Le graphe étant compact et le domaine $\mathscr{U}$ étant ouvert, il est possible de trouver $\delta_1 > 0$ tel que $\mathscr{C}(\psi,\delta_1) \subset \mathscr{U}$ (on pourra écrire les détails à titre d'exercice). Choisissons alors

$$\delta < e^{-k(b-a)}\delta_1,$$

où k désigne la constante de Lipschitz de f sur $\mathscr{U}$ dans l'hypothèse 7.1. Comme $\mathscr{C}(\psi,\delta) \subset \mathscr{C}(\psi,\delta_1) \subset \mathscr{U}$, pour tout $(\tau,\xi) \in \mathscr{C}(\psi,\delta)$, il existe d'après le Théorème 7.1 une unique solution locale à l'edo (7.1) vérifiant $\varphi(\tau) = \xi$. L'estimation de Grönwall (7.7) founit, sur l'intervalle d'existence de la fonction $t \mapsto \varphi(t,\tau,\xi)$, l'estimation :

$$\|\varphi(t,\tau,\xi) - \psi(t)\| \leqslant e^{k|t-\tau|}\|\xi - \psi(\tau)\| < e^{k(b-a)}\delta = \delta_1.$$

Ceci prouve que le point $(t,\varphi(t,\tau,\xi))$ reste à l'intérieur de $\mathscr{C}(\psi,\delta_1) \subset \mathscr{U}$ et donc, d'après le Théorème 7.5, que l'intervalle d'existence de la fonction $t \mapsto \varphi(\cdot,\tau,\xi)$ contient l'intervalle $]a,b[$. On peut donc affirmer que toute solution qui passe par un point de $\mathscr{C}(\psi,\delta)$ est entièrement contenue dans $\mathscr{C}(\psi,\delta_1)$ (cf. Fig. 7.4).

Pour montrer la continuité de le fonction φ sur $\mathscr{V} =]a,b[\times\mathscr{C}(\psi,\delta)$, nous allons montrer qu'elle est la limite uniforme d'une suite de fonctions continues sur $\mathscr{V}$. Introduisons pour cela la suite $(\varphi_n)_n$ définie de la façon suivante pour tout $(t,\tau,\xi) \in \mathscr{V}$:

$$\begin{aligned}
\varphi_0(t,\tau,\xi) &= \psi(t) + \xi - \psi(\tau),\\
\varphi_{n+1}(t,\tau,\xi) &= \xi + \int_\tau^t f(s,\varphi_n(s,\tau,\xi))\,\mathrm{d}s, \qquad \forall\, n \geqslant 0,
\end{aligned}$$

et montrons par récurrence qu'elle a les bonnes propriétés. Il est clair que φ_0 est continue sur $\mathscr{V}$ et que pour tout n, la continuité sur $\mathscr{V}$ de φ_n entraîne la continuité de φ_{n+1}. Toutes les fonctions

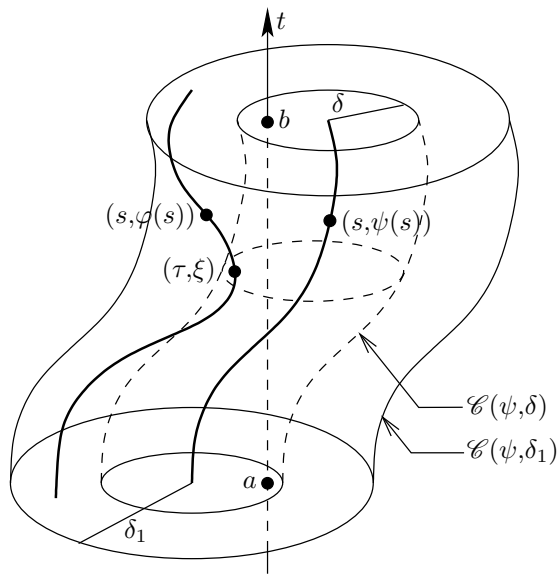


FIG. 7.4 – *Les tubes* $\mathscr{C}(\psi,\delta)$ *et* $\mathscr{C}(\psi,\delta_1)$

φ_n sont donc bien continues sur $\mathscr{V}$. Montrons maintenant que, pour tout $n \geqslant 1$, la propriété $\mathscr{P}_n$ suivante est vérifiée :

$$\begin{cases} \text{Pour tout } (t,\tau,\xi) \in \mathscr{V} \text{ et pour tout } 1 \leqslant m \leqslant n, \text{ on a :} \\ (t,\varphi_m(t,\tau,\xi)) \in \mathscr{C}(\psi,\delta_1) \quad \text{et} \quad \|\varphi_m(t,\tau,\xi) - \varphi_{m-1}(t,\tau,\xi)\| \leqslant k^m \dfrac{|t-\tau|^m}{m!} \|\xi - \psi(\tau)\|. \end{cases} \tag{$\mathscr{P}_n$}$$

Vérifions que $\mathscr{P}_1$ est vraie. D'une part, pour tout $(t,\tau,\xi) \in \mathscr{V}$:

$$\|\varphi_0(t,\tau,\xi) - \psi(t)\| = \|\xi - \psi(\tau)\| < \delta \leqslant \delta_1,$$

ce qui prouve que $(t,\varphi_0(t,\tau,\xi)) \in \mathscr{C}(\psi,\delta_1) \subset \mathscr{U}$ pour tout $(t,\tau,\xi) \in \mathscr{V}$. D'autre part, comme:

$$\psi(t) = \psi(\tau) + \int_\tau^t f(s,\psi(s))\,\mathrm{d}s, \qquad \forall\, t \in [a,b],$$

on obtient que

$$\begin{aligned} \|\varphi_1(t,\tau,\xi) - \varphi_0(t,\tau,\xi)\| &\leqslant \left| \int_\tau^t \|f(s,\varphi_0(s,\tau,\xi)) - f(s,\psi(s))\|\,\mathrm{d}s \right| \\ &\leqslant \left| \int_\tau^t k\|\varphi_0(s,\tau,\xi) - \psi(s)\|\,\mathrm{d}s \right| = k|t-\tau|\|\xi - \psi(\tau)\|, \end{aligned}$$

ce qui est la relation $\mathscr{P}_1$. On obtient également :

$$\begin{aligned} \|\varphi_1(t,\tau,\xi) - \psi(t)\| &\leqslant \|\varphi_1(t,\tau,\xi) - \varphi_0(t,\tau,\xi)\| + \|\varphi_0(t,\tau,\xi) - \psi(t)\| \\ &\leqslant (1 + k|t-\tau|)\|\xi - \psi(\tau)\| \leqslant e^{k|t-\tau|}\|\xi - \psi(\tau)\| \leqslant \delta_1, \end{aligned}$$

ce qui prouve que $(t,\varphi_1(t,\tau,\xi)) \in \mathscr{C}(\psi,\delta_1)$ pour tout $(t,\tau,\xi) \in \mathscr{V}$.

Montrons maintenant que pour tout $n \geqslant 1$, $(\mathscr{P}_n \Rightarrow \mathscr{P}_{n+1})$. Supposons donc que $\mathscr{P}_n$ est vérifiée pour un certain rang $n \geqslant 1$. Alors

$$\|\varphi_{n+1}(t,\tau,\xi) - \varphi_n(t,\tau,\xi)\| \leqslant \left| \int_\tau^t \|f(s,\varphi_n(s,\tau,\xi)) - f(s,\varphi_{n-1}(s,\tau,\xi))\| \, \mathrm{d}s \right|.$$

D'après $\mathscr{P}_n$, $(t,\varphi_n(t,\tau,\xi))$ et $(t,\varphi_{n-1}(t,\tau,\xi))$ sont dans $\mathscr{C}(\psi,\delta_1) \subset \mathscr{U}$, on peut donc écrire que :

$$\left| \int_\tau^t \|f(s,\varphi_n(s,\tau,\xi)) - f(s,\varphi_{n-1}(s,\tau,\xi))\| \, \mathrm{d}s \right| \leqslant \left| \int_\tau^t k\|\varphi_n(s,\tau,\xi) - \varphi_{n-1}(s,\tau,\xi)\| \, \mathrm{d}s \right|.$$

On utilise l'inégalité de $(\mathscr{P}_n)$, pour obtenir:

$$\begin{aligned} \left| \int_\tau^t k\|\varphi_n(s,\tau,\xi) - \varphi_{n-1}(s,\tau,\xi)\| \, \mathrm{d}s \right| &\leqslant \left| \int_\tau^t k^{n+1} \frac{|s-\tau|^n}{n!} \|\xi - \psi(\tau)\| \, \mathrm{d}s \right| \\ &= k^{n+1} \frac{|t-\tau|^{n+1}}{(n+1)!} \|\xi - \psi(\tau)\|, \end{aligned}$$

et on a ainsi montré que :

$$\|\varphi_{n+1}(t,\tau,\xi) - \varphi_n(t,\tau,\xi)\| \leqslant k^{n+1} \frac{|t-\tau|^{n+1}}{(n+1)!} \|\xi - \psi(\tau)\|,$$

ce qui est la deuxième inégalité de $\mathscr{P}_{n+1}$. L'inégalité triangulaire conduit à :

$$\begin{aligned} \|\varphi_{n+1}(t,\tau,\xi) - \psi(t)\| &\leqslant \sum_{p=1}^{n+1} \|\varphi^*(t,\tau,\xi) - \varphi_{p-1}(t,\tau,\xi)\| + \|\varphi_0(t,\tau,\xi) - \psi(t)\| \\ &= \left(\sum_{p=0}^{n+1} k^p \frac{|t-\tau|^p}{p!} \right) \|\xi - \psi(\tau)\| \leqslant e^{k|t-\tau|} \|\xi - \psi(\tau)\| \leqslant \delta_1, \end{aligned}$$

et la propriété $\mathscr{P}_{n+1}$ est donc bien vérifiée. Ceci termine la démonstration par récurrence et nous en déduisons que la propriété $\mathscr{P}_n$ est vraie pour tout $n \geqslant 1$.

On en déduit que, pour tout $n \geqslant 0$, tout $m \geq 1$ et tout $(t,\tau,\xi) \in \mathscr{V}$:

$$\begin{aligned} \|\varphi_{n+m}(t,\tau,\xi) - \varphi_n(t,\tau,\xi)\| &\leqslant \sum_{p=n+1}^{n+m} k^p \frac{|t-\tau|^p}{p!} \|\xi - \psi(\tau)\| \\ &\leqslant \delta \sum_{p=n+1}^{n+m} k^p \frac{|b-a|^p}{p!} = \delta |u_{m+n} - u_n|, \end{aligned}$$

où la suite $(u_n)_n$ est définie pour tout $n \geqslant 0$ par :

$$u_n = \sum_{p=0}^{n} k^p \frac{|b-a|^p}{p!}.$$

Cette suite converge vers $e^{k|b-a|}$, elle est donc de Cauchy, ce qui entraîne que la suite $(\varphi_n)_n$ est de Cauchy uniforme sur $\mathscr{V}$. Sa limite, notée $\widetilde{\varphi}$ est continue sur $\mathscr{V}$. Vérifions que c'est bien la solution au problème de Cauchy cherchée. En passant à la limite quand $n \to \infty$ dans la relation:

$$\varphi_{n+1}(t,\tau,\xi) = \xi + \int_\tau^t f(s,\varphi_n(s,\tau,\xi))\,\mathrm{d}s,$$

on obtient que

$$\widetilde{\varphi}(t,\tau,\xi) = \xi + \int_\tau^t f(s,\widetilde{\varphi}(s,\tau,\xi))\,\mathrm{d}s,$$

ce qui prouve que, en effet, $\widetilde{\varphi} = \varphi$. □

7.8 Différentiabilité par rapport aux conditions initiales

Nous admettrons le théorème suivant (dont la démonstration est longue et technique) :

Théorème 7.8. *On suppose maintenant que $f \in \mathscr{C}^1(\mathscr{U})$ et soit ψ une solution de (7.1) définie sur un intervalle $[a,b]$. Alors, en reprenant les mêmes notations que dans le Théorème 7.7, $\varphi \in \mathscr{C}^1(\mathscr{V})$.*

La fonction $\varphi : \mathscr{V} \to \mathbb{R}^n$ est parfaitement déterminée comme solution du problème de Cauchy :

$$\begin{cases} \dfrac{\partial \varphi}{\partial t}(t,\tau,\xi) = f(t,\varphi(t,\tau,\xi)), \\ \varphi(\tau,\tau,\xi) = \xi, \end{cases} \tag{7.9}$$

On cherche maintenant à déterminer des edo satisfaites par les dérivées partielles de φ, c'est à dire, pour tout $(t,\tau,\xi) \in \mathscr{V}$, par les fonctions :

$$\frac{\partial \varphi}{\partial \xi_j}(t,\tau,\xi) \quad (\text{pour } 1 \leqslant j \leqslant N) \quad \text{et} \quad \frac{\partial \varphi}{\partial \tau}(t,\tau,\xi).$$

En dérivant formellement par rapport à ξ_j $(1 \leqslant j \leqslant N)$ les deux égalités du problème de Cauchy (7.9), on obtient :

$$\begin{cases} \dfrac{\partial}{\partial t}\left(\dfrac{\partial \varphi}{\partial \xi_j}\right)(t,\tau,\xi) = D_x f(t,\varphi(t,\tau,\xi))\dfrac{\partial \varphi}{\partial \xi_j}(t,\tau,\xi), \\ \dfrac{\partial \varphi}{\partial \xi_j}(\tau,\tau,\xi) = e_j, \end{cases}$$

où e_j et le j–ème élément de la base canonique et où, pour tout $(t,x) \in \mathscr{U}$:

$$D_x f(t,x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(t,x) \cdots\cdots \frac{\partial f_1}{\partial x_N}(t,x) \\ \vdots \qquad\qquad \vdots \\ \frac{\partial f_N}{\partial x_1}(t,x) \cdots\cdots \frac{\partial f_N}{\partial x_N}(t,x) \end{pmatrix} \qquad \text{et} \qquad \frac{\partial \varphi}{\partial \xi_j}(t,\tau,\xi) = \begin{pmatrix} \frac{\partial \varphi_1}{\partial \xi_j}(t,\tau,\xi) \\ \vdots \\ \frac{\partial \varphi_N}{\partial \xi_j}(t,\tau,\xi) \end{pmatrix}.$$

On en déduit l'énoncé suivant (que nous admettrons également) :

Proposition 7.1. *On suppose toujours que $f \in \mathscr{C}^1(\mathscr{U})$. Pour tout $(t,\tau,\xi) \in \mathscr{V}$, notons $M(t,\tau,\xi)$ la matrice $D_x f(t,\varphi(t,\tau,\xi))$ où φ est la solution du problème de Cauchy (7.9). Alors la fonction*

$$t \longmapsto \frac{\partial \varphi}{\partial \xi_j}(\cdot,\tau,\xi),$$

est l'unique solution du problème de Cauchy (dont l'edo est linéaire) :

$$\begin{cases} X'(t) = M(t,\tau,\xi)X(t), \\ X(\tau) = e_j. \end{cases} \tag{7.10}$$

D'après le théorème 7.8, la fonction φ est de classe $\mathscr{C}^1$ sur $\mathscr{V}$. Par composition, les cœfficients de la matrice $M(t,\tau,\xi)$ sont donc des fonctions continues sur $\mathscr{V}$ à valeurs dans $\mathbb{R}$. L'étude du problème de Cauchy (7.10) relève donc de la section 7.6 (en particulier, l'existence et l'unicité d'une solution maximale sont assurés par le théorème 7.6).

En utilisant les mêmes arguments, on vérifie que la fonction

$$t \longmapsto \frac{\partial \varphi}{\partial \tau}(\cdot,\tau,\xi),$$

satisfait la même edo linéaire que celle du système (7.10). La seule difficultée consiste à déterminer la valeur en $t = \tau$ dans le problème de Cauchy. On l'obtient par le calcul suivant:

$$\begin{aligned} \varphi(\tau,\tau+h,\xi) - \varphi(\tau,\tau,\xi) &= \varphi(\tau,\tau+h,\xi) - \xi \\ &= \int_{\tau+h}^{\tau} f(s,\varphi(s,\tau+h,\xi))\,\mathrm{d}s, \end{aligned}$$

puis, en divisant par $h \neq 0$ (h petit) et en faisant un changement de variables dans l'intégrale, il vient :

$$\begin{aligned} \frac{\varphi(\tau,\tau+h,\xi) - \varphi(\tau,\tau,\xi)}{h} &= -\frac{1}{h}\int_{\tau}^{\tau+h} f(s,\varphi(s,\tau+h,\xi))\,\mathrm{d}s \\ &= -\frac{1}{h}\int_{0}^{h} f(\tau+s,\varphi(\tau+s,\tau+h,\xi))\,\mathrm{d}s. \end{aligned}$$

On en déduit que :

$$\left|\frac{\varphi(\tau,\tau+h,\xi) - \varphi(\tau,\tau,\xi)}{h} + f(\tau,\xi)\right| \leqslant \frac{1}{h}\int_0^h \big|f(\tau+s,\varphi(\tau+s,\tau+h,\xi)) - f(\tau,\xi)\big|\,\mathrm{d}s. \tag{7.11}$$

Les variables τ et ξ étant fixées, la fonction

$$(s,h) \longmapsto f(\tau+s,\varphi(\tau+s,\tau+h,\xi)),$$

est bien définie sur un voisinage de l'origine dans $\mathbb{R}^2$ et est continue en $(0,0)$ où elle vaut $f(\tau,\xi)$. Pour tout $\varepsilon > 0$, il existe donc $\delta > 0$ tel que, pour tout (s,h) dans un voisinage de $(0,0)$:

$$\|(s,h)\| < \delta \quad \Rightarrow \quad \big|f(\tau+s,\varphi(\tau+s,\tau+h,\xi)) - f(\tau,\xi)\big| < \varepsilon.$$

En passant à la limite quand $h \to 0$ dans l'égalité (7.11), on obtient finalement :

$$\frac{\partial \varphi}{\partial \tau}(\tau,\tau,\xi) = -f(\tau,\xi),$$

et on peut énoncer (toujours en omettant la démonstration) :

Proposition 7.2. *On suppose toujours que $f \in \mathscr{C}^1(\mathscr{U})$. Alors, avec les mêmes notations que dans la proposition 7.1, la fonction*

$$t \longmapsto \frac{\partial \varphi}{\partial \tau}(\cdot,\tau,\xi),$$

est l'unique solution du problème de Cauchy :

$$\begin{cases} X'(t) = M(t,\tau,\xi)X(t), \\ X(\tau) = -f(\tau,\xi). \end{cases}$$

7.9 Flot d'une edo autonome

Soit f une fonction $\mathscr{C}^1$ sur un ouvert $\mathscr{U}$ de $\mathbb{R}^N$ et à valeurs dans $\mathbb{R}^N$ (une telle fonction est appelée un champ de vecteurs). Pour tout $x_0 \in \mathscr{U}$, notons $J(x_0)$ l'intervalle (ouvert) de temps maximal d'existence de la solution du problème de Cauchy :

$$\begin{cases} x'(t) = f(x(t)), \\ x(0) = x_0, \end{cases} \tag{7.12}$$

dont la solution associée est (en faisant apparaître la dépendance en x_0) :

$$\begin{aligned} \phi(\cdot,x_0) : J(x_0) &\longrightarrow \mathbb{R}^N \\ t &\longmapsto \phi(t,x_0). \end{aligned}$$

L'image de cette fonction, c'est à dire la courbe :

$$\big\{\phi(t,x_0) \ : \ t \in J(x_0)\big\},$$

s'appelle l'orbite (ou la trajectoire) de l'edo passant par x_0. La fonction ϕ vue comme fonction des deux variables (t,x_0) est définie sur l'ensemble :

$$\mathscr{V} = \{(t,x) \in \mathbb{R} \times \mathscr{U} \ : \ t \in J(x)\} = \bigcup_{x \in \mathscr{U}} \Big(J(x) \times \{x\}\Big). \tag{7.13}$$

Nous pouvons alors poser :

Définition 7.4. *La fonction :*

$$\begin{aligned} \phi : \mathscr{V} &\longrightarrow \mathbb{R}^N \\ (t,x) &\longmapsto \phi(t,x), \end{aligned}$$

s'appelle le flot de l'edo (7.12).

Remarquer que l'on peut aussi définir le flot d'un système non–autonome, mais comme expliqué dans la section 7.3, tout système non–autonome peut se réécrire comme un système autonome. L'hypothèse d'autonomie ne constitue donc pas une restriction.

Pour tout $t \in \mathbb{R}$, introduisons encore le sous–ensemble de $\mathscr{U}$:

$$\mathscr{U}_t = \{x \in \mathscr{U} \ : \ t \in J(x)\}, \tag{7.14}$$

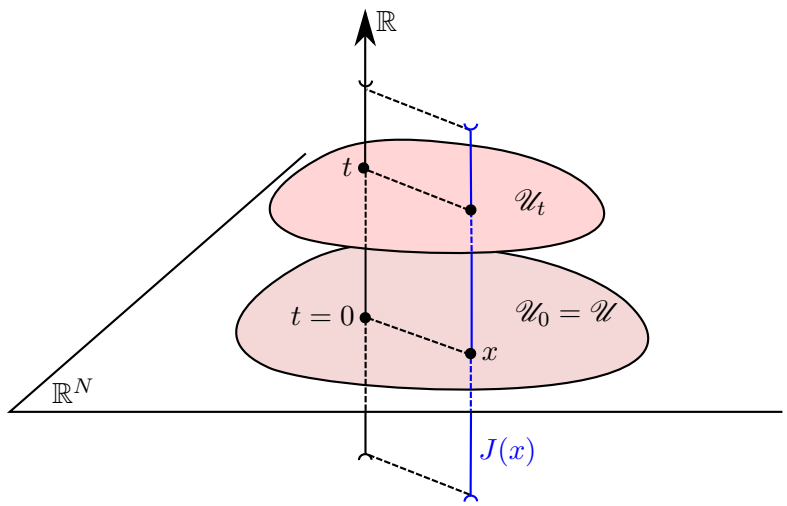


FIG. 7.5 – *L'ouvert $\mathscr{V}$ est défini par* $\mathscr{V}=\bigcup_{x\in\mathscr{U}}\Big(J(x)\times\{x\}\Big)=\bigcup_{t\in\mathbb{R}}\Big(\{t\}\times\mathscr{U}_t\Big)$.

c'est à dire l'ensemble des valeurs initiales x pour lesquelles la solution de (7.12) est définie au temps t. Évidemment $\mathscr{U}_0=\mathscr{U}$ et si t n'est dans aucun $J(x)$ alors $\mathscr{U}_t=\varnothing$. On a alors l'égalité (à comparer avec (7.13)) :

$$\mathscr{V}=\bigcup_{t\in\mathbb{R}}\Big(\{t\}\times\mathscr{U}_t\Big).$$

La valeur de t étant fixée, on peut définir la fonction :

$$\begin{aligned}\phi_t:\mathscr{U}_t&\longrightarrow\mathbb{R}^N\\ x&\longmapsto\phi(t,x).\end{aligned}$$

Remarquer en particulier que $\phi_0=\mathrm{Id}$. Donnons un exemple faisant apparaitre toutes ces notions :

Exemple 7.4. Soit la fonction de classe $\mathscr{C}^1$ (en fait $\mathscr{C}^\infty$) :

$$\begin{aligned}f:\mathbb{R}^2&\longrightarrow\mathbb{R}^2\\ (x,y)&\longmapsto\begin{pmatrix}1\\2xy^2\end{pmatrix},\end{aligned}$$

et pour tout $(x_0,y_0)\in\mathbb{R}^2$, considérons le problème de Cauchy :

$$\begin{cases}\dfrac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}x(t)\\y(t)\end{pmatrix}=f(x(t),y(t)),\\ (x(0),y(0))=(x_0,y_0),\end{cases}$$

dont la solution est (en faisant apparaître la condition initiale (x_0,y_0) dans la notation) :

$$\begin{aligned}\phi(\cdot,(x_0,y_0)) : J(x_0,y_0) &\longrightarrow \mathbb{R}\\ t &\longmapsto \begin{pmatrix} t+x_0 \\ \dfrac{y_0}{1-ty_0(t+2x_0)}\end{pmatrix}.\end{aligned}$$

Des calculs élémentaires permettent de déterminer les intervalles d'existence $J(x_0,y_0)$ (pour (x_0,y_0) donné, c'est tout simplement l'intervalle contenant 0 sur lequel la fonction $t \mapsto 1 - ty_0(t+2x_0)$ ne s'annule pas). On trouve que :

- Si $y_0 > 0$ alors $J(x_0,y_0) = \left]\left(-x_0 - \sqrt{x_0^2 + \frac{1}{y_0}}\right), \left(-x_0 + \sqrt{x_0^2 + \frac{1}{y_0}}\right)\right[$;
- Si $-\frac{1}{x_0^2} < y_0 \leqslant 0$ alors $J(x_0,y_0) = \mathbb{R}$;
- Si $y_0 = -1/x_0^2$ et $x_0 < 0$ alors $J(x_0,y_0) =]-\infty, -x_0[$;
- Si $y_0 = -1/x_0^2$ et $x_0 > 0$ alors $J(x_0,y_0) =]-x_0, +\infty[$;
- Si $y_0 < -1/x_0^2$ et $x_0 < 0$ alors $J(x_0,y_0) = \left]-\infty, \left(-x_0 - \sqrt{x_0^2 + \frac{1}{y_0}}\right)\right[$;
- Si $y_0 < -1/x_0^2$ et $x_0 > 0$ alors $J(x_0,y_0) = \left]\left(-x_0 + \sqrt{x_0^2 + \frac{1}{y_0}}\right), +\infty\right[$.

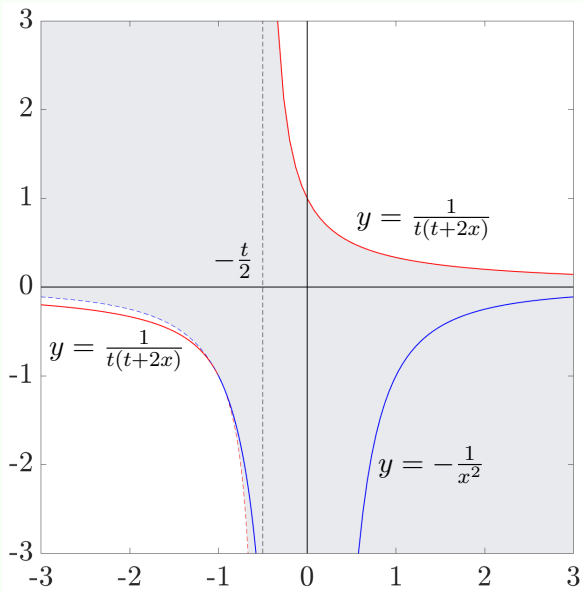


FIG. 7.6 – *L'ensemble $\mathscr{U}_t$ (en gris) ici avec $t = 1$.*

On fixe maintenant t (par exemple $t = 1$) et on cherche l'ensemble des points (x_0,y_0) pour lesquels $J(x_0,y_0)$ contient t. On trouve l'ensemble $\mathscr{U}_t$ représenté sur la figure 7.6.

Les principales propriétés du flot sont les suivantes :

Théorème 7.9 (Propriétés du flot)**.**

1. *L'ensemble $\mathscr{V}$ est ouvert dans $\mathbb{R} \times \mathbb{R}^N$ et ϕ est de classe $\mathscr{C}^1$ sur $\mathscr{V}$. Pour tout $t \in \mathbb{R}$, $\mathscr{U}_t$ est ouvert dans $\mathbb{R}^N$.*

2. *$\mathscr{U}_0 = \mathscr{U}$ et $\phi_0 = \mathrm{Id}$.*
3. *Soit $t \in \mathbb{R}$. Alors ϕ_t est un difféomorphisme de classe $\mathscr{C}^1$ de $\mathscr{U}_t$ sur $\mathscr{U}_{-t}$ dont l'inverse est ϕ_{-t}.*
4. *Soient s et t dans $\mathbb{R}$. Dans l'ensemble ouvert $\phi_t^{-1}(\mathscr{U}_s)$ de $\mathscr{U}$, on a :*

$$\phi_{s+t} = \phi_s \circ \phi_t.$$

La démonstration du théorème repose sur le lemme élémentaire suivant :

Lemme 7.2. *Soit $t_0 \in \mathbb{R}$ et $x_0 \in \mathscr{U}$. Notons T_{t_0} la translation $t \mapsto t + t_0$ de $\mathbb{R}$. Si (φ,J) est la solution maximale du problème de Cauchy :*

$$\begin{cases} x'(t) = f(x(t)) \\ x(t_0) = x_0, \end{cases} \tag{7.15}$$

alors $(\varphi \circ T_{t_0}, T_{-t_0}(J))$ est la solution maximale du problème de Cauchy :

$$\begin{cases} x'(t) = f(x(t)) \\ x(0) = x_0. \end{cases} \tag{7.16}$$

Démonstration. Notons $(\tilde{\varphi},\tilde{J})$ la solution maximale du problème de Cauchy (7.16). On vérifie facilement que $\varphi \circ T_{t_0}$ est aussi une solution du problème de Cauchy (7.16) définie sur l'intervalle $T_{-t_0}(J)$ et que donc, par unicité de la solution, $T_{-t_0}(J) \subset \tilde{J}$ et $\tilde{\varphi} = \varphi \circ T_{t_0}$ sur $T_{-t_0}(J)$. Supposons que l'inclusion $T_{-t_0}(J) \subset \tilde{J}$ soit stricte. Alors l'inclusion $J \subset T_{t_0}(\tilde{J})$ serait stricte également et $(\tilde{\varphi} \circ T_{-t_0}, T_{t_0}(\tilde{J}))$ serait une solution du problème de Cauchy (7.15), ce qui contredirait la maximalité de l'intervalle J. □

Démonstration du théorème 7.9.

1. Soit $(t_0,x_0) \in \mathscr{V}$. Par définition de $\mathscr{V}$, t_0 est donc dans l'intervalle ouvert $J(x_0)$ qui contient aussi 0. Par conséquent, il est possible de trouver deux réels a et b ($a < b$) tels que l'intervalle $]a,b[$ soit inclus dans $J(x_0)$ et contienne les points 0 et t_0. Cela signifie en particulier que la solution de l'edo $x'(t) = f(x(t))$ satisfaisant $x(0) = x_0$ est définie sur l'intervalle $]a,b[$. Le point 1 du théorème 7.7 nous assure alors de l'existence d'un réel $\delta > 0$ tel que pour tout $\xi \in \mathscr{U}$ vérifiant $\|\xi - x_0\| < \delta$ alors la solution de l'edo $x'(t) = f(x(t))$, satisfaisant cette fois la condition initiale $x(0) = \xi$, est définie elle aussi sur l'intervalle $]a,b[$. On en déduit que le cylindre ouvert :
$$\big\{(t,\xi) \in \mathbb{R} \times \mathscr{U} \ : \ t \in]a,b[\,, \|\xi - x_0\| < \delta\big\},$$
qui contient (t_0,x_0) est inclus dans $\mathscr{V}$ et donc que $\mathscr{V}$ est ouvert. La même démonstration prouve aussi que dans $\mathbb{R}^N$, la boule ouverte centrée en x_0 et de rayon δ est contenue dans $\mathscr{U}_{t_0}$ et donc que $\mathscr{U}_{t_0}$ est ouvert.
2. Ce point est évident.
3. Soit $t \in \mathbb{R}$. D'après le théorème 7.8 (différentiabilité en fonction des conditions initiales), ϕ_t est une application de classe $\mathscr{C}^1$ sur l'ouvert $\mathscr{U}_t$.
Soit $x \in \mathscr{U}_t$. D'après le lemme 7.2, on a :
$$J(\phi_t(x)) = T_{-t}(J(x)), \tag{7.17a}$$
$$\phi_s(\phi_t(x)) = \phi_{s+t}(x), \quad \forall\, s \in J(\phi_t(x)). \tag{7.17b}$$

Comme 0 est dans l'intervalle $J(x)$, la relation (7.17a) nous assure que $-t$ est dans $J(\phi_t(x))$ ou de manière équivalente que $\phi_t(x)$ est dans $\mathscr{U}_{-t}$. Ceci prouve que la fonction ϕ_t est bien à valeurs dans l'ouvert $\mathscr{U}_{-t}$. On utilise ensuite la relation (7.17b) avec $s = -t$ pour obtenir que :

$$\phi_{-t} \circ \phi_t(x) = \phi_0(x) = x \quad \text{pour tout } x \in \mathscr{U}_t.$$

En remplaçant t par $-t$, on a aussi montré que :

$$\phi_t \circ \phi_{-t}(\tilde{x}) = \tilde{x} \quad \text{pour tout } \tilde{x} \in \mathscr{U}_{-t},$$

et le point 3 est ainsi démontré.

4. Soit (t,x) dans $\mathscr{V}$ et soit s un réel tel que $\phi_t(x)$ soit dans $\mathscr{U}_s$. On a donc $s \in J(\phi_t(x))$ et l'on en déduit le dernier point du théorème avec la relation (7.17b).

□

Définition 7.5 (Champ de vecteurs complet)**.** *On dit qu'un champ de vecteur défini par une fonction f d'un ouvert $\mathscr{U}$ de $\mathbb{R}^N$ dans $\mathbb{R}^N$ est complet si toute solution maximale du problème de Cauchy* (7.12) *est définie sur $\mathbb{R}$ tout entier. L'ensemble $\mathscr{V}$ défini par l'égalité* (7.13) *est alors égal à :*

$$\mathscr{V} = \mathbb{R} \times \mathscr{U},$$

et pour tout $t \in \mathbb{R}$, $\mathscr{U}_t = \mathscr{U}$.

Notons maintenant $\mathrm{Diff}^1(\mathscr{U})$ l'ensemble des difféomorphismes de classe $\mathscr{C}^1$ de $\mathscr{U}$ dans lui même. Muni de la composition, l'ensemble $\mathrm{Diff}^1(\mathscr{U})$ est un groupe. Selon le point 3 du théorème 7.9, pour tout t dans $\mathbb{R}$, ϕ_t appartient à $\mathrm{Diff}^1(\mathscr{U})$. Les points 2, 3 et 4 prouvent que l'application :

$$\begin{aligned} (\mathbb{R},+) &\longrightarrow (\mathrm{Diff}^1(\mathscr{U}),\circ) \\ t &\longmapsto \phi_t \end{aligned}$$

est un homomorphisme de groupe.

Nous poursuivrons l'étude des edo homogènes un peu plus loin, après avoir traité un peu plus en détails le cas des edo linéaires.

7.10 Équations différentielles linéaires

Nous avons déjà posé quelques définitions et donné quelques résultats concernant les edo linéaires dans la section 7.6. Avant d'entrer plus avant dans les détails, faisons quelques rappels d'algèbre linéaire.

7.10.1 Rappels d'algèbre linéaire, exponentielle de matrices

On note $\mathscr{M}_n(\mathbb{C})$ (respectivement $\mathscr{M}_n(\mathbb{R})$) (pour tout entier $n \geqslant 1$) l'espace des matrices carrées à cœfficients dans $\mathbb{C}$ (respectivement à cœfficients dans $\mathbb{C}$). Cet espace est muni d'une norme subordonnées $\|\cdot\|_{\mathscr{M}_n(\mathbb{C})}$.

Une matrice A de $\mathscr{M}_n(\mathbb{R})$ est diagonalisable dans $\mathscr{M}_n(\mathbb{C})$ s'il existe une matrice inversible P dans $\mathscr{M}_n(\mathbb{C})$ telle que :

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix},$$

où les n nombres complexes $\lambda_1, \ldots, \lambda_n$ sont les valeurs propres de la matrice A. Les cœfficients de A étant réels, si une valeur propre λ_j n'est pas réelle, alors son conjugué $\bar{\lambda}_j$ est également une valeur propre de A.

Dans la suite de ce cours, on considèrera essentiellement des matrices A de $\mathscr{M}_n(\mathbb{R})$ diagonalisable dans $\mathscr{M}_n(\mathbb{C})$.

L'exponentielle de la matrice A de $\mathscr{M}_n(\mathbb{C})$ est défini comme étant la somme de la série absolument uniformément convergente sur tout ensemble borné de $\mathscr{M}_n(\mathbb{C})$:

$$e^A = \mathrm{Id}_n + \sum_{p=1}^{+\infty} \frac{A^p}{p!}, \tag{7.18}$$

où Id_n est la matrice identité de $\mathscr{M}_n(\mathbb{R})$.

Proposition 7.3. *Les principales propriétés de l'exponentielle de matrices sont :*

1. *Pour toute matrice A dans $\mathscr{M}_n(\mathbb{C})$, $\|e^A\|_{\mathscr{M}_n(\mathbb{R})} \leqslant e^{\|A\|_{\mathscr{M}_n(\mathbb{R})}}$.*
2. *Pour A et B dans $\mathscr{M}_n(\mathbb{C})$, la relation $e^{A+B} = e^A e^B$ n'est assurée que lorsque les matrices A et B commutent.*
3. *Pour tout A dans $\mathscr{M}_n(\mathbb{C})$, la matrice e^A est toujours inversible et $(e^A)^{-1} = e^{-A}$.*
4. *Si P est un matrice inversible dans $\mathscr{M}_n(\mathbb{C})$, on a pour tout $A \in \mathscr{M}_n(\mathbb{C})$, $P^{-1}e^A P = e^{P^{-1}AP}$.*
5. *Si D est une matrice diagonale dans $\mathscr{M}_n(\mathbb{C})$ alors e^D est une matrice diagonale également, et plus précisément :*

$$\textit{Si} \quad D = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix} \quad \textit{alors} \quad e^D = \begin{pmatrix} e^{\lambda_1} & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & e^{\lambda_n} \end{pmatrix},$$

où $\lambda_1, \ldots, \lambda_n$ sont des nombres complexes.

Nous utiliserons dans la démonstration de la proposition le lemme suivant :

Lemme 7.3. *Pour toute matrice A dans $\mathscr{M}_n(\mathbb{C})$, la fonction :*

$$\begin{aligned} \xi_A : \mathbb{R} &\longrightarrow \mathscr{M}_n(\mathbb{C}) \\ t &\longmapsto e^{tA}, \end{aligned}$$

est $\mathscr{C}^\infty$ et pour tout t dans $\mathbb{R}$:

$$\xi_A'(t) = A\xi_A(t). \tag{7.19}$$

Démonstration. Par définition, pour tout t dans $\mathbb{R}$:

$$\xi_A(t) = \mathrm{Id}_n + \sum_{p=1}^{+\infty} t^p \frac{A^p}{p!}.$$

Notons $\{e_1, \ldots, e_n\}$ la base canonique de $\mathbb{R}^n$ et introduisons, pour tous les indices i,j dans $\{1, \ldots, n\}$ et pour tout t dans $\mathbb{R}$:

$$\xi_A^{i,j}(t) = \langle \xi_A(t)e_j, e_i\rangle = \langle e_j, e_i\rangle + \sum_{p=1}^{+\infty} \frac{t^p}{p!}\langle A^p e_j, e_i\rangle.$$

On vérifie ainsi que chaque fonction $\xi_A^{i,j}$ est la somme d'une série entière de rayon de convergence infini. Ces fonctions sont donc $\mathscr{C}^\infty$ sur $\mathbb{R}$. Or ce sont les cœfficients de la matrice $\xi_A(t)$ donc la fonction ξ_A est elle même de classe $\mathscr{C}^\infty$ sur $\mathbb{R}$. D'autre part, un résultat classique sur les séries entières nous assure que pour tout t dans $\mathbb{R}$:

$$\big(\xi_A^{i,j}\big)'(t) = \sum_{p=1}^{+\infty} \frac{t^{p-1}}{(p-1)!}\langle A^p e_j, e_i\rangle = \langle Ae_j, e_i\rangle + \sum_{p=1}^{+\infty} \frac{t^p}{p!}\langle A^{p+1} e_j, e_i\rangle = \langle A\xi_A(t)e_j, e_i\rangle.$$

On a ainsi montré que pour tout t dans $\mathbb{R}$ et tous les indices i,j dans $\{1, \ldots, n\}$:

$$\langle \xi_A'(t)e_j, e_i\rangle = \big(\xi_A^{i,j}\big)'(t) = \langle A\xi_A(t)e_j, e_i\rangle,$$

ce qui est l'égalité des cœfficients de la relation (7.19). □

Démonstration de la proposition 7.3. ***1.*** Ce premier point découle de l'écriture de e^A sous la forme d'une série, de l'inégalité triangulaire et du fait que pour une norme subordonnée et pour toute matrice A dans $\mathscr{M}_n(\mathbb{C})$ et tout entier $p \geqslant 1$:

$$\|A^p\|_{\mathscr{M}_n(\mathbb{C})} \leqslant \|A\|^p_{\mathscr{M}_n(\mathbb{C})}.$$

2. Soient A et B deux matrices dans $\mathscr{M}_n(\mathbb{C})$ telles que $AB = BA$ et soient i dans $\{1, \ldots, n\}$. Posons :

$$\varphi(t) = \xi_{A+B}(t)e_j \qquad \text{et} \qquad \psi(t) = \xi_A(t)\xi_B(t)e_j.$$

D'après le lemme, ces deux fonctions sont de classe $\mathscr{C}^\infty$ sur $\mathbb{R}$ et à valeurs dans $\mathbb{C}^n$. Toujours d'après le lemme :

$$\varphi'(t) = (A+B)\xi_{A+B}(t)e_j = (A+B)\varphi(t) \quad \text{et} \quad \psi'(t) = \big[A\xi_A(t)\xi_B(t) + \xi_A(t)B\xi_B(t)\big]e_j.$$

Puisque les matrices A et B commutent $\xi_A(t)B = B\xi_A(t)$ pour tout t dans $\mathbb{R}$ et donc :

$$\psi'(t) = (A+B)\xi_A(t)\xi_B(t)e_j = (A+B)\psi(t).$$

D'autre part, $\varphi(0) = \psi(0) = e_j$. Par unicité de la solution d'un problème de Cauchy (théorème 7.4), on en déduit que $\varphi(t) = \psi(t)$ pour tout t dans $\mathbb{R}$ et en particulier que $\varphi(1) = \psi(1)$ c'est à dire que $e^{A+B}e_j = e^A e^B e_j$ pour tout j dans $\{1, \ldots, n\}$. Ceci entraîne que $e^{A+B} = e^A e^B$ (remarquons que le théorème 7.4 est énoncé pour des fonctions à valeurs dans $\mathbb{R}^n$ et que nous l'utilisons ici pour des fonctions à valeurs dans $\mathbb{C}^n$. Pour être rigoureux, il faudrait considérer les parties réelles et imaginaires de φ et ψ).

3. Soit A une matrice de $\mathscr{M}_n(\mathbb{R})$. Selon le point précédent $e^A e^{-A} = e^{-A} e^A = e^0 = \mathrm{Id}_n$ ce qui prouve que e^A est inversible et que son inverse est e^{-A}.

4. Soit A et P deux matrices de $\mathscr{M}_n(\mathbb{C})$ avec P une matrice inversible. On note $B = P^{-1}AP$. Calculons :

$$P^{-1}e^A P = P^{-1}\left(\mathrm{Id}_n + \sum_{p=1}^{+\infty} \frac{A^p}{p!}\right) P = \mathrm{Id}_n + \sum_{p=1}^{+\infty} \frac{1}{p!} P^{-1}A^p P.$$

Or pour tout entier p :

$$P^{-1}A^p P = (P^{-1}AP)^p = B^p,$$

et la relation est ainsi démontrée.

5. Ce dernier point est évident. Il suffit d'appliquer l'égalité (7.18) dans le cas où A est une matrice diagonale.

□

Les points 4 et 5 de la proposition 7.3 permettent de calculer très facilement l'exponentielle d'une matrice diagonalisable, comme cela est illustré dans l'exemple suivant :

Exemple 7.5. Soit la matrice de $\mathscr{M}_3(\mathbb{R})$:

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

On vérifie que cette matrice est diagonalisable dans $\mathscr{M}_3(\mathbb{C})$ et que $P^{-1}AP = D$ avec :

$$P = \frac{1}{\sqrt{2}} \begin{pmatrix} i & -i & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} \qquad \text{et} \qquad D = \begin{pmatrix} 1+i & 0 & 0 \\ 0 & 1-i & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Les points 4 et 5 de la proposition 7.3 nous permettent d'écrire que pour tout $t \in \mathbb{R}$, $e^{tA} = Pe^{tD}P^{-1}$, c'est à dire :

$$e^{tA} = \frac{1}{2} \begin{pmatrix} i & -i & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} e^{t(1+i)} & 0 & 0 \\ 0 & e^{t(1-i)} & 0 \\ 0 & 0 & e^t \end{pmatrix} \begin{pmatrix} -i & 1 & 1 \\ i & 1 & 1 \\ 0 & 1 & 0 \end{pmatrix} = e^t \begin{pmatrix} \cos t & -\sin t & -\sin t \\ 0 & 1 & 0 \\ \sin t & \cos t + 1/2 & \cos t \end{pmatrix}.$$

7.10.2 Équations linéaires homogènes

Rappelons brièvement le contexte de la section 7.6. On désigne par I un intervalle de $\mathbb{R}$ et on considère :

$$\begin{aligned} M : I &\longrightarrow \mathscr{M}_n(\mathbb{R}) \\ t &\longmapsto M(t), \end{aligned} \tag{7.20}$$

une application <u>continue</u>, ce qui est équivalent à dire que tous les cœfficients $m_{ij}(t)$ de la matrice $M(t)$ sont des fonctions réelles continues sur I. L'edo linéaire homogène associée à $M(t)$ s'écrit comme une égalité dans $\mathbb{R}^n$:

$$x'(t) = M(t)x(t), \qquad t \in I. \tag{7.21}$$

On a vu que pour toute donnée initiale $(t_0,x_0) \in I \times \mathbb{R}^n$, il existe une unique solution φ à (7.21) vérifiant la condition $\varphi(t_0) = x_0$ et définie sur I tout entier.

On peut énoncer un premier résultat classique concernant la structure de l'ensemble des solutions de l'edo (7.21) :

Théorème 7.10. *L'ensemble des solutions sur I de l'edo linéaire homogène* (7.21) *est un sous espace vectoriel de $\mathscr{C}^1(I)$ de dimension n.*

Démonstration. Si φ_1 et φ_2 sont deux solutions de (7.21) sur I alors pour toute constante $c \in \mathbb{R}$ il est clair que la fonction $\varphi_1 + c\varphi_2$ est encore une solution de (7.21) sur I, ce qui prouve que l'ensemble des solutions est bien un sous ev de $\mathscr{C}^1(I)$.

Pour montrer que ce sous ev est de dimension n, nous allons en exhiber une base. Choisissons $t_0 \in I$ et notons (pour tout j dans $\{1,\ldots,n\}$) φ_j la solution de (7.21) vérifiant la condition de Cauchy $\varphi_j(t_0) = e_j$ où $\mathscr{B} = \{e_1,\ldots,e_n\}$ est la base canonique de $\mathbb{R}^n$. Soit maintenant φ une solution de (7.21). Le vecteur $\varphi(t_0)$ peut se décomposer dans la base $\mathscr{B}$: Il existe donc n nombres réels $c_1,\ldots,c_n$ tels que

$$\varphi(t_0) = c_1e_1 + \ldots + c_ne_n.$$

Or, les deux solutions φ et $c_1\varphi_1 + \ldots + c_n\varphi_n$ coïncident au point $t = t_0$ et par unicité globale de la solution du problème de Cauchy, elles coïncident partout. Ainsi $\varphi = c_1\varphi_1 + \ldots + c_N\varphi_n$ sur I et $\{\varphi_1,\ldots,\varphi_n\}$ est un système générateur dans $\mathscr{C}(I)$ de l'ensemble des solutions. Vérifions que ce système est aussi libre. Considérons pour cela $c_1,\ldots,c_n$, n nombres réels tels que $c_1\varphi_1+\ldots+c_n\varphi_n = 0$ sur I. Alors en particulier $c_1\varphi_1(t_0)+\ldots+c_n\varphi_n(t_0) = c_1e_1+\ldots+c_ne_n = 0$. On utilise encore une fois le fait que $\mathscr{B}$ soit une base pour conclure que $c_1 = \ldots = c_n = 0$. □

La connaissance de n solutions de (7.21) linéairement indépendantes permet donc de résoudre tous les problèmes de Cauchy associés à (7.21). Cela nous amène à poser :

Définition 7.6 (Matrice fondamentale)**.** *Soit n solutions linéairement indépendantes de l'edo* (7.21) *sur I :*

$$\begin{aligned} \varphi_j : I &\longrightarrow \mathbb{R}^n \\ t &\longmapsto \varphi_j(t), \qquad (j = 1,\ldots,n). \end{aligned}$$

La matrice $\Phi(t)$ de $\mathscr{M}_n(\mathbb{R})$ obtenue pour tout t dans I en concatenant les vecteurs colonnes $\varphi_j(t)$ $(j = 1,\ldots,n)$ est appelée matrice fondamentale de l'edo (7.21)*.*

Elle satisfait l'égalité dans $\mathscr{M}_n(\mathbb{R})$ suivante :

$$\Phi'(t) = M(t)\Phi(t), \qquad \forall\, t \in I. \tag{7.22}$$

Observer qu'il existe un grand nombre de matrices fondamentales. Dans la démonstration du Théorème 7.10 on a en effet prouvé que pour toute base $\{f_1,\ldots,f_n\}$ de $\mathbb{R}^n$ et pour tout $t_0 \in I$, on pouvait construire une matrice fondamentale $\Phi(t)$ telle que ses vecteurs colonnes $\varphi_1(t),\ldots,\varphi_n(t)$ vérifient $\varphi_j(t_0) = f_j$ pour tout $j = 1,\ldots,n$. Si on choisit en particulier pour base la base canonique $\mathscr{B}$ de $\mathbb{R}^n$, on en déduit:

Proposition 7.4. *Pour tout t_0 dans I il existe une unique matrice fondamentale $\Phi(t)$ à l'edo* (7.21) *définie pour tout t dans I et qui vérifie $\Phi(t_0) = \mathrm{Id}_n$.*

Considérons maintenant le problème de Cauchy :

$$\begin{cases} x'(t) = M(t)x(t) \\ x(t_0) = x_0, \end{cases}$$

où (t_0,x_0) est dans $I\times\mathbb{R}^n$. Soit pour tout t dans I, $\Phi(t)$ la matrice fondamentale de la proposition 7.4. Alors, on vérifie sans peine que la solution du problème de Cauchy ci–dessus est :

$$x(t) = \Phi(t)x_0, \qquad t \in I.$$

Le déterminant de la matrice fondamentale jouera un rôle particulier dans la suite, c'est pourquoi on définit plus généralement :

Définition 7.7 (Wronskien)**.** *Pour toute solution*

$$\begin{aligned} \Phi : I &\longrightarrow \mathscr{M}_n(\mathbb{R}) \\ t &\longmapsto \Phi(t), \end{aligned}$$

de l'edo matricielle (7.22) *sur I, le déterminant*

$$W_\Phi(t) = \det(\Phi(t)), \qquad t \in I,$$

est appelé le Wronskien de $\Phi(t)$.

La proposition suivante fournit une formule permettant de calculer le Wronskien connaissant l'application (7.20).

Proposition 7.5. *Pour tout t et t_0 dans l'intervalle I, le Wronskien vérifie :*

$$W_\Phi(t) = W_\Phi(t_0)\exp\left(\int_{t_0}^{t} \operatorname{Tr} M(s)\,\mathrm{d}s\right). \tag{7.23}$$

Démonstration. Pour tout t dans I, notons $\varphi_j(t)$ $(j = 1,\ldots,n)$ les vecteurs colonnes de la matrice $\Phi(t)$. Comme $t \mapsto \Phi(t)$ est une solution de l'edo matricielle (7.22) sur I alors chaque fonction φ est une solution de l'edo (7.21) sur I également.

Supposons qu'il existe t_1 dans I tel que

$$W_\Phi(t_1) = \det\Phi(t_1) = 0.$$

Cela signifie qu'il existe n constantes réelles $c_1,\ldots,c_n$ telles que :

$$c_1\varphi_1(t_1) + \ldots + c_n\varphi_n(t_1) = 0.$$

Or la fonction $t \mapsto c_1\varphi_1 + \ldots + c_n\varphi_n$ est également une solution de l'edo (7.21) sur I. Comme elle s'annule en $t = t_1$, par unicité de la solution, elle est nulle pour tout t dans I. Les vecteurs colonnes de la matrice $\Phi(t)$ forment ainsi une famille liée pour tout t dans I et donc le Wronskien est nul pour tout t dans I. La formule (7.23) est bien vérifiée dans ce cas.

Supposons maintenant que :

$$W_\Phi(t) = \det\Phi(t) \neq 0 \qquad \text{pour tout } t \in I.$$

Rappelons que l'application det : $\mathrm{GL}_n(\mathbb{R}) \to \mathbb{R}$ est différentiable en tout point et pour tout $A \in \mathrm{GL}_n(\mathbb{R})$ et tout $H \in \mathscr{M}_n(\mathbb{R})$:

$$D\det(A)H = \det(A))\,\mathrm{Tr}(A^{-1}H).$$

On en déduit que pour tout t dans I :

$$\begin{aligned} W'_\Phi(t) = \frac{\mathrm{d}}{\mathrm{d}t}\det(\Phi(t)) &= \det(\Phi(t))\,\mathrm{Tr}(\Phi^{-1}(t)\Phi'(t)) \\ &= \det(\Phi(t))\,\mathrm{Tr}(\Phi^{-1}(t)M(t)\Phi(t)) \\ &= \det(\Phi(t))\,\mathrm{Tr}(M(t)) = W_\Phi(t)\,\mathrm{Tr}(M(t)). \end{aligned}$$

Soit t_0 dans I. La fonction $t \mapsto W_\Phi(t)$ est donc la solution sur I du problème de Cauchy :

$$\begin{cases} x'(t) = \mathrm{Tr}(M(t))x(t) \\ x(t_0) = W_\Phi(t_0). \end{cases}$$

Cette solution est unique et donnée par la formule (7.23). □

Exemple 7.6. Nous allons utiliser la formule de la proposition 7.5 avec le flot étudié dans la section 7.9. Rappelons brièvement le contexte :

On considère une fonction f de classe $\mathscr{C}^1$ d'un ouvert $\mathscr{U}$ de $\mathbb{R}^N$ dans $\mathbb{R}^N$ et pour tout x dans $\mathscr{U}$, on note $t \mapsto \phi(t,x)$ l'unique solution du problème de Cauchy :

$$\begin{cases} y'(t) = f(y(t)), \\ y(0) = x. \end{cases} \tag{7.24}$$

On note $J(x)$ l'intervalle maximal d'existence de cette solution et pour tout t dans $\mathbb{R}$, on définit le sous ensemble de $\mathscr{U}$:

$$\mathscr{U}_t = \{x \in \mathscr{U} \ : \ t \in J(x)\}. \tag{7.25}$$

La fonction ϕ vue comme fonction des deux variables (t,x) est définie sur l'ensemble :

$$\mathscr{V} = \{(t,x) \in \mathbb{R} \times \mathscr{U} \ : \ t \in J(x)\} = \bigcup_{x \in \mathscr{U}} \Big(J(x) \times \{x\}\Big) = \bigcup_{t \in \mathbb{R}} \Big(\{t\} \times \mathscr{U}_t\Big), \tag{7.26}$$

et le théorème 7.9 nous assure que ϕ est de classe $\mathscr{C}^1$ sur cet ensemble. Pour tout indice j dans $\{1,\ldots,N\}$ et tout (t,x) dans $\mathscr{V}$, notons $\partial\phi/\partial x_j(t,x)$ le vecteur colonne des dérivées partielles de ϕ par rapport à la variable x_j au point (t,x). La fonction $\partial\phi/\partial x_j$ est définie sur $\mathscr{V}$, à valeurs dans $\mathbb{R}^N$ et d'après la proposition 7.1, la fonction $t \mapsto \partial\phi/\partial x_j(t,x)$ est la solution du problème de Cauchy :

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\left(\dfrac{\partial\phi}{\partial x_j}\right)(t,x) = Df(\phi(t,x))\dfrac{\partial\phi}{\partial x_j}(t,x) \\ \qquad \dfrac{\partial\phi}{\partial x_j}(0,x) = e_j, \end{cases}$$

où $\{e_1,\dots,e_N\}$ désigne comme toujours la base canonique de $\mathbb{R}^N$. C'est un problème linéaire homogène et la fonction $t \mapsto Df(\phi(t,x))$ est continue sur $J(x)$ donc la solution est également définie sur $J(x)$. De nouveau d'après le théorème 7.9, pour tout t dans $\mathbb{R}$, la fonction :

$$\begin{aligned} \phi_t : \mathscr{U}_t &\longrightarrow \mathscr{U}_{-t} \\ x &\longmapsto \phi(t,x), \end{aligned} \tag{7.27}$$

est un $\mathscr{C}^1$ difféormophisme. D'après la proposition 7.5 :

$$\det D\phi_t(x) = \exp\left(\int_0^t \operatorname{Tr}\big(Df(\phi(s,x))\big)\,\mathrm{d}s\right) = \exp\left(\int_0^t \operatorname{div}\big(f(\phi(s,x))\big)\,\mathrm{d}s\right).$$

Cela nous donne le déterminant de la Jacobienne associée au changement de variables défini par le difféomorphsime (7.27). Dans la formule ci–dessus, l'opérateur divergence est défini pour toute fonction ξ de classe $\mathscr{C}^1$ d'un ouvert de $\mathbb{R}^N$ dans $\mathbb{R}^N$ par :

$$\operatorname{div}\xi = \operatorname{Tr} D\xi = \sum_{j=1}^N \frac{\partial \xi}{\partial x_j}.$$

Continuons d'énoncer les propriétés des matrices fondamentales :

Théorème 7.11. *Soit*

$$\begin{aligned} \Phi : I &\longrightarrow \mathscr{M}_n(\mathbb{R}) \\ t &\longmapsto \Phi(t), \end{aligned}$$

une solution de l'edo matricielle (7.22) *et soit t dans I. Alors la matrice $\Phi(t)$ est une matrice fondamentale si et seulement si il existe $t_0 \in I$ tel que $W_\Phi(t_0) \neq 0$. Dans ce cas $W_\Phi(t) \neq 0$, pour tout $t \in I$.*

Si pour tout t dans I, $\Phi(t)$ est une matrice fondamentale de (7.21) *et si $P \in \mathscr{M}_n(\mathbb{C})$ est inversible, alors $\Phi(t)P$ est encore une matrice fondamentale de* (7.21)*. Réciproquement, si pour tout t dans I, $\Phi_1(t)$ et $\Phi_2(t)$ sont deux matrices fondamentales de* (7.21) *alors la fonction*

$$\begin{aligned} I &\longrightarrow \mathscr{M}_n(\mathbb{R}) \\ t &\longmapsto \Phi_1(t)^{-1}\Phi_2(t) \end{aligned}$$

est une fonction constante sur I.

Démonstration. La première assertion est une conséquence immédiate de la proposition 7.5.

Soit P une matrice inversible de $\mathscr{M}_n(\mathbb{R})$ et $\Phi(t)$ une matrice fondamentale pour tout t dans I. On vérifie alors facilement que l'application $t \mapsto \Phi(t)P$ est encore une solution de l'edo matricielle (7.22) sur I. Comme d'autre part $\det(\Phi(t)P) = \det(\Phi(t))\det(P) \neq 0$ pour tout t dans I, la matrice $\Phi(t)P$ est donc encore bien une matrice fondamentale.

Considérons maintenant deux matrices fondamentales $\Phi_1(t)$ et $\Phi_2(t)$ pour tout t dans $\mathbb{R}$. Par définition, les fonctions $t \mapsto \Phi_k(t)$ $(k = 1,2)$ sont de classe $\mathscr{C}^1$ sur I car ce sont des solutions de l'edo (7.22). Rappelons d'autre part que l'application :

$$\begin{aligned} G : \mathrm{GL}_n(\mathbb{R}) &\longrightarrow \mathrm{GL}_n(\mathbb{R}) \\ A &\longmapsto A^{-1}, \end{aligned}$$

est de classe $\mathscr{C}^1$ et que, pour tout A dans $\mathrm{GL}_n(\mathbb{R})$ et tout H dans $\mathscr{M}_n(\mathbb{R})$:

$$DG(A)H = -A^{-1}HA^{-1}.$$

On en déduit que le fonction

$$\begin{aligned} F : I &\longrightarrow \mathscr{M}_n(\mathbb{R}) \\ t &\longmapsto \Phi_1(t)^{-1}\Phi_2(t), \end{aligned}$$

est elle aussi de classe $\mathscr{C}^1$ et que sa dérivée vaut, pour tout t dans I :

$$F'(t) = -\Phi_1(t)^{-1}\Phi_1'(t)\Phi_1(t)^{-1}\Phi_2(t) + \Phi_1(t)^{-1}\Phi_2'(t).$$

Comme les deux fonctions $t \mapsto \Phi_1(t)$ et $t \mapsto \Phi_2(t)$ sont des solutions de (7.22), il vient, pour tout t dans I :

$$F'(t) = -\Phi_1(t)^{-1}M(t)\Phi_2(t) + \Phi_1(t)^{-1}M(t)\Phi_2(t) = 0,$$

ce qui prouve la deuxième assertion du théorème. □

On peut remarquer que :

- Si $\Phi(t)$ est une matrice fondamentale et si P est une matrice constante inversible alors $P\Phi(t)$ n'est pas forcément une matrice fondamentale.
- Toute matrice fondamentale $\Phi(t)$ satisfait par définition l'edo (7.22), ce qui entraîne que

$$M(t) = \Phi'(t)\Phi(t)^{-1} \quad \text{pour tout } t \in I.$$

  Deux edo (7.21) différentes ne peuvent donc pas avoir la même matrice fondamentale.

7.10.3 Équations linéaires avec second membre

On considère cette fois deux fonctions <u>continues</u> :

$$\begin{aligned} M : I &\longrightarrow \mathscr{M}_n(\mathbb{R}) \\ t &\longmapsto M(t) \end{aligned} \qquad \text{et} \qquad \begin{aligned} b : I &\longrightarrow \mathbb{R}^n \\ t &\longmapsto b(t). \end{aligned} \tag{7.28}$$

Si la fonction b n'est pas identiquement nulle, alors l'edo suivante est appellée edo linéaire avec second membre :

$$x'(t) = M(t)x(t) + b(t), \qquad t \in I. \tag{7.29}$$

A ce stade, on pourrait appliquer le Théorème de Cauchy–Lipschitz pour montrer l'existence et l'unicité d'une solution à (7.29). Nous allons procéder différement et nous ramener à des résultats déjà démontrés pour l'edo homogène (7.21).

Théorème 7.12. *Soit $(t_0,x_0) \in I \times \mathbb{R}^n$ et pour tout t dans I, on note $\Phi(t)$ la matrice fondamentale de (7.21) vérifiant $\Phi(t_0) = \mathrm{Id}_n$ (dont l'existence est assurée par la proposition 7.4). Alors la fonction φ définie sur I par :*

$$\varphi(t) = \Phi(t)x_0 + \Phi(t)\int_{t_0}^{t} \Phi(s)^{-1}b(s)\,\mathrm{d}s, \qquad t \in I,$$

est l'unique solution de l'edo (7.29) qui vaut x_0 en t_0.

Démonstration. Il suffit de dériver l'expression de φ ci–dessus pour vérifier que c'est bien une solution de (7.29). On montre l'unicité de façon classique : Considérons φ_1 et φ_2 deux solutions telles que $\varphi_1(t_0) = \varphi_2(t_0) = x_0$. Alors $\varphi_1 - \varphi_2$ est une solution de (7.21) vérifiant $\varphi_1(t_0) - \varphi_2(t_0) = 0$ et par unicité de la solution pour (7.21), $\varphi_1 - \varphi_2 = 0$ sur I. □

Méthode de la variation de la constante

L'expression de φ donnée dans ce théorème peut être retrouvée facilement en cherchant la solution φ sous la forme

$$\varphi(t) = \Phi(t)\xi(t),$$

où $\xi : I \longmapsto \mathbb{R}^n$ est une fonction inconnue vérifiant $\xi(t_0) = x_0$ (rappelons que $\Phi(t_0) = \mathrm{Id}_n$). En dérivant φ on obtient

$$\varphi'(t) = \Phi'(t)\xi(t) + \Phi(t)\xi'(t),$$

et comme $\Phi(t)$ vérifie l'edo matricielle (7.22), il vient

$$\varphi'(t) = M(t)\Phi(t)\xi(t) + \Phi(t)\xi'(t).$$

Pour que $\varphi(t)$ soit une solution de (7.29), il faut et il suffit donc que $\Phi(t)\xi'(t) = b(t)$ pour tout t dans I, c'est à dire que

$$\xi'(t) = \Phi(t)^{-1}b(t).$$

En intégrant cette relation entre t_0 et t, on obtient que

$$\xi(t) = x_0 + \int_{t_0}^{t} \Phi(s)^{-1}b(s)\,\mathrm{d}s.$$

Cette méthode s'appelle la *méthode de la variation de la constante.*

Si l'on dispose d'une solution particulière à l'edo avec second membre (7.29), on pourra utiliser la proposition suivante :

Proposition 7.6. *Soit $t_0 \in I$. Pour tout t dans I, on note $\Phi(t)$ la matrice fondamentale de l'edo* (7.21) *vérifiant $\Phi(t_0) = \mathrm{Id}_n$. Si l'on dispose déjà d'une solution particulière φ^* à l'edo* (7.29) *sur I, alors pour tout $x_0 \in \mathbb{R}^n$, la fonction $\varphi : I \to \mathbb{R}^n$ définie par :*

$$\varphi(t) = \varphi^*(t) + \Phi(t)(x_0 - \varphi^*(t_0)) \qquad \textit{pour tout } t \in I,$$

est la solution de l'edo (7.29) *qui vérifie $\varphi(t_0) = x_0$.*

Démonstration. Il suffit de dériver φ pour vérifier que c'est bien la solution cherchée. □

7.10.4 Équations linéaires à cœfficients constants

Dans ce paragraphe, on étudiera le cas particulier où la matrice M dans l'équation (7.21) (edo linéaire homogène) ou dans l'équation (7.29) (edo linéaire avec second membre) ne dépend pas de t (c'est à dire que l'edo est linéaire autonome). Ainsi, on va considérer les edo :

$$x'(t) = Mx(t), \tag{7.30a}$$

$$x'(t) = Mx(t) + b(t), \tag{7.30b}$$

où la fonction

$$\begin{aligned} b : \mathbb{R} &\longrightarrow \mathbb{R}^n \\ t &\longmapsto b(t), \end{aligned}$$

est continue. La Proposition suivante rassemble les premiers résultats sur les edo (7.30) qui se déduisent directement de l'étude des edo plus générales (7.21) et (7.29) :

Proposition 7.7. *Pour toute matrice $M \in \mathscr{M}_n(\mathbb{R})$ et pour tout couple (t_0,x_0) dans $\mathbb{R} \times \mathbb{R}^n$:*

- *Il existe une unique solution maximale φ à l'edo* (7.30a) *définie sur $\mathbb{R}$ tout entier et qui vérifie $\varphi(t_0) = x_0$.*
- *L'expression de la matrice fondamentale $\Phi(t)$ ($t \in \mathbb{R}$) associée à l'edo* (7.30a) *et vérifiant $\Phi(t_0) = \mathrm{Id}_n$ (cf. proposition 7.4) est*

$$\Phi(t) = e^{(t-t_0)M} \qquad \textit{pour tout } t \in \mathbb{R}.$$

- *Il existe une unique solution maximale φ à l'edo* (7.30b) *vérifiant $\varphi(t_0) = x_0$. Cette solution est définie sur $\mathbb{R}$ tout entier par*

$$\varphi(t) = e^{(t-t_0)M}x_0 + \int_{t_0}^{t} e^{(t-s)M} b(s)\,\mathrm{d}s \qquad \textit{pour tout } t \in \mathbb{R}. \tag{7.31}$$

Traitons un exemple simple ($N = 2$) pour illustrer tout cela :

Exemple 7.7. Soit l'edo linéaire homogène autonome :

$$x'(t) = Mx(t), \qquad (t \in \mathbb{R}),$$

où la matrice M est définie par :

$$M = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Afin de déterminer, pour tout t dans $\mathbb{R}$, la matrice fondamentale $\Phi(t) = e^{tM}$, il faut diagonaliser (si c'est possible) la matrice M (comme expliqué dans la proposition 7.3). On calcule le polynôme caractéristique :

$$P(\lambda) = \lambda^2 + 1 = (\lambda - i)(\lambda + i).$$

La matrice M admet donc deux valeurs propres distinctes (complexes conjuguées) $\lambda_1 = i$ et $\lambda_2 = -i$ et comme $N = 2$, on en déduit que M est bien diagonalisable. On détermine les sous espaces propres en résolvant :

$$(M - \lambda_1 \mathrm{Id}_2)x = \begin{pmatrix} -i & 1 \\ -1 & -i \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Les solutions forment un sous–ev de dimension 1 engendré par $u_1 = (1/\sqrt{2}, i/\sqrt{2})$. De même les solutions de

$$(M - \lambda_2 \mathrm{Id}_2)x = \begin{pmatrix} i & 1 \\ -1 & i \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

forment un sous–ev de dimension 1 engendré par $u_2 = (1/\sqrt{2}, -i/\sqrt{2})$. Remarquer que l'on a choisi de normer les vecteurs propres u_1 et u_2 (pour la norme euclidienne classique). En notant

$$P = \begin{pmatrix} 1/\sqrt{2} & -i/\sqrt{2} \\ i/\sqrt{2} & -i/\sqrt{2} \end{pmatrix} \qquad \text{et} \qquad D = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix},$$

on a la relation $P^{-1}MP = D$. La matrice D étant diagonale

$$e^{tD} = \begin{pmatrix} e^{it} & 0 \\ 0 & e^{-it} \end{pmatrix},$$

et en appliquant ensuite la formule $Pe^{tD}P^{-1} = e^{tPDP^{-1}} = e^{tM}$ (voir la proposition 7.3), on obtient :

$$\begin{aligned} \Phi(t) = Pe^{tD}P^{-1} &= \begin{pmatrix} 1/\sqrt{2} & -i/\sqrt{2} \\ i/\sqrt{2} & -i/\sqrt{2} \end{pmatrix} \begin{pmatrix} e^{it} & 0 \\ 0 & e^{-it} \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & -i/\sqrt{2} \\ 1/\sqrt{2} & i/\sqrt{2} \end{pmatrix} \\ &= \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix}. \end{aligned}$$

Pour tout couple $(t_0,x_0) \in \mathbb{R} \times \mathbb{R}^2$ (en notant $x_0 = (x_0^1,x_0^2)$), la solution φ de l'edo vérifiant $\varphi(t_0) = x_0$ est donc :

$$\varphi(t) = \begin{pmatrix} \cos(t-t_0) & \sin(t-t_0) \\ -\sin(t-t_0) & \cos(t-t_0) \end{pmatrix} \begin{pmatrix} x_0^1 \\ x_0^2 \end{pmatrix}.$$

On remarque qu'elle est périodique de période 2π. Quelques solutions sont représentées sur la figure 7.7.

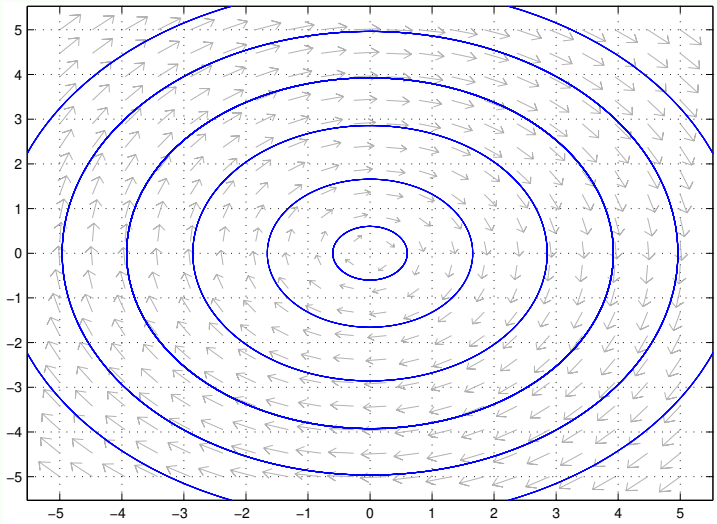


FIG. 7.7 – *Quelques trajectoires de l'edo $x'(t) = Mx(t)$*

On souhaite maintenant résoudre l'équation avec second membre :

$$x'(t) = Mx(t) + b \qquad \text{pour tout } t \in \mathbb{R}, \tag{7.32}$$

avec b un vecteur constant, par exemple :

$$b = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Une première méthode consiste à chercher une solution particulière "à la main". Ainsi, en cherchant une solution constante, on est amené à résoudre

$$M\varphi^* = -b \quad \text{c'est à dire} \quad \varphi^* = -M^{-1}b = \begin{pmatrix} 2 \\ -1 \end{pmatrix}.$$

L'unique solution φ de l'edo (7.32) vérifiant $\varphi(t_0) = x_0$ est donc obtenue en appliquant la Proposition 7.6, c'est à dire, pour tout t dans $\mathbb{R}$:

$$\varphi(t) = \begin{pmatrix} \cos(t-t_0) & \sin(t-t_0) \\ -\sin(t-t_0) & \cos(t-t_0) \end{pmatrix} \begin{pmatrix} x_0^1 - 2 \\ x_0^2 + 1 \end{pmatrix} + \begin{pmatrix} 2 \\ -1 \end{pmatrix}.$$

Une autre approche consiste à appliquer la méthode de la variation de la constante qui conduit à l'expression :

$$\varphi(t) = e^{(t-t_0)M}\xi + \int_{t_0}^{t} e^{(t-s)M} b \, \mathrm{d}s.$$

Comme b est un vecteur constant, on peut le sortir de l'intégrale. On doit donc calculer :

$$\begin{aligned} \int_{t_0}^{t} e^{(t-s)M} \, \mathrm{d}s &= \int_{t_0}^{t} \begin{pmatrix} \cos(t-s) & \sin(t-s) \\ -\sin(t-s) & \cos(t-s) \end{pmatrix} \mathrm{d}s \\ &= \begin{pmatrix} \sin(t-t_0) & -\cos(t-t_0)+1 \\ \cos(t-t_0)-1 & \sin(t-t_0) \end{pmatrix}. \end{aligned}$$

Finalement, la solution s'écrit :

$$\varphi(t) = \begin{pmatrix} \cos(t-t_0) & \sin(t-t_0) \\ -\sin(t-t_0) & \cos(t-t_0) \end{pmatrix} \begin{pmatrix} x_0^1 \\ x_0^2 \end{pmatrix} + \begin{pmatrix} \sin(t-t_0) & -\cos(t-t_0)+1 \\ \cos(t-t_0)-1 & \sin(t-t_0) \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix},$$

et on vérifie quelle coïncide bien avec celle trouvée précédement.

7.10.5 Équations linéaires d'ordre n à cœfficients constants

Soient $a_0, a_1, \ldots, a_{n-1}$, n réels dont au moins un est différent de 0. Considérons l'edo linéaire d'ordre n :

$$x^{(n)}(t) + a_{n-1}x^{(n-1)}(t) + \ldots + a_1 x'(t) + a_0 x(t) = 0. \tag{7.33}$$

Comme expliqué dans la section 7.2, cette équation peut se mettre sous la forme résolue :

$$X'(t) = AX(t),$$

où

$$X(t) = \begin{pmatrix} x(t) \\ x'(t) \\ \vdots \\ x^{(n-2)}(t) \\ x^{(n-1)}(t) \end{pmatrix} \qquad \text{et} \qquad A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & 1 \\ -a_0 & -a_1 & \cdots & -a_{n-2} & -a_{n-1} \end{pmatrix}.$$

On commence par établir :

Lemme 7.4. *Le polynôme caractéristique de la matrice A est :*

$$P_A(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + \ldots + a_1\lambda + a_0.$$

(La matrice A est la matrice compagnon de son polynôme caractéristique).

Démonstration. D'après la définition, on doit calculer pour tout λ dans $\mathbb{R}$:

$$P_A(\lambda) = \det(\lambda \mathrm{Id}_n - A) = \begin{vmatrix} \lambda & -1 & 0 & \cdots & 0 \\ 0 & \lambda & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda & -1 \\ a_0 & a_1 & \cdots & a_{n-2} & \lambda + a_{n-1} \end{vmatrix}. \tag{7.34}$$

Notons L_1,$L_2, \ldots$,L_n les lignes de la matrice apparaissant à droite dans l'expression ci–dessus. En supposant dans un premier temps que $\lambda \neq 0$, on peut remplacer la dernière ligne par

$$L_n - \left[\frac{a_0}{\lambda} L_1 + \frac{1}{\lambda}\left(a_1 + \frac{a_0}{\lambda}\right) L_2 + \cdots + \frac{1}{\lambda}\left(a_{n-2} + \cdots + \frac{a_0}{\lambda^{n-2}}\right) L_{n-1}\right],$$

sans changer la valeur du déterminant. Le déterminant à calculer est alors :

$$\begin{vmatrix} \lambda & -1 & 0 & \cdots & 0 \\ 0 & \lambda & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \lambda & -1 \\ 0 & \cdots & \cdots & 0 & Q(\lambda) \end{vmatrix},$$

où $Q(\lambda) = \lambda + a_{n-1} + \lambda^{-1}a_{n-2} + \ldots + \lambda^{-(n-1)}a_0$. La matrice étant maintenant diagonale, on obtient facilement le résultat annoncé.

Si $\lambda = 0$ dans l'égalité (7.34), on calcule le déterminant en développant par rapport à la première colonne et on trouve bien que $P_A(0) = a_0$. □

Le corps $\mathbb{C}$ étant algébriquement clos, le polynôme caractéristique de A se factorise en :

$$P_A(\lambda) = (\lambda - \lambda_1)^{d_1} \cdots (\lambda - \lambda_m)^{d_m},$$

où les racines $\lambda_1, \ldots, \lambda_m$ sont les valeurs propres (complexes) de A et où les entiers d_k $(k = 1, \ldots, m)$ vérifient $d_1 + \cdots + d_m = n$. Le Théorème suivant donne explicitement les solutions de l'edo (7.33) en fonction des valeurs propres λ_k et des entiers d_k $(k = 1, \ldots, m)$.

Théorème 7.13. *L'edo linéaire d'ordre n (7.33) admet n solutions linéairement indépendantes. Les expressions des solutions se déduisent des valeurs propres λ_j de la matrice A et des indices d_j comme suit (pour tout $j = 1,\ldots,m$) :*

- *Si la valeur propre λ_j est réelle alors pour tout $k_j = 0,1,\ldots,d_j - 1$, les fonctions :*

$$\begin{aligned} \mathbb{R} &\longrightarrow \mathbb{R} \\ t &\longmapsto t^{k_j} e^{\lambda_j t}, \end{aligned} \tag{7.35a}$$

  sont des solutions à l'edo (7.33).
- *Si la valeur propre λ_j est complexe alors $\bar{\lambda}_j$ est aussi une valeur propre de A et pour tout $k_j = 0,1,\ldots,d_j - 1$ les fonctions :*

$$\begin{aligned} \mathbb{R} &\longrightarrow \mathbb{R} \\ t &\longmapsto t^{k_j} e^{\mathrm{Re}(\lambda_j)t} \cos\big(\mathrm{Im}(\lambda_j)t\big) \end{aligned} \qquad \textit{et} \qquad \begin{aligned} \mathbb{R} &\longrightarrow \mathbb{R} \\ t &\longmapsto t^{k_j} e^{\mathrm{Re}(\lambda_j)t} \sin\big(\mathrm{Im}(\lambda_j)t\big), \end{aligned} \tag{7.35b}$$

  sont des solutions à l'edo (7.33).

Les fonctions (7.35) forment une base de l'espace des solutions de l'edo (7.33).

Démonstration. Notons L l'opérateur différentiel associé à l'edo (7.33), c'est à dire que pour toute fonction φ, n fois dérivable sur $\mathbb{R}$:

$$L\varphi(t) = \varphi^{(n)}(t) + a_{n-1}\varphi^{(n-1)}(t) + \ldots + a_1\varphi'(t) + a_0\varphi(t), \qquad (t \in \mathbb{R}).$$

Introduisons ensuite la fonction :

$$\begin{aligned} \phi : \mathbb{R} \times \mathbb{C} &\longrightarrow \mathbb{C} \\ (t,\lambda) &\longmapsto e^{\lambda t}, \end{aligned}$$

ainsi que $\xi(t,\lambda) = L\phi(t,\lambda) = P_A(\lambda)e^{\lambda t}$ (pour tout $(t,\lambda) \in \mathbb{R} \times \mathbb{C}$). D'autre part, pour tout entier k posons :

$$\psi_k(t,\lambda) = \frac{\partial^k \phi}{\partial \lambda^k}(t,\lambda) = t^k e^{\lambda t} \qquad (t \in \mathbb{R}, \lambda \in \mathbb{C}).$$

On a alors, d'après le lemme de Schwarz (toutes les fonctions sont de classe $\mathscr{C}^\infty$ ici) :

$$L\psi_k(t,\lambda) = L\left(\frac{\partial^k \phi}{\partial \lambda^k}\right)(t,\lambda) = \frac{\partial^k (L\phi)}{\partial \lambda^k}(t,\lambda) = \frac{\partial^k \xi}{\partial \lambda^k}(t,\lambda). \tag{7.36a}$$

En appliquant la formule de Newton pour dériver un produit de fonctions, nous obtenons:

$$\frac{\partial^k \xi}{\partial \lambda^k}(t,\lambda) = \left[\sum_{p=0}^{k} \binom{k}{p} P_A^{(p)}(\lambda) t^{k-p}\right] e^{\lambda t} \qquad (t \in \mathbb{R}, \lambda \in \mathbb{C}). \tag{7.36b}$$

Dire que λ est une racine d'ordre d du polynôme P_A signifie que

$$P_A(\lambda) = P_A'(\lambda) = \cdots = P_A^{(d-1)}(\lambda) = 0,$$

ce qui entraîne avec les formules (7.36) ci–dessus que pour tout $j = 1,\ldots,m$ et tout $k_j = 0,1,\ldots,d_j - 1$, les fonctions

$$\begin{aligned} \mathbb{R} &\longrightarrow \mathbb{C} \\ t &\longmapsto t^{k_j} e^{\lambda_j t}, \end{aligned}$$

sont bien des solutions (à valeurs complexes a priori) de l'edo (7.33). Si λ_j est réelle, on retrouve l'expression (7.35a) dans l'énoncé du théorème. Si λ_j est dans $\mathbb{C}$ et pas dans $\mathbb{R}$ alors, comme indiqué plus haut, il existe un indice ℓ tel que $\lambda_\ell = \bar{\lambda}_j$ et $d_\ell = d_j$ (car les cœfficients du polynôme P_A sont réels). Pour tout $k_j = 0,1,\ldots,d_j - 1$, par combinaison linéaire des fonctions

$$\begin{array}{l} \mathbb{R} \longrightarrow \mathbb{C} \\ t \longmapsto t^{k_j} e^{\lambda_j t} \end{array} \qquad \text{et} \qquad \begin{array}{l} \mathbb{R} \longrightarrow \mathbb{C} \\ t \longmapsto t^{k_j} e^{\bar{\lambda}_j t}, \end{array}$$

on retrouve les expressions des solutions réelles (7.35b) de l'énoncé.

Notons maintenant $\mathscr{F}$ la famille dans $\mathscr{C}^1(\mathbb{R})$ composée des n fonctions (à valeurs complexes) :

$$\begin{array}{l} \mathbb{R} \longrightarrow \mathbb{C} \\ t \longmapsto t^{k_j} e^{\lambda_j t}, \end{array} \tag{7.37}$$

pour tout $j = 1,\ldots,m$ et $k_j = 0,1,\ldots,d_j - 1$ (on rappelle qu'en effet $d_1 + \cdots + d_m = n$) et vérifions que $\mathscr{F}$ est une famille libre. Classiquement, on suppose donc qu'il existe n nombres complexes c_{jk} (pour tous les couples d'indices (j,k) tels que $1 \leqslant j \leqslant m$ et $1 \leqslant k \leqslant d_j$) vérifiant :

$$\sum_{j=1}^{m} \sum_{k=1}^{d_j - 1} c_{jk} t^k e^{\lambda_j t} = 0 \qquad \text{pour tout } t \in \mathbb{R}.$$

Si les nombres complexes c_{jk} ne sont pas tous égaux à 0, cette égalité se réécrit pour tout t dans $\mathbb{R}$ sous la forme :

$$\sum_{j=1}^{\sigma} Q_j(t) e^{\lambda_j t} = 0 \qquad \text{où l'on a posé pour tout } j = 1,\ldots,m \qquad Q_j(t) = \sum_{k=1}^{d_j - 1} c_{jk} t^k,$$

et où $\sigma \geqslant 1$ est le plus grand indice (inférieur ou égal à m) tel que le polynôme Q_σ n'est pas identiquement nul ($\sigma = 0$ signifierait que tous les polynômes Q_j sont nuls et donc que tous les cœfficients c_{jk} sont nuls). En multipliant l'égalité ci–dessus par $e^{-\lambda_1 t}$, on obtient :

$$Q_1(t) + \sum_{j=2}^{\sigma} Q_j(t) e^{(\lambda_j - \lambda_1)t} = 0 \qquad \text{pour tout } t \in \mathbb{R}.$$

Comme $\deg(Q_1) \leqslant d_1 - 1$, en dérivant d_1 fois par rapport à t cette égalité devient :

$$\sum_{i=2}^{\sigma} R_j(t) e^{(\lambda_j - \lambda_1)t} = \sum_{i=2}^{\sigma} R_j(t) e^{\lambda_j t} = 0 \qquad \text{pour tout } t \in \mathbb{R},$$

où chaque R_j est un polynôme vérifiant $\deg(R_j) = \deg(Q_j)$ pour tout $j = 2,\ldots,\sigma$ (utiliser la formule de Newton pour dériver un produit). Réitérant cette opération, on finit par obtenir un polynôme identiquement nul et du même degré que Q_σ (dont on sait pas définition de l'indice σ qu'il est non nul). Les nombres c_{jk} sont donc nécessairement tous nuls et la famille $\mathscr{F}$ formée par les n fonctions (7.37) est bien une famille libre dans $\mathscr{C}^1(\mathbb{R})$.

Toutes les fonctions (7.35) sont des combinaisons linéaires des fonctions de la famille $\mathscr{F}$. Réciproquement, toute fonction de la famille $\mathscr{F}$ s'obtient comme combinaison linéaire de fonctions (7.35). Ceci prouve que les fonctions (7.35) génèrent un ev de dimension n et donc qu'elles forment bien une base de l'espace des solutions de l'edo (7.33). $\square$

Exemple 7.8. Soit l'edo linéaire homogène d'ordre 4 :

$$x^{(4)}(t) - x''(t) + 2x'(t) + 2 = 0.$$

D'après les résultats établis plus haut, on est amené à factoriser le polynôme :

$$P(\lambda) = \lambda^4 - \lambda^2 + 2\lambda + 2,$$

ce qui donne :

$$P(\lambda) = (\lambda - (1+i))(\lambda - (1-i))(\lambda+1)^2.$$

D'après le théorème 7.13, l'espace vectoriel des solutions de l'edo est de dimension 4 et il est engendré par les fonctions suivantes :

$$\begin{array}{l} \mathbb{R} \longrightarrow \mathbb{R} \\ t \longmapsto e^t \sin(t), \end{array} \qquad \begin{array}{l} \mathbb{R} \longrightarrow \mathbb{R} \\ t \longmapsto e^t \cos(t), \end{array} \qquad \begin{array}{l} \mathbb{R} \longrightarrow \mathbb{R} \\ t \longmapsto e^{-t}, \end{array} \quad \text{et} \quad \begin{array}{l} \mathbb{R} \longrightarrow \mathbb{R} \\ t \longmapsto te^{-t}. \end{array}$$

Terminons ce chapitre par l'étude d'un exemple qui utilise une partie des résultats que nous venons d'établir. Le but de ce paragraphe est de montrer comment il est possible de déterminer les propriétés de la solution d'une edo sans que cette solution puisse être calculée explicitement.

7.10.6 L'équation de Sturm–Liouville

Soit I un intervalle ouvert dans $\mathbb{R}$ et :

$$\begin{array}{l} a_0 : I \longrightarrow \mathbb{R} \\ \quad t \longmapsto a_0(t) \end{array} \quad \text{et} \quad \begin{array}{l} a_0 : I \longrightarrow \mathbb{R} \\ \quad t \longmapsto a_0(t), \end{array}$$

deux fonctions continues. On souhaite étudier la position des zéros des solutions de l'edo linéaire homogène d'odre 2 :

$$x''(t) + a_1(t)x'(t) + a_0(t)x(t) = 0, \tag{7.38}$$

appelée équation de Sturm–Liouville. En posant :

$$X(t) = \begin{pmatrix} x(t) \\ x'(t) \end{pmatrix} \qquad \text{pour tout } t \in I,$$

cette équation se met sous forme résolue :

$$X'(t) = \begin{pmatrix} 0 & 1 \\ -a_0(t) & -a_1(t) \end{pmatrix} X(t). \tag{7.39}$$

D'après le théorème 7.10, l'ensemble des solutions est un espace vectoriel de dimension 2. On peut donc considérer deux solutions linéairement indépendantes de l'edo (7.38) :

$$\begin{array}{l} \varphi_1 : I \longrightarrow \mathbb{R} \\ \quad t \longmapsto \varphi_1(t) \end{array} \quad \text{et} \quad \begin{array}{l} \varphi_2 : I \longrightarrow \mathbb{R} \\ \quad t \longmapsto \varphi_2(t), \end{array}$$

qui correspondent à deux solutions linéairement indépendantes de l'edo (7.39) :

$$\begin{array}{l} X_1 : I \longrightarrow \mathbb{R}^2 \\ \quad t \longmapsto \begin{pmatrix} \varphi_1(t) \\ \varphi_1'(t) \end{pmatrix} \end{array} \quad \text{et} \quad \begin{array}{l} X_2 : I \longrightarrow \mathbb{R}^2 \\ \quad t \longmapsto \begin{pmatrix} \varphi_2(t) \\ \varphi_2'(t) \end{pmatrix} \end{array}$$

de sorte qu'en posant, toujours pour tout t dans I :

$$\Phi(t) = \begin{pmatrix} \varphi_1(t) & \varphi_2(t) \\ \varphi_1'(t) & \varphi_2'(t) \end{pmatrix},$$

on définit une matrice fondamentale pour l'edo (7.39).

Théorème 7.14. *Soient φ_1 et φ_2 deux solutions linéairement indépendantes de l'edo* (7.38) *sur I. Alors, entre deux zéros consécutifs de φ_1 il existe exactement un zéro de φ_2.*

Démonstration. Soient u et v deux zéros consécutifs de φ_1 dans I (c'est à dire que $\varphi_1(u) = \varphi_1(v) = 0$ et $\varphi_1(t) \neq 0$ pour tout $t \in]u,v[$). Le Wronskien de la matrice fondamentale $\Phi(t)$ est, pour tout t dans I :

$$W_\Phi(t) = \det \Phi(t) = \varphi_1(t)\varphi_2'(t) - \varphi_2(t)\varphi_1'(t).$$

D'après le théorème 7.11, le Wronskien d'une matrice fondamentale ne s'annule jamais donc $W_\Phi(u) \neq 0$ et $W_\Phi(v) \neq 0$, ce qui entraîne que $\varphi_2(u) \neq 0$ et $\varphi_2(v) \neq 0$. Supposons que φ_2 ne s'annule pas sur $]u,v[$, alors la fonction $f = \varphi_1/\varphi_2$ est bien définie sur cet intervalle et $f(u) = f(v) = 0$. D'après le théorème des accroissements finis, il existerait un point $c \in]u,v[$ tel que $f'(c) = 0 = -W_\Phi(c)/\varphi_2^2(c)$ ce qui n'est pas possible (le Wronskien ne peut pas s'annuler sur I). Il existe donc au moins un zéro de φ_2 entre u et v. S'il en existait deux (notons les α et β), alors par symétrie des rôles joués par les fonctions φ_1 et φ_2, en suivant le même raisonnement que ci–dessus on montrerait l'existence d'un autre zéro de φ_1 entre α et β. Or ceci n'est pas possible par hypothèse. □

Exemple 7.9. Pour illustrer le théorème 7.14 examinons le cas où les fonctions a_0 et a_1 sont des fonctions constantes. Nous disposons alors du théorème 7.13 qui nous permet de calculer explicitement les solutions. Considérons ainsi l'edo :

$$x''(t) + a_1 x'(t) + a_0 x(t) = 0.$$

Les solutions s'obtiennent à partir des racines du polynôme :

$$P(\lambda) = \lambda^2 + a_1\lambda + a_1.$$

On est amené à distinguer les cas suivants :

- Si le polynôme P a deux racines réelles λ_1 et λ_2. Dans ce cas les solutions définies sur $\mathbb{R}$ forment un ev de dimension 2 engendré par $\varphi_1 : t \mapsto e^{\lambda_1 t}$ et $\varphi_2 : t \mapsto e^{\lambda_2 t}$. Aucune de ces fonctions ne s'annule sur $\mathbb{R}$.
- Si le polynôme P a une racine double (réelle) λ alors les solutions sur $\mathbb{R}$ sont engendrées par $\varphi_1 : t \mapsto e^{\lambda t}$ et $\varphi_2 : t \mapsto te^{\lambda t}$. La solution φ_1 ne s'annule pas et la solution φ_2 s'annule une seule fois.
- Si le polynôme P a deux racines complexes conjuguées λ et $\bar{\lambda}$. L'espace des solutions est engendré par $\varphi_1 : t \mapsto e^{\mathrm{Re}(\lambda)t} \cos\big(\mathrm{Im}(\lambda)t\big)$ et $\varphi_2 : t \mapsto e^{\mathrm{Re}(\lambda)t} \sin\big(\mathrm{Im}(\lambda)t\big)$. On retrouve ce qui est annoncé dans le théorème 7.14 : Entre deux zéros de φ_1 il existe exactement un zéro de φ_2.

On suppose maintenant que a_1 n'est plus seulement continue mais de classe $\mathscr{C}^1$ sur I (la fonction a_0 étant toujours supposée continue). Soient $(t_0,x_0,x_1) \in I \times \mathbb{R} \times \mathbb{R}$ et φ la solution sur I du problème de Cauchy :

$$\begin{cases} x''(t) + a_1(t)x'(t) + a_0(t)y(t) = 0 \\ \qquad (x(t_0),x'(t_0)) = (x_0,x_1). \end{cases} \tag{7.40}$$

dans lequel on reconnait une edo de Sturm–Liouville (7.38). On vérifie par un calcul direct que la fonction ϕ définie sur I par

$$\phi(t) = \varphi(t) \exp\left(\frac{1}{2}\int_{t_0}^{t} a_1(s)\,\mathrm{d}s\right) \qquad \text{pour tout } t \in I, \tag{7.41}$$

est l'unique solution du problème de Cauchy :

$$\begin{cases} y''(t) + q(t)y(t) = 0 \\ \quad (y(t_0),y'(t_0)) = (x_0,x_1 + a_1(t_0)x_0/2), \end{cases} \tag{7.42}$$

où la fonction réelle q est définie pour tout t dans I par :

$$q(t) = -\frac{1}{2}a_1'(t) - \frac{1}{4}a_1^2(t) + a_0(t).$$

La fonction q est continue et l'edo apparaissant dans la système (7.42) est appelée équation différentielle de Sturm–Liouville réduite. Si l'on connait les solutions du problème de Cauchy (7.42) on peut donc facilement en déduire l'expression des solutions du problème de Cauchy (7.40) grâce à l'identité (7.41). Remarquer de plus que les fonctions φ et ϕ ont les mêmes zéros.

Théorème 7.15. *Soit $[a,b] \subset I$ un intervalle sur lequel $q(t) \leqslant 0$ (avec $a < b$). Alors toute solution non identiquement nulle de l'équation de Sturm–Liouville réduite (apparaissant dans le système* (7.42)*) a au plus un zéro dans $[a,b]$.*

Démonstration. Soit ϕ une solution non identiquement nulle de l'équation de Sturm–Liouville réduite (celle du système (7.42)) et soit θ un zéro de ϕ dans $[a,b]$.

Supposons que $\theta \neq b$. Comme ϕ n'est pas identiquement nulle, $\phi'(\theta) \neq 0$ (en effet, la seule solution de (7.42) vérifiant les conditions de Cauchy $\phi(\theta) = \phi'(\theta) = 0$ est la solution identiquement nulle). Supposons par exemple que $\phi'(\theta) > 0$. Par continuité de ϕ', il existe $\delta > 0$ (vérifiant $\theta+\delta < b$) tel que sur $]\theta,\theta + \delta[$ on a encore $\phi'(t) > 0$. Posons alors :

$$J = \big\{c \in]\theta,b] \text{ tels que } \phi'(t) > 0 \text{ sur }]\theta,c[\big\}.$$

Nous venons de montrer que $\theta+\delta \in J$ et donc J est non vide. Comme J est par ailleurs majoré par b, il admet une borne supérieure que nous noterons b^* et nous allons vérifier que $b^* = b$. En effet, si $b^* < b$ alors par continuité de ϕ' sur I, $\phi'(b^*) = 0$ (si $\phi'(b^*)$ était strictement positif, par continuité de ϕ' il serait encore strictement positif au delà de b^* ce qui n'est pas possible). Or $\phi(\theta) = 0$ et $\phi'(t) > 0$ sur $]\theta,b^*[$ donc $\phi(t) > \phi(\theta) > 0$ sur $]\theta,b^*[$. D'après l'edo de Sturm–Liouville réduite (7.42), $\phi''(t) = -q(t)\phi(t) \geqslant 0$ sur $]\theta,b^*[$. Ceci entraîne que la fonction ϕ' est croissante sur $]\theta,b^*[$ et donc $\phi'(b^*) \geqslant \phi'(\theta) > 0$ ce qui contredit la maximalité de b^*. Notre supposition $b^* < b$ conduit à une contraction et ainsi, nécessairement $b = b^*$.

Ainsi on a montré que $\phi'(t) > 0$ sur $]\theta,b[$, ϕ est donc strictement croissante sur cet interval et donc ne s'annule pas sur $]\theta,b]$ (si on avait supposé $\phi'(\theta) < 0$, on aurait obtenu que ϕ était strictement décroissante).

Supposons maintenant que $\theta \neq a$ et que $\phi'(\theta) > 0$. Dans ce cas, on pose :

$$J = \big\{c \in [a,\theta[\text{ tels que } \phi'(t) > 0 \text{ sur }]c,\theta[\big\}.$$

On montrerait alors que a est l'infimum de J, donc que ϕ est strictement croissante sur $]a,\theta]$ et donc que ϕ ne peut pas s'annuler sur $[a,\theta[$.

En conclusion, si $\theta = a$ alors ϕ ne peut pas s'annuler que $]a,b]$, si $\theta = b$ alors ϕ ne peut pas s'annuler sur $[a,b[$ et si $\theta \neq a$ et $\theta \neq b$ alors ϕ ne peut s'annuler ni sur $[a,\theta[$ ni sur $]\theta,b]$. Le théorème est bien démontré. □

Le Théorème suivant permet de comparer la position des zéros pour deux solutions d'edo de Sturm–Liouville sous forme réduite.

Théorème 7.16. *Soient φ_1 et φ_2 deux solutions sur I respectivement de*

$$x''(t) + q_1(t)x(t) = 0, \tag{7.43a}$$

$$x''(t) + q_2(t)x(t) = 0. \tag{7.43b}$$

On suppose que $q_1(t) \geqslant q_2(t)$ sur I. Alors, entre deux zéros consécutifs de φ_2 il existe au moins un zéro de φ_1.

Démonstration. Soient u et v deux zéros consécutifs de φ_2 sur I. Supposons par exemple que $\varphi_2(t) > 0$ sur $]u,v[$. Cela entraîne que $\varphi_2'(u) \geqslant 0$ et $\varphi_2'(v) \leqslant 0$. Supposons également que φ_1 ne s'annule pas sur $]u,v[$, par exemple $\varphi_1(t) > 0$. Introduisons alors, pour tout t dans $[u,v]$:

$$W(\varphi_1,\varphi_2)(t) = \varphi_1(t)\varphi_2'(t) - \varphi_2(t)\varphi_1'(t).$$

Remarquer que W n'est pas un Wronskien ici puisque φ_1 et φ_2 sont des solutions d'edo différentes. On déduit des hypothèses que

$$W(\varphi_1,\varphi_2)(u) = \varphi_1(u)\varphi_2'(u) \geqslant 0 \qquad \text{et} \qquad W(\varphi_1,\varphi_2)(v) = \varphi_1(v)\varphi_2'(v) \leqslant 0.$$

Or un simple calcul nous donne :

$$W'(\varphi_1,\varphi_2)(t) = \varphi_1(t)\varphi_2''(t) - \varphi_2(t)\varphi_1''(t) = \varphi_1(t)\varphi_2(t)(q_1(t) - q_2(t)) \geqslant 0,$$

sur $]u,v[$ de sorte que $W(\varphi_1,\varphi_2)$ devrait être croissante, d'où la contradiction. Les autres cas se traîtent de la même façon. □

Exemple 7.10. *(Fonctions de Bessel)*. Soient $\nu \in \mathbb{R}_+$. Considérons l'edo de Sturm–Liouville suivante, définie sur $]0,\infty[$:

$$x''(t) + \frac{1}{t}x'(t) + \left(1 - \frac{\nu^2}{t^2}\right)x(t) = 0, \tag{7.44}$$

appelée équation de Bessel. Sous forme réduite, elle s'écrit :

$$y''(t) + \left(1 + \frac{1 - 4\nu^2}{4t^2}\right)y(t) = 0. \tag{7.45}$$

En appliquant le théorème 7.16, nous allons comparer les zéros ce cette edo avec ceux de

$$x''(t) + x(t) = 0,$$

dont les solutions sont $\varphi(t) = A\,\sin(t-\theta)$, $A \in \mathbb{R}$, $\theta \in \mathbb{R}$. Distinguons trois cas :

- Soit $0 \leqslant \nu \leqslant 1/2$, alors $1 + (1 - 4\nu^2/4t^2) > 1$ et toute solution de (7.45) admet au moins un zéro dans tout intervalle de longueur π.
- Soit $\nu > 1/2$, alors $1 + (1 - 4\nu^2/4t^2) \leqslant 1$ et les zéros (éventuels) des solutions de (7.45) sont distants d'une longueur plus grande que π.
- Si $\nu = 1/2$, les solutions de (7.45) sont $A\,\sin(t+\theta)$ et les zéros sont écartés d'une distance exactement égale à π.

On démontre qu'en fait, les solutions des edo de Bessel, quelque soit la valeur de ν, ont toujours une infinité de zéros sur $]0,\infty[$. Pour tout $\nu > 0$, l'edo de Bessel (7.44) admet deux solutions linéairement indépendantes notées J_ν (fonction de Bessel de première espèce) et Y_ν (fonction de Bessel de deuxième espèce). Les fonctions J_ν sont prolongeables par continuité en $t = 0$ alors les fonctions Y_ν ont en ce point une asymptote verticale.

Selon the théorème 7.14, entre deux zéros de J_ν il y a exactement un zéro de Y_ν et réciproquement. Les graphes des fonctions J_ν et Y_ν (pour différentes valeurs de ν) sont représentés sur les figures 7.8 et 7.9 ci–dessous.

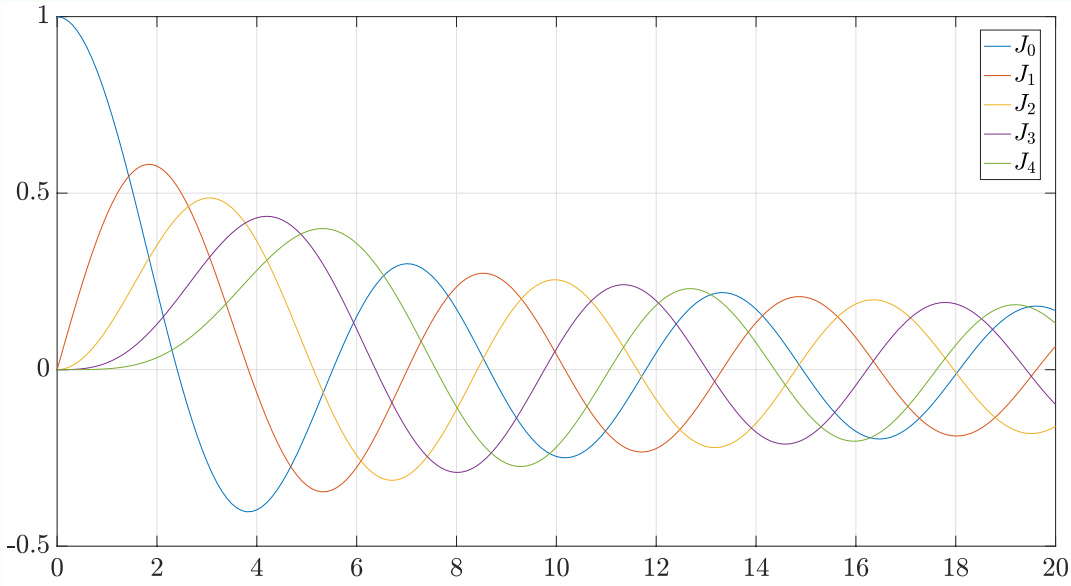


FIG. 7.8 – *Les fonctions de Bessel de première espèce J_ν, solutions de* (7.45) *pour* $\nu = 0,1,2,3,4$.

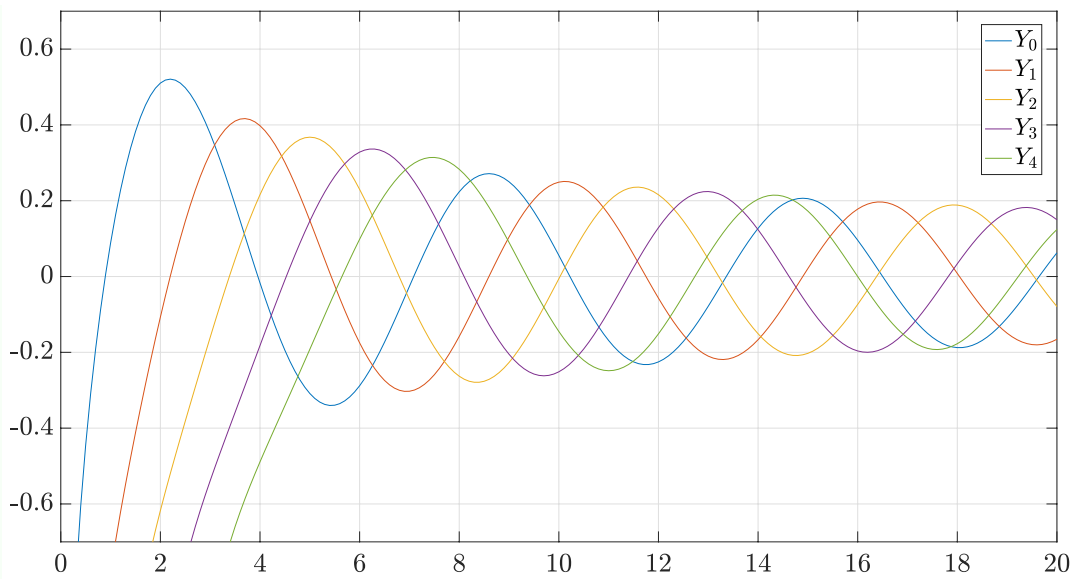


FIG. 7.9 – *Les fonctions de Bessel de deuxième espèce Y_ν, solutions de* (7.45) *pour* $\nu = 0,1,2,3,4$.

7.11 Stabilité des équations différentielles autonomes

Nous poursuivons dans cette section l'étude des edo autonomes initiée dans la section 7.9. Rappelons que les edo autonomes sont les edo qui s'écrivent sous la forme :

$$x'(t) = f(x(t)), \tag{7.46}$$

où f est une fonction de classe $\mathscr{C}^1$ d'un ouvert $\mathscr{U}$ de $\mathbb{R}^N$ à valeurs dans $\mathbb{R}^N$. Pour tout x dans $\mathscr{U}$, la solution de l'edo (7.46) qui vaut x au temps $t = 0$ est notée $\phi(\cdot,x)$. Elle est définie sur un intervalle de $\mathbb{R}$ noté $J(x)$. On définit ensuite l'ouvert de $\mathbb{R} \times \mathbb{R}^N$:

$$\mathscr{V} = \{(t,x) \in \mathbb{R} \times \mathscr{U} \ : \ t \in J(x)\} = \bigcup_{x \in \mathscr{U}} \Big(J(x) \times \{x\} \Big),$$

ainsi que le flot associé à l'edo :

$$\begin{aligned} \phi : \mathscr{V} &\longrightarrow \mathbb{R}^N \\ (t,x) &\longmapsto \phi(t,x). \end{aligned}$$

Les principales propriétés du flots sont regroupées dans le théorème 7.9. Introduisons maintenant une nouvelle notion :

Définition 7.8. *Un point d'équilibre (ou point critique) de l'edo* (7.46) *est un point $x_0 \in \mathscr{U}$ pour lequel $f(x_0) = 0$.*

La suite de cette section est dédiée essentiellement à l'étude de l'allure des orbites de l'edo (7.46) au voisinage des points d'équilibre, mais auparavant traitons rapidement le cas des points x pour lesquels $f(x) \neq 0$. Le résultat suivant affirme qu'au voisinage de tels points, les trajectoires de l'edo (7.46) sont les images par un difféomorphisme de droites parallèles.

Théorème 7.17. *Soit $x^* = (x_1^*,\dots,x_N^*)$ un point de $\mathscr{U}$ pour lequel $f(x^*) \neq 0$ (c'est à dire que $f_k(x^*) \neq 0$ pour un certain k dans $\{1,\dots,N\}$). Alors il existe un voisinage $\mathscr{U}^*$ de x^* dans $\mathscr{U}$, un réel $\delta > 0$ et un voisinage Ω de $y^* = (x_1^*,\dots,x_{k-1}^*,x_{k+1}^*,x_N^*)$ dans $\mathbb{R}^{N-1}$ tels que la fonction :*

$$\begin{aligned} F :]-\delta,\delta[\times\Omega &\longrightarrow \mathscr{U}^* \\ (t,y) &\longmapsto \phi(t,y_1,\dots,y_{k-1},x_k^*,y_k,\dots,y_{N-1}), \end{aligned}$$

soit un $\mathscr{C}^1$ difféomorphisme.

Comme indiqué plus haut, pour interpréter ce résultat, on pourra considérer le difféomorphisme F^{-1} qui envoie, sur un voisinage du point x^*, les trajectoires de l'edo sur des droites parallèles. Illustrons cela avec un exemple.

Exemple 7.11. On considère la même edo que celle de l'exemple 7.4, à savoir :

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = f(x(t),y(t)), \\ (x(0),y(0)) = (x_0,y_0), \end{cases}$$

où $(x_0,y_0) \in \mathbb{R}^2$ et f est la fonction de classe $\mathscr{C}^1$ (en fait $\mathscr{C}^\infty$) définie par :

$$\begin{aligned} f : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x,y) &\longmapsto \begin{pmatrix} 1 \\ 2xy^2 \end{pmatrix}. \end{aligned}$$

La solution du problème de Cauchy ci–dessus est :

$$\begin{aligned} \phi(\cdot,(x_0,y_0)) : J(x_0,y_0) &\longrightarrow \mathbb{R} \\ t &\longmapsto \begin{pmatrix} t + x_0 \\ \dfrac{y_0}{1 - ty_0(t+2x_0)} \end{pmatrix}. \end{aligned}$$

On souhaite appliquer le théorème 7.17 au voisinage du point (0,0) (c'est le point x^* dans l'énoncé du théorème). Un calcul explicite conduit aux expressions suivantes :

$$\begin{aligned} F :]-1,1[\times]-\tfrac{1}{2},\tfrac{1}{2}[&\longrightarrow \mathscr{U}^* \\ (t,y) &\longmapsto \begin{pmatrix} t \\ \frac{y}{1-t^2y} \end{pmatrix} \end{aligned} \quad \text{et} \quad \begin{aligned} F^{-1} : \mathscr{U}^* &\longrightarrow]-1,1[\times]-\tfrac{1}{2},\tfrac{1}{2}[\\ (x,y) &\longmapsto \begin{pmatrix} x \\ \frac{y}{1+x^2y} \end{pmatrix}. \end{aligned} \tag{7.47}$$

L'ouvert $\mathscr{U}^*$ est l'image par F du pavé $]-1,1[\times]-\frac{1}{2},\frac{1}{2}[$, il est représenté sur la figure 7.10.

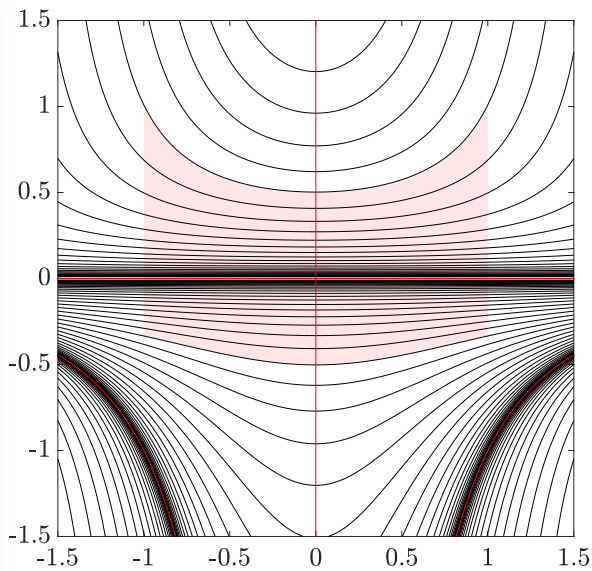


FIG. 7.10 – *L'ouvert $\mathscr{U}^*$ (en rose) intervenant dans la définition de la fonction F (cf.* (7.47)*). La fonction F^{-1} est un difféomorphisme qui envoie les orbites de l'edo contenues dans $\mathscr{U}^*$ (lignes en noir sur le dessin) sur des droites horizontales dans le pavé $]-1,1[\times]-\frac{1}{2},\frac{1}{2}[$.*

Démonstration du théorème 7.17. Le flot ϕ est de classe $\mathscr{C}^1$ sur $\mathscr{V}$ d'après le théorème 7.9. Introduisons l'ensemble

$$\mathscr{V}_k^* = \mathscr{V} \cap \big\{(t,x_1,\ldots,x_N) \in \mathbb{R}\times\mathbb{R}^N \,:\, x_k = x_k^*\big\},$$

qui est un ouvert de $\mathbb{R}\times\mathbb{R}^{N-1}$. La fonction :

$$\begin{aligned} F : \mathscr{V}_k^* &\longrightarrow \mathbb{R}^N \\ (t,y) &\longmapsto \phi(t,y_1,\ldots,y_{k-1},x_k^*,y_k,\ldots,y_{N-1}), \end{aligned}$$

est alors de classe $\mathscr{C}^1$ (comme le flot ϕ). D'autre part, par définition du flot

$$F(0,y) = (y_1,\ldots,y_{k-1},x_k^*,y_k,\ldots,y_{N-1}),$$

pour tout y dans l'ouvert de $\mathbb{R}^{N-1}$:

$$\mathscr{U}_k^* = \mathscr{U} \cap \big\{(x_1,\ldots,x_N) \in \mathbb{R}^N \,:\, x_k = x_k^*\big\}.$$

Notant $y^* = (x_1^*,\ldots,x_{k-1}^*,x_{k+1}^*,\ldots,x_N^*)$, on en déduit l'expression de la matrice Jacobienne de F au point $(0,y^*)$, à savoir :

$$JF(0,y^*) = \left[\begin{array}{c|ccc|ccc} \frac{\partial\phi_1}{\partial t}(0,x^*) & & & & 0 & \cdots & 0 \\ \vdots & & \mathrm{Id}_{k-1} & & \vdots & & \vdots \\ \frac{\partial\phi_{k-1}}{\partial t}(0,x^*) & & & & 0 & \cdots & 0 \\ \hline \frac{\partial\phi_k}{\partial t}(0,x^*) & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \hline \frac{\partial\phi_{k+1}}{\partial t}(0,x^*) & 0 & \cdots & 0 & & & \\ \vdots & \vdots & & \vdots & & \mathrm{Id}_{N-k} & \\ \frac{\partial\phi_N}{\partial t}(0,x^*) & 0 & \cdots & 0 & & & \end{array}\right] = \left[\begin{array}{c|ccc|ccc} f_1(x^*) & & & & 0 & \cdots & 0 \\ \vdots & & \mathrm{Id}_{k-1} & & \vdots & & \vdots \\ f_{k-1}(x^*) & & & & 0 & \cdots & 0 \\ \hline f_k(x^*) & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \hline f_{k+1}(x^*) & 0 & \cdots & 0 & & & \\ \vdots & \vdots & & \vdots & & \mathrm{Id}_{N-k} & \\ f_N(x^*) & 0 & \cdots & 0 & & & \end{array}\right],$$

dont le déterminant est égal à $(-1)^k f_k(x^*) \neq 0$. On peut alors appliquer le théorème d'inversion locale à la fonction F au voisinage du point $(0,y^*)$, ce qui nous donne la conclusion du théorème. $\square$

Intéressons nous maintenant au comportement des orbites au voisinage d'un points d'équilibre x_0 dans $\mathscr{U}$. On remarque que pour un tel point, $\phi(t,x_0) = x_0$ pour tout t dans $\mathbb{R}$ d'où la dénomination "point d'équilibre". Parmis les points d'équilibre, on peut distinguer plusieurs catégories :

Définition 7.9. *Un point d'équilibre $x_0 \in \mathscr{U}$ est dit*

- *Stable si pour tout $\varepsilon > 0$ il existe $\delta_\varepsilon > 0$ tel que, pour tout $x \in \mathscr{U}$*

$$\Big(\|x - x_0\| \leqslant \delta_\varepsilon\Big) \Rightarrow \Big([0, +\infty[\subset J(x) \text{ et pour tout } t \in [0, +\infty[, \quad \|\phi(t,x) - x_0\| < \varepsilon\Big).$$

- *Instable s'il n'est pas stable.*
- *Asymptotiquement stable s'il est stable et s'il existe $\delta > 0$ tel que pour tout x dans $\mathscr{U}$:*

$$\Big(\|x - x_0\| < \delta\Big) \Rightarrow \Big([0, +\infty[\subset J(x) \text{ et } \lim_{t\to\infty} \phi(t,x) = x_0\Big).$$

Autrement dit, un point x_0 est stable si l'on peut rendre la solution $\phi(t,x)$ aussi proche que l'on veut de x_0 pour tout $t \geq 0$ pourvu que x soit suffisamment proche de x_0.

Commençons pas considérer le cas le plus simple, celui des edo (autonomes) linéaires à cœfficients constants en dimension 2.

7.11.1 Stabilité des edo linéaires à cœfficients constants

Dans tout ce paragraphe, on supposera donc que M est une matrice de $\mathscr{M}_2(\mathbb{R})$ qui vérifie de plus $\det M \neq 0$. On considère alors l'edo linéaire :

$$x'(t) = Mx(t), \tag{7.48}$$

de sorte que $x = (0,0)$ est le seul point d'équilibre de l'équation (7.48). L'allure des trajectoires dépend essentiellement de la nature des valeurs propres de M. Ainsi, on est amené à distinguer les cas suivants :

A. La matrice M a deux valeurs propres réelles distinctes

Soient λ_1 et λ_2 les valeurs propres de la matrice M et u_1,u_2 des vecteurs propres associés. Pour tout $x \neq (0,0)$, la solution $t \mapsto \phi(t,x)$ à l'edo (7.48) s'écrit dans la base $\{u_1,u_2\}$:

$$\phi(t,x) = \alpha_1(t)u_1 + \alpha_2(t)u_2.$$

Dérivant cette expression, on déduit que :

$$\phi'(t,x) = \alpha_1'(t)u_1 + \alpha_2'(t)u_2 = M\phi(t,x) = \alpha_1(t)Mu_1 + \alpha_2(t)Mu_2,$$

et donc finalement que :

$$\alpha_1'(t)u_1 + \alpha_2'(t)u_2 = \alpha_1(t)\lambda_1 u_1 + \alpha_2(t)\lambda_2 u_2.$$

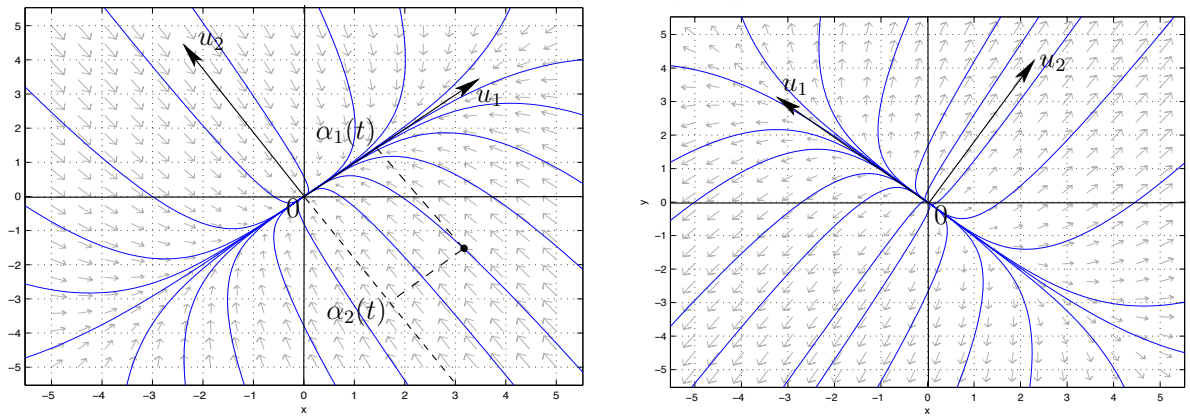
Comme $\{u_1,u_2\}$ est un système libre, cette égalité implique que, pour tout t dans $\mathbb{R}$:

$$\begin{cases} \alpha_1'(t) - \lambda_1\alpha_1(t) = 0 \\ \alpha_2'(t) - \lambda_2\alpha_1(t) = 0. \end{cases}$$

Sachant que $\phi(0,x) = x = \widetilde{x}_1 u_1 + \widetilde{x}_2 u_2$ ($\tilde{x}_1$ et $\tilde{x}_2$ sont donc les coordonnées de x dans la base $\{u_1,u_2\}$), on obtient que pour tout t dans $\mathbb{R}$:

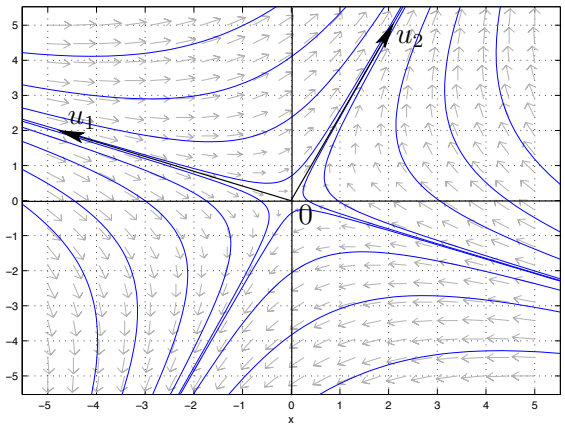
$$\begin{cases} \alpha_1(t) = \widetilde{x}_1 e^{\lambda_1 t} \\ \alpha_2(t) = \widetilde{x}_2 e^{\lambda_2 t}. \end{cases} \tag{7.49a}$$

De cette étude on déduit que le point d'équilibre (0,0) est asymptotiquement stable si les valeurs propres λ_1 et λ_2 sont strictement négatives (on parle alors de nœud stable) et instable si $\lambda_1 > 0$ ou $\lambda_2 > 0$ (on parle alors de nœud instable si les deux valeurs propres sont positives et de point selle ou col si l'une seulement des deux est positive). Ces cas sont illustrés sur les figure 7.11 et 7.12.



$\lambda_1 = -1, \quad \lambda_2 = -4$ $\qquad$ $\lambda_1 = 1, \quad \lambda_2 = 4$

FIG. 7.11 – *Quelques orbites de l'edo* (7.48) *dans le cas d'un nœud stable à gauche et d'un nœud instable à droite (considérer le champ de gradients).*



$\lambda_1 = -\sqrt{2}, \quad \lambda_2 = \sqrt{2}$

FIG. 7.12 – *Quelques orbites de l'edo* (7.48) *dans le cas d'un point selle (ou col).*

B. La matrice M a une seule valeur propre réelle

On note alors $\lambda \neq 0$ cette valeur propre double et $d = \dim \mathrm{Ker}\,(M - \lambda \mathrm{Id}_2)$. Il faut encore distinguer deux sous–cas : $d = 2$ (la matrice est diagonalisable) et $d = 1$ (la matrice n'est pas diagonalisable).

La matrice M est diagonalisable ($d = 2$). Dans ce cas, la matrice M est en fait tout simplement diagonale et s'écrit $M = \lambda \mathrm{Id}_2$. Il est facile de montrer que, pour tout $x \in \mathbb{R}^2$, $x \neq (0,0)$,

$$\phi(t,x) = e^{\lambda t} x, \qquad \text{pour tout } t \in \mathbb{R}. \tag{7.49b}$$

Les orbites sont des demis–droites issues de l'origine (exclue) et passant par x. Le point critique (0,0) est stable si $\lambda < 0$ (on parle de puit) et instable si $\lambda > 0$ (on parle alors de point source). Ces cas sont illustrés sur la figure 7.13.

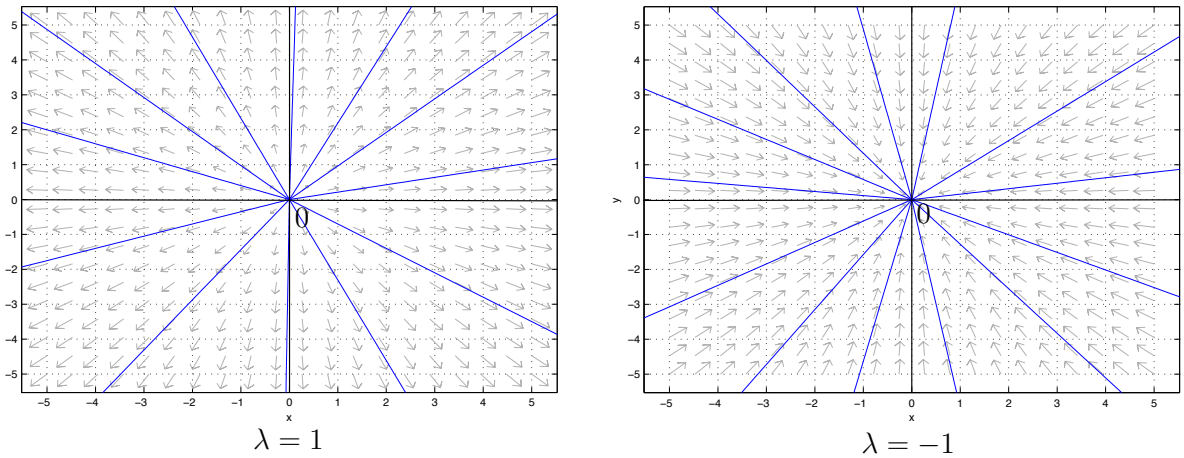


FIG. 7.13 – *Quelques orbites de l'edo* (7.48) *dans les cas où l'origine est un point source (instable) à gauche et un puit (stable) à droite.*

La matrice M n'est pas diagonalisable ($d = 1$). Il existe une base $\{u_1,u_2\}$ telle que $Mu_1 = \lambda u_1$ et $Mu_2 = u_1 + \lambda u_2$ (nous admettrons ce résultat d'algèbre linéaire). Dans cette base, les solutions $t \mapsto \phi(t,x)$ s'écrivent $\phi(t,x) = \alpha_1(t)u_1 + \alpha_2(t)u_2$. On dérive cette égalité pour obtenir :

$$\phi'(t,x) = \alpha_1'(t)u_1 + \alpha_2'(t)u_2 = \alpha_1(t)Mu_1 + \alpha_2(t)Mu_2 = \big(\lambda\alpha_1(t) + \alpha_2(t)\big)u_1 + \lambda\alpha_2(t)u_2,$$

ce qui permet d'en déduire le système d'edo suivant vérifié par les fonctions α_1 et α_2 :

$$\begin{cases} \alpha_1'(t) = \lambda\alpha_1(t) + \alpha_2(t) \\ \alpha_2'(t) = \lambda\alpha_2(t). \end{cases}$$

En tenant compte de la condition au temps $t = 0$, on peut intégrer ce système pour trouver :

$$\begin{cases} \alpha_1(t) = (\widetilde{x}_2\, t + \widetilde{x}_1)e^{\lambda t} \\ \alpha_2(t) = \widetilde{x}_2\, e^{\lambda t}, \end{cases} \tag{7.49c}$$

où comme précédemment $x = \widetilde{x}_1 u_1 + \widetilde{x}_2 u_2$ est la décomposition de x dans la base $\{u_1,u_2\}$. On en conclut ensuite assez facilement que le point d'équilibre (0,0) est asymptotiquement stable si $\lambda < 0$ (on parle de nœud dégénéré stable) et instable si $\lambda > 0$ (on parle alors de nœud dégénéré instable). Ces cas sont illustrés sur la figure 7.14.

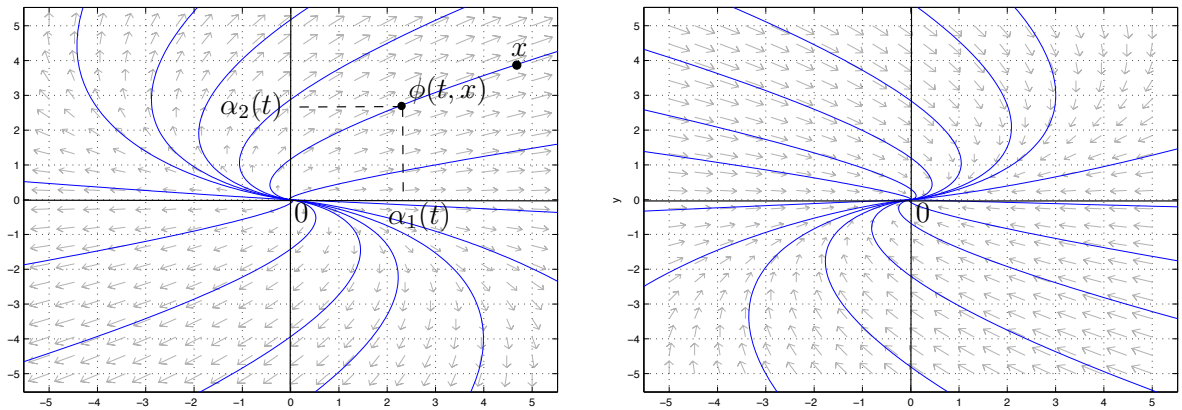


FIG. 7.14 – *Quelques trajectoires de l'edo* (7.48) *dans les cas où l'origine est un nœud dégénéré instable (à gauche) et un nœud dégénéré stable (à droite).*

C. La matrice M a deux valeurs propres, non réelles, conjuguées

Notons $\lambda = a + ib$ et $\bar{\lambda} = a - ib$ (a et b des réels avec $b \neq 0$) les valeurs propres complexes de M et considérons $z = u + iv$ un vecteur propre associé à la valeur propre λ (u et v des vecteurs de $\mathbb{R}^2$). Alors $\bar{z} = u - iv$ est un vecteur propre associé à la valeur propre $\bar{\lambda}$ et les vecteurs $u = (z + \bar{z})/2$ et $v = (z - \bar{z})/(2i)$ forment un système libre (et donc une base) dans $\mathbb{R}^2$. D'autre part :

$$Mz = Mu + iMv = (a + ib)(u + iv) = (au - bv) + i(av + bu),$$

c'est à dire que

$$\begin{cases} Mu = au - bv \\ Mv = av + bu. \end{cases}$$

Dans la base $\{u,v\}$ la solution $t \mapsto \phi(t,x)$ (pour x dans $\mathbb{R}^2$) s'écrit :

$$\phi(t,x) = \alpha_1(t)u + \alpha_2(t)v \qquad \text{pour tout } t \in \mathbb{R}.$$

On dérive cette égalité pour obtenir :

$$\alpha_1'(t)u + \alpha_2'(t)v = \phi'(t,x) = \alpha_1(t)Mu + \alpha_2(t)Mv = \big(a\alpha_1(t) + b\alpha_2(t)\big)u + \big(a\alpha_2(t) - b\alpha_1(t)\big)v.$$

Les fonctions $t \mapsto \alpha_1(t)$ et $t \mapsto \alpha_2(t)$ sont ainsi solutions du système d'edo :

$$\begin{cases} \alpha_1'(t) - a\,\alpha_1(t) = b\,\alpha_2(t), \\ \alpha_2'(t) - a\,\alpha_2(t) = -b\,\alpha_1(t) \qquad \text{pour tout } t \in \mathbb{R}. \end{cases}$$

Pour résoudre ce système introduisons les fonctions β_k ($k = 1,2$) définies pour tout t dans $\mathbb{R}$ par $\beta_k(t) = \alpha_k(t)e^{-a\,t}$. Ces fonctions vérifient :

$$\begin{cases} \beta_1'(t) = -b\,\beta_2(t), \\ \beta_2'(t) = b\,\beta_1(t) \qquad \text{pour tout } t \in \mathbb{R}. \end{cases}$$

Dans la base $\{u,v\}$, la donnée initiale $x \neq (0,0)$ se décompose en $x = \widetilde{x}_1 u + \widetilde{x}_2 v$. Introduisons $R \in]0,\infty[$ et $\theta \in [-\pi,\pi[$ tels que $\widetilde{x}_1 = R\cos(\theta)$ et $\widetilde{x}_2 = R\sin(\theta)$. En intégrant le système d'edo

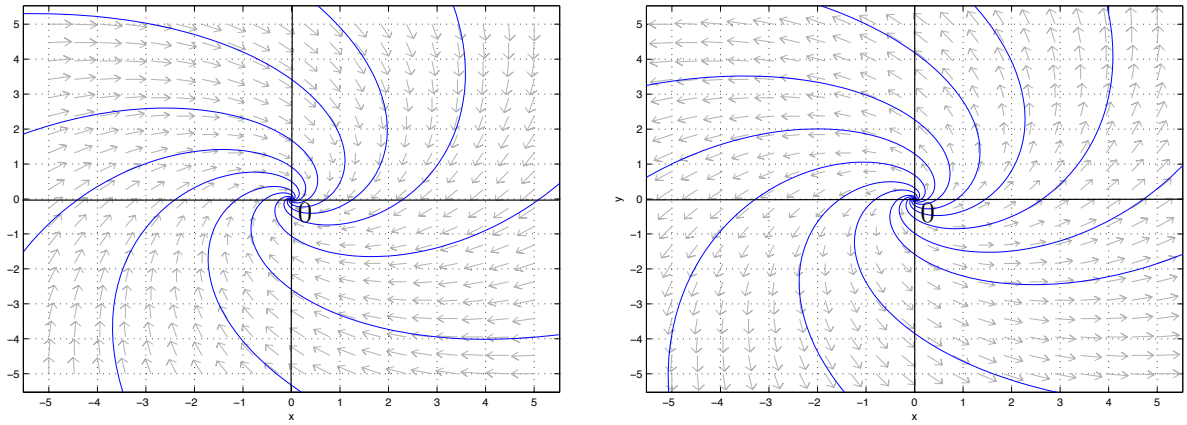
ci-dessus, on trouve que pour tout t dans $\mathbb{R}$, $\beta_1(t) = R\cos(b\,t+\theta)$ et $\beta_2(t) = R\sin(bt+\theta)$ ce qui conduit aux expressions :

$$\begin{cases} \alpha_1(t) = R\,e^{a\,t}\cos(bt+\theta), \\ \alpha_2(t) = R\,e^{a\,t}\sin(bt+\theta) \qquad \text{pour tout } t\in\mathbb{R}. \end{cases} \tag{7.49d}$$

On en déduit, en fonction de $a = \mathrm{Re}(\lambda)$:

- Si $a = 0$ alors dans la base $\{u,v\}$ les trajectoires sont des cercles centrés à l'origine et de rayons R. L'origine est alors appelée un centre. C'est un point d'équilibre stable mais pas asymptotiquement stable. Attention, les orbites ne sont pas des cercles dans la base canonique d'origine (ce sont les images d'ellipses par une application linéaire, donc des ellipses un peu "tordues").
- Si $a > 0$. Les orbites sont des spirales qui émanent de l'origine. Le point d'équilibre (0,0) est alors appelé un foyer instable.
- Si $a < 0$. Les trajectoires sont des spirales qui convergent vers l'origine. Le point (0,0) est appelé un foyer stable. Il est même asymptotiquement stable.

En guise d'illustration, quelques trajectoires de l'edo (7.48) sont représentées sur les figures 7.15 et 7.16.



$\lambda_1 = -1-i, \quad \lambda_2 = -1+i$ $\qquad$ $\lambda_1 = 1-i, \quad \lambda_2 = 1+i$

FIG. 7.15 – *Quelques trajectoires de l'edo* (7.48) *dans le cas où la matrice M a deux valeurs propres complexes conjuguées λ_1 et $\lambda_2 = \bar{\lambda}_1$ avec* $\mathrm{Re}(\lambda_1) \neq 0$. *Le point* (0,0) *est un foyer stable à gauche (considérer le champ de gradients) et un foyer instable à droite.*

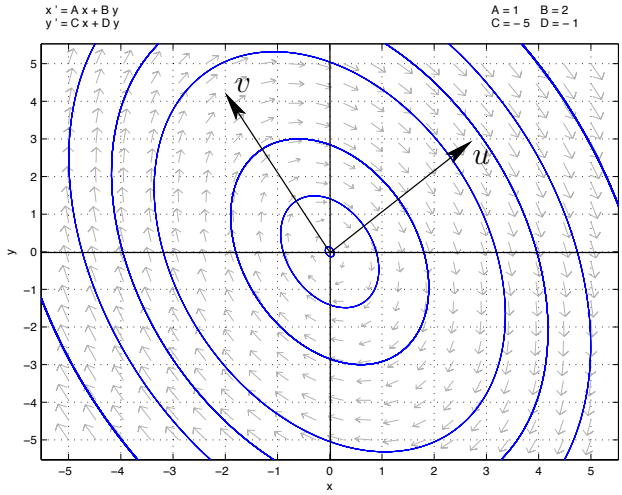
On déduit des différentes expressions (7.49) établies ci–dessus le corollaire suivant :

Corollaire 7.2. *Soient λ_1 et λ_2 les valeurs propres dans $\mathbb{C}$ de la matrice M (avec éventuellement $\lambda_1 = \lambda_2$). Notons*

$$\lambda = \max\big\{\,\mathrm{Re}(\lambda_1),\,\mathrm{Re}(\lambda_2)\big\}.$$

Alors pour tout réel $\sigma > \lambda$, il existe un réel $C_\sigma > 0$ tel que pour tout x dans $\mathbb{R}^2$:

$$\|\phi(t,x)\| \leqslant C_\sigma\|x\|e^{\sigma t} \qquad \textit{pour tout } t \geqslant 0.$$



$$\lambda_1 = i, \quad \lambda_2 = -i$$

FIG. 7.16 – *Quelques trajectoires de l'edo* (7.48) *dans le cas où la matrice M a deux valeurs propres complexes conjuguées λ_1 et $\lambda_2 = \bar{\lambda}_1$ avec* $\mathrm{Re}(\lambda_1) = 0$. *Le point d'équilibre* (0,0) *est appelé un centre. Il est stable mais pas asymptotiquement stable.*

En particulier, si $\lambda < 0$ alors on peut choisir $\sigma < 0$ et l'origine est un point d'équilibre asymptotiquement stable de l'edo (7.50).

7.11.2 Stabilité des edo autonomes non linéaires

Nous allons maintenant expliquer comment les résultats détaillés dans le paragraphe précédent (le cas des edo autonomes linéaires en dimension 2) peuvent nous servir pour décrire l'allure des orbites de l'edo (7.46) autour d'un point d'équilibre dans le cas général considéré au début de cette section.

L'idée principale consiste à linéariser l'edo (7.46) au voisinage du point d'équilibre $x_0 \in \mathbb{R}^2$ qui nous intéresse. En effet, comme la fonction f est de classe $\mathscr{C}^1$, on peut écrire que pour tout x dans un voisinage de x_0 :

$$f(x) = Df(x_0)(x - x_0) + g(x - x_0),$$

où la fonction g, définie sur un voisinage $\mathscr{O}$ de l'origine, est de classe $\mathscr{C}^1$ et vérifie :

$$\lim_{\substack{\|x\| \to 0 \\ x \neq 0}} \frac{\|g(x)\|}{\|x\|} = 0.$$

L'edo (7.46) s'écrit alors de façon équivalente sous la forme :

$$x'(t) = Df(x_0)(x(t) - x_0) + g(x(t) - x_0).$$

On effectue le changement de variables $y(t) = x(t) - x_0$ pour obtenir :

$$y'(t) = Df(x_0)y(t) + g(y(t)),$$

et on peut donc sans perte de généralité se ramener au cas où $x_0 = 0$. L'idée est maintenant de déduire la nature du point d'équilibre $x_0 = 0$ en fonctions des valeurs propres de la matrice Jabcobienne $Df(x_0)$. Énonçons un résultat dans ce sens :

Théorème 7.18 (de Perron)**.** *Soient λ_1 et λ_2 les valeurs propres dans $\mathbb{C}$ d'une matrice M dans $\mathscr{M}_2(\mathbb{R})$ (avec éventuellement $\lambda_1 = \lambda_2$). On note*

$$\lambda = \max\big\{\operatorname{Re}(\lambda_1), \operatorname{Re}(\lambda_2)\big\},$$

et on suppose que $\lambda < 0$. Soit g une fonction $\mathscr{C}^1$ définie sur un voisinage $\mathscr{O}$ de 0 dans $\mathbb{R}^2$, à valeurs dans $\mathbb{R}^2$ et vérifiant

$$\lim_{\substack{\|x\|\to 0\\ x\neq 0}} \frac{\|g(x)\|}{\|x\|} = 0.$$

Pour tout x dans $\mathscr{O}$, on note $t \mapsto \phi(t,x)$ la solution maximale de l'edo autonome :

$$x'(t) = Mx(t) + g(x(t)), \tag{7.50}$$

vérifiant $\phi(0,x) = x$. Alors pour tout $\varepsilon > 0$ et pour tout réel σ dans l'intervalle $]\lambda,0[$ il existe $\delta > 0$ tel que pour tout x de $\mathscr{O}$

$$(\|x\| < \delta) \quad \Rightarrow \quad \big(\|\phi(t,x)\| \leq \varepsilon e^{\sigma t} \qquad \text{pour tout } t \in [0,T_+[\big)\,,$$

où $J(x) =]T_-,T_+[$ est l'intervalle maximal d'existence de la solution $t \mapsto \phi(t,x)$.

Si $\overline{B(0,\varepsilon)} \subset \mathscr{O}$, alors $T_+ = +\infty$ et l'origine est un point d'équilibre asymptotiquement stable de l'edo (7.50).

Démonstration. Soit $\varepsilon > 0$ et σ comme dans l'énoncé. Introduisons également σ_0 tel que

$$\lambda < \sigma_0 < \sigma < 0,$$

ce qui est toujours possible. Le Corollaire 7.2 nous assure de l'existence d'une constante $C_{\sigma_0} > 0$ telle que pour tout x dans $\mathbb{R}^2$:

$$\big\|e^{tM}x\big\| = \|\phi(t,x)\| \leqslant C_{\sigma_0}\|x\|e^{\sigma_0 t} \qquad \text{pour tout } t \geqslant 0.$$

On en déduit, en prenant le supremum sur tout les x de norme 1, que :

$$\big\|e^{tM}\|_{\mathscr{L}(\mathbb{R}^2)} \leqslant C_{\sigma_0}e^{\sigma_0 t} \qquad \text{pour tout } t \geqslant 0. \tag{7.51a}$$

D'après l'hypothèse vérifiée par g, il existe $\delta_0 > 0$ tel que pour tout x dans $\mathscr{O}$:

$$(\|x\| \leqslant \delta_0) \quad \Rightarrow \quad \left(\|g(x)\| \leqslant \left(\frac{\sigma - \sigma_0}{C_{\sigma_0}}\right)\|x\|\right). \tag{7.51b}$$

Posons alors :

$$\delta = \min\left\{\frac{\delta_0}{2}, \frac{\delta_0}{C_{\sigma_0}}, \frac{\varepsilon}{C_{\sigma_0}}\right\},$$

et choisissons x dans $\mathscr{O}$ tel que $\|x\| \leq \delta$. Pour un tel x, notons comme toujours $t \mapsto \phi(t,x)$ la solution maximale de l'edo (7.50) telle que $\phi(0,x) = x$ (remarquer qu'une telle solution existe toujours car g est $\mathscr{C}^1$ sur $\mathscr{O}$). Son intervalle de définition maximal est $]T_-,T_+[$ avec $-\infty \leqslant T_- < 0 < T_+ \leqslant \infty$. D'après l'identité (7.31) de la proposition 7.7 :

$$\phi(t,x) = e^{tM}x + \int_0^t e^{(t-s)M}g(\phi(s,x))\,\mathrm{d}s \qquad \text{pour tout } t \in [0,T_+[.$$

On en déduit que

$$\|\phi(t,x)\| \leqslant \left\|e^{tM}\right\|_{\mathscr{L}(\mathbb{R}^2)}\|x\| + \int_0^t \left\|e^{(t-s)M}\right\|_{\mathscr{L}(\mathbb{R}^2)}\|g(\phi(s,x))\|\,\mathrm{d}s \qquad \text{pour tout } t \in [0,T_+[. \tag{7.51c}$$

Rappelons que $\|x\| \leq \delta \leq \delta_0/2$ et notons :

$$T_+^* = \sup\left\{T \in [0,T_+[: \|\phi(t,x)\| \leqslant \delta_0 \text{ pour tout } t \in [0,T[\right\}.$$

La continuité de la fonction $t \mapsto \phi(t,x)$ entraîne que $T_+^* > 0$. En combinant les trois assertions (7.51), on obtient l'estimation :

$$\|\phi(t,x)\| \leqslant C_{\sigma_0}e^{\sigma_0 t}\delta + (\sigma - \sigma_0)e^{t\sigma_0}\int_0^t e^{-s\sigma_0}\|\phi(s,x)\|\,\mathrm{d}s \qquad \text{pour tout } 0 \leqslant t \in [0,T_+^*[.$$

Pour tout t dans $[0,T_+^*[$ posons $\varphi(t,x) = \phi(t,x)e^{-t\sigma_0}$ et réécrivons l'inégalité ci–dessus sous la forme :

$$\|\varphi(t,x)\| \leqslant C_{\sigma_0}\delta + (\sigma - \sigma_0)\int_0^t \|\varphi(s,x)\|\,\mathrm{d}s \qquad \text{pour tout } t \in [0,T_+^*[.$$

En appliquant alors l'inégalité de Grönwall du Lemme 7.1, on obtient :

$$\|\varphi(t,x)\| \leqslant C_{\sigma_0}\delta e^{(\sigma-\sigma_0)t} \qquad \text{pour tout } t \in [0,T_+^*[,$$

et donc :

$$\|\phi(t,x)\| \leqslant C_{\sigma_0}\delta e^{\sigma t} \qquad \text{pour tout } t \in [0,T_+^*[. \tag{7.52}$$

On en déduit d'une part, selon la définition de δ, que

$$\|\phi(t,x)\| \leqslant \delta_0 e^{\sigma_0 t} \leqslant \delta_0 e^{\sigma_0 T_+^*} \qquad \text{pour tout } t \in [0,T_+^*[.$$

Si on avait $T_+^* < T_+$ alors par un argument de continuité, on pourrait en déduire que $\|\phi(t,x)\|$ reste inférieur à δ_0 au delà de T_+^*, ce qui contredirait la définition de T_+^*. On a donc nécessairement $T_+^* = T_+$.

D'autre part, toujours d'après la définition de δ, l'estimation (7.52) entraîne que :

$$\|\phi(t,x)\| \leqslant \varepsilon e^{\sigma t} \qquad \text{pour tout } t \in [0,T_+[,$$

ce qui est l'estimation du Théorème. Enfin, lorsque $\overline{B(0,\varepsilon)} \subset \mathscr{O}$, alors le théorème 7.7 nous assure que $T_+ = +\infty$ et l'origine est un point d'équilibre asymptotiquement stable. □

Le théorème 7.18 a une portée limitée dans la mesure où il faut que les parties réelles des valeurs propres de la matrice Jacobienne au point d'équilibre soient toutes strictement négatives. On peut énoncer un résultat un peu plus général pour lequel il faut seulement supposer que les parties réelles des valeurs propres de la matrice Jacobienne sont non nulles.

Théorème 7.19 (Hartman–Grobman). *Soit $\mathscr{U}$ un ouvert de $\mathbb{R}^2$ contenant l'origine et f une fonction de classe $\mathscr{C}^1$ sur $\mathscr{U}$ à valeurs dans $\mathbb{R}^2$ et telle que $f(0,0) = (0,0)$. On note λ_1 et λ_2 les deux valeurs propres dans $\mathbb{C}$ de la matrice Jacobienne M de f au point $(0,0)$ et on suppose que*

$$\mathrm{Re}(\lambda_1) \neq 0 \quad \textit{et} \quad \mathrm{Re}(\lambda_2) \neq 0. \tag{7.53}$$

Pour tout $x \in \mathscr{U}$, on note $t \mapsto \phi(t,x)$ la solution de l'edo autonome

$$x'(t) = f(x(t)), \tag{7.54a}$$

qui vérifie $\phi(0,x) = x$ *.*
Alors il existe deux ouverts $\mathscr{V}$ *et* $\mathscr{W}$ *de* $\mathbb{R}^2$ *contenant* $(0,0)$ *et un homéomorphisme* H *de* $\mathscr{V}$ *dans* $\mathscr{W}$ *tel que* $H(0,0) = (0,0)$ *et pour tout* $x \in \mathscr{V} \cap \mathscr{U}$ *,*

$$H(\phi(t,x)) = e^{tM} H(x) \qquad \text{pour tout } t \in J,$$

où J *est un intervalle réel contenant* 0*. En particulier* H *envoie localement (autour du point d'équilibre* $(0,0)$*) les trajectoires de l'edo* (7.54a) *sur les trajectoires de l'edo linéaire à cœfficients constants*

$$x'(t) = Mx(t), \tag{7.54b}$$

en préservant la paramétrisation. De ce résultat on peut déduire que sous l'hypothèse (7.53)*, le point d'équilibre* $(0,0)$ *est de même nature pour le système* (7.54a) *et pour le système* (7.54b)*.*

La démonstration est longue et difficile et nous l'admettrons. On peut en trouver les grandes lignes dans le livre de L. Perko *Differential equations and dynamical systems*, Springer Verlag.

7.12 Portrait de phase

Définition 7.10. *On appelle portrait de phase la représentation graphique des orbites (sous la forme de courbes paramétrées) d'une edo autonome en dimension 2.*

Exemple 7.12. On souhaite tracer le portrait de phase de l'edo autonome :

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = f(x_1(t),x_2(t)), \tag{7.55}$$

où l'expression de la fonction f est :

$$\begin{aligned} f : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x_1,x_2) &\longmapsto \begin{pmatrix} x_1 - x_1x_2 - x_1^2 \\ -4x_2 + 2x_1x_2 \end{pmatrix}. \end{aligned}$$

On commence par chercher les points d'équilibre de l'edo c'est à dire les points (x_1,x_2) vérifiant $f(x_1,x_2) = 0$. On doit résoudre le système (non linéaire) :

$$\begin{cases} x_1 - x_1x_2 - x_1^2 = 0 \\ -4x_2 + 2x_1x_2 = 0. \end{cases}$$

La deuxième équation conduit à $x_2(x_1 - 2) = 0$. Soit $x_1 = 2$ et avec la première équation on obtient $x_2 = -1$, soit $x_2 = 0$ et avec la première équation on trouve cette fois $x_1 = 0$ ou $x_1 = 1$. Finalement, les points d'équilibre sont :

$$(0,0), \quad \alpha = (1,0) \quad \text{et} \quad \beta = (2,-1).$$

Que ce soit pour utiliser le théorème de Perron ou le théorème de Hartman–Grobman, nous devons calculer la matrice Jacobienne de f en ces points. La fonction f est de classe $\mathscr{C}^1$ sur

$\mathbb{R}^2$ et pour tout $(x,y) \in \mathbb{R}^2$:

$$\mathrm{Jac}_f(x,y) = \begin{pmatrix} 1 - x_2 - 2x_1 & -x_1 \\ 2x_2 & -4 + 2x_1 \end{pmatrix}.$$

– Le point d'équilibre (0,0). En ce point, la matrice Jacobienne est diagonale et s'écrit :

$$\mathrm{Jac}_f(0,0) = \begin{pmatrix} 1 & 0 \\ 0 & -4 \end{pmatrix}.$$

Ses valeurs propres sont donc $\lambda_1 = 0$ et $\lambda_2 = -4$. On peut appliquer le théorème de Hartman–Grobman : Le point d'équilibre (0,0) de l'edo (7.55) est de même nature que le point d'équilibre (0,0) pour l'edo linéaire :

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -4 \end{pmatrix}\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix}.$$

C'est donc à un homéomorphisme près un point selle (ou col) comme celui représenté sur la figure 7.12.

– Le point d'équilibre α. En ce point, la matrice Jacobienne de f est triangulaire supérieure et s'écrit :

$$\mathrm{Jac}_f(1,0) = \begin{pmatrix} -1 & -1 \\ 0 & -2 \end{pmatrix}.$$

Les valeurs propres de cette matrice sont $\lambda_1 = -1$ et $\lambda_2 = -2$. Selon le théorème de Perron, le point d'équilibre α est asymptotiquement stable. Selon le théorème de Hartman–Grobman, c'est à un homéomorphisme près un nœud stable comme celui représenté à gauche sur la figure 7.11.

– Le point d'équilibre β. En ce point on trouve :

$$\mathrm{Jac}_f(2,-1) = \begin{pmatrix} -2 & -2 \\ -2 & 0 \end{pmatrix},$$

et les valeurs propres de cette matrice sont (après un calcul élémentaire) $\lambda_1 = 2 - 2\sqrt{5}$ et $\lambda_2 = 2 + 2\sqrt{5}$. D'après le théorème de Hartman–Grobman c'est à un homéomorphisme près un point selle (ou col) comme celui représenté sur la figure 7.12.

On cherche ensuite à déterminer les signes des composantes de la fonction f, c'est à dire à résoudre le système d'inéquations :

$$\begin{cases} x_1(1 - x_2 - x_1) \geqslant 0 \\ \qquad x_2(x_1 - 2) \geqslant 0. \end{cases}$$

On est amené à introduire les droites d'équations $1 - x_2 - x_1 = 0$ et $x_1 = 2$. Avec les 2 axes (des ordonnées et des abscisses), ces droites délimitent 10 domaines du plan dans lesquels les composantes de f gardent un signe constant. Ces domaines, numérotés de 1 à 10, sont représentés sur la figure 7.17. Si l'on représente la fonction f sous la forme d'un champ de vecteurs, alors les vecteurs seront orientés comme indiqué sur la figure.

De la figure 7.17, on peut déduire l'allure générales de orbites de l'edo 7.55. Notons (x_1^0,x_2^0) la condition de Cauchy au temps $t = 0$. Sans faire une analyse exhaustive, on peut remarquer par exemple que :

- Toute orbite originaire du domaine ① (c'est à dire telle que que (x_1^0,x_2^0) soit dans le domaine ①) se retrouve au bout d'un temps fini dans le domaine ④ et n'en sort plus.
- Toute orbite originaire du domaine ③ se poursuit dans le domaine ② puis dans le domaine ⑤ et enfin tend vers le point d'équilibre α.
- La situation est moins claire pour les trajectoires provenant du domaine ⑦. Elles peuvent se prolonger soit dans le domaine ⑧ et tendre vers le point d'équilibre α, soit se prolonger dans le domaine ⑩. Sur la figure 7.18, la courbe en rouge divise le domaine ⑦ en deux sous–domaines suivant la destination des trajectoires.
- De même, les orbites originaires du domaine ⑨ se prolongent soit dans le domaine ⑧ et tendent vers le point d'équilibre α, soit se prolongent dans le domaine ⑩. Sur la figure 7.18, la courbe en rouge divise le domaine ⑨ en deux sous–domaines suivant la destination des trajectoires.

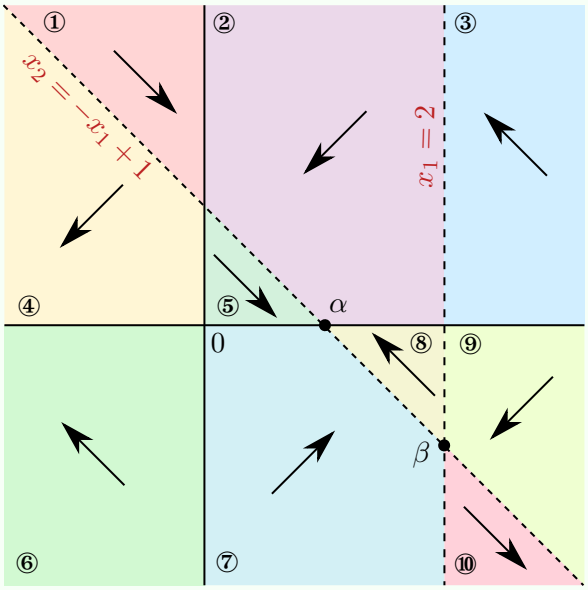


FIG. 7.17 – *Les domaines du plan dans lesquels les composantes de la fonction f gardent un signe constant (délimités par les axes et les droites $x_2 = -x_1 + 1$ et $x_1 = 2$). En représentant f comme un champ de vecteurs, les vecteurs dans chaque domaine sont orientés comme indiqué ci–dessus.*

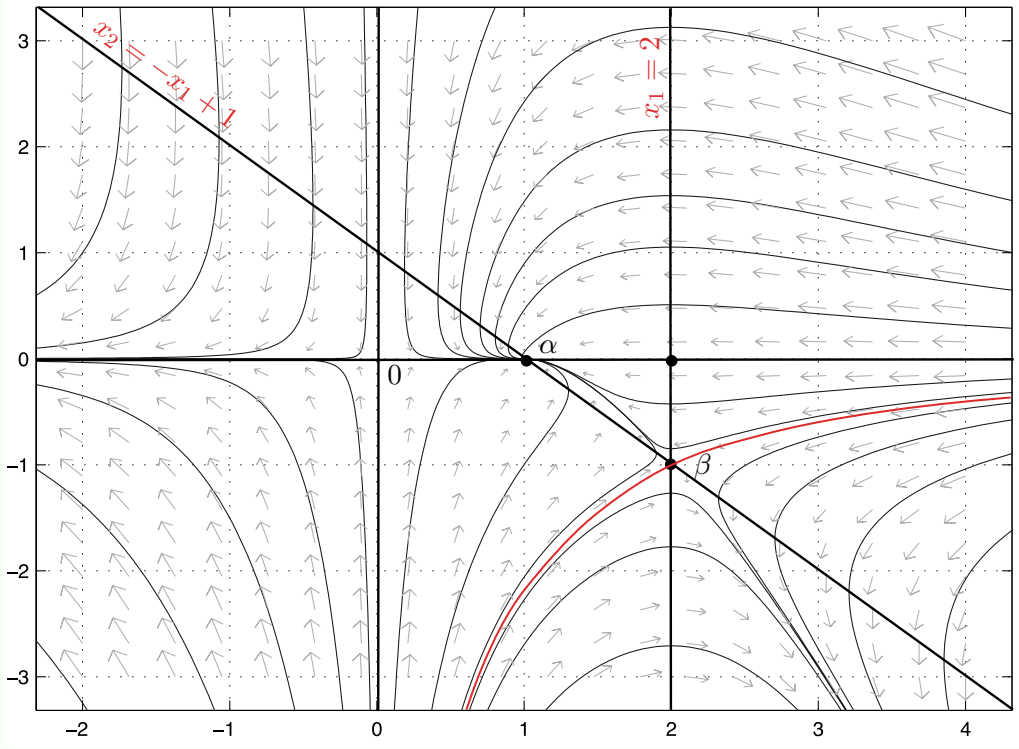


FIG. 7.18 – *Portrait de phase de l'edo autonome* (7.55).