

Analyse II

Partie B



Calcul Différentiel Equations Différentielles

Ernst Hairer

Université de Genève
Section de mathématiques
Case postale 240
CH-1211 Genève 24

Juin 1999

Table des matières

I	Calcul différentiel dans les espaces de Banach	1
I.1	Rappel sur la différentiabilité dans \mathbb{R}^n	1
I.2	Espaces de Banach et de Hilbert	2
I.3	Applications linéaires	4
I.4	Différentiabilité dans les espaces normés	8
I.5	Théorème des accroissements finis	10
I.6	Théorème du point fixe de Banach	12
I.7	Théorème d'inversion locale	17
I.8	Théorème des fonctions implicites	20
I.9	Applications bilinéaires et multilinéaires	22
I.10	Dérivées d'ordre supérieur	23
I.11	Exercices	26
II	Maxima et minima relatifs et calcul de variations	33
II.1	Maxima et minima relatifs	33
II.2	Multiplicateurs de Lagrange	35
II.3	Calcul de variations	38
II.4	Problèmes isopérimétriques	42
II.5	Géodésiques des surfaces	43
II.6	Exercices	47
III	Equations différentielles ordinaires	51
III.1	Terminologie et quelques exemples	51
III.2	Existence et unicité du problème de Cauchy	54
III.3	Théorème de Peano	56
III.4	Prolongement des solutions et existence globale	59
III.5	Inégalités différentielles et solutions approchées	61
III.6	Systèmes d'équations différentielles linéaires	64
III.7	Systèmes linéaires à coefficients constants	67
III.8	Différentiabilité par rapport aux valeurs initiales	69
III.9	Stabilité	72
III.10	Exercices	78

IV	Equations différentielles linéaires d'ordre 2	85
IV.1	Problèmes aux limites – introduction	85
IV.2	La fonction de Green	86
IV.3	Problèmes adjoints et auto-adjoints	88
IV.4	Le problème de Sturm–Liouville	90
IV.5	Théorème de comparaison de Sturm	93
IV.6	Equations différentielles avec des singularités	95
IV.7	Exercices	100
V	Sous-variétés de \mathbb{R}^n	105
V.1	Théorème du rang	105
V.2	Définition des sous-variétés de \mathbb{R}^n	107
V.3	Espace tangent	109
V.4	Equations différentielles sur des sous-variétés	110
V.5	Exercices	111
	Bibliographie	113

Avant-propos

Cette brochure est une version corrigée des polycopiés distribués pendant le cours “Analyse II, Partie B” (2 heures par semaine) donné en 1998/99. Plusieurs parties ont profité des notes de cours de Pierre de la Harpe. Les portraits sur la page de couverture (S. Banach (1892–1945) et D. Hilbert (1862–1943) à gauche, G. Peano (1858–1932) et C.-F. Sturm (1803–1855) à droite) ont été copiés sur Internet par les soins de Stéphane Cirilli (site <http://www-groups.dcs.st-and.ac.uk/~history/Mathematicians/>). Quelques figures (surtout celles avec un flair artistique) ont été produites en utilisant des programmes de Gerhard Wanner. Je remercie ces collègues sincèrement.

En ce lieu, j’aimerais remercier Luc Rey-Bellet et Stéphane Cirilli pour leur aide dans la préparation des exercices et pour leur lecture soignée d’une version préliminaire. Je tiens aussi à remercier mon fils Martin pour avoir vérifié l’orthographe de ces polycopiés.

Bibliographie

- [AMR88] R. Abraham, J.E. Marsden and T. Ratiu, *Manifolds, Tensor Analysis, and Applications*, Appl. Math. Sciences **75**, 2nd edition, Springer-Verlag, 1988.
- [Ar74] V. Arnold, *Equations Différentielles Ordinaires*, Editions Mir (traduction française), Moscou, 1974.
- [A83] A. Avez, *Calcul Différentiel*, Masson, Paris, 1983.
- [C67] H. Cartan, *Calcul Différentiel*, Hermann, Paris, 1967.
- [D63] J. Dieudonné, *Fondements de l'Analyse Moderne*, Gauthier-Villars, Paris, 1963.
- [Fox] C. Fox, *An Introduction to The Calculus of Variations*, Dover Publications New York, 1987.
- [GH96] M. Giaquinta and S. Hildebrandt, *Calculus of Variations I*, Springer Berlin, 1996.
- [HW95] E. Hairer and G. Wanner, *Analysis by Its History*, Undergraduate Texts in Mathematics, Springer New York, 1995.
- [HNW93] E. Hairer, S.P. Nørsett and G. Wanner, *Solving Ordinary Differential Equations. Nons-tiff Problems, 2nd edition*, Springer Series in Computational Mathematics **8**, Springer Berlin, 1993.
- [H64] P. Hartman, *Ordinary Differential Equations*, John Wiley & Sons, New York, 1964.
- [HS74] M.W. Hirsch and S. Smale, *Differential Equations, Dynamical Systems, and Linear Algebra*, Academic Press, 1974.
- [S68] S. Lang, *Analysis II*, Addison-Wesley Publishing Company, Massachusetts, 1968.
- [LS68] L.H. Loomis and S. Sternberg, *Advanced Calculus*, Addison-Wesley Publishing Company, Massachusetts, 1968.
- [Tr83] J.L. Troutman, *Variational Calculus with Elementary Convexity*, Undergraduate Texts in Mathematics, Springer New York, 1983.

Chapitre I

Calcul différentiel dans les espaces de Banach

Le calcul différentiel dans \mathbb{R} et dans \mathbb{R}^n était un des sujets traités au cours ‘Analyse I’ (voir le livre [HW95]). Le premier chapitre du cours ‘Analyse II’ a comme but d’étendre ce calcul aux espaces plus généraux. Ceci nous permet non seulement d’obtenir des résultats plus généraux avec des applications intéressantes, mais nous obtenons en même temps une meilleure compréhension du calcul différentiel dans \mathbb{R}^n .

Après un rappel sur la différentiabilité dans \mathbb{R}^n , nous donnons la définition d’un espace de Banach et nous discutons les différences essentielles entre \mathbb{R}^n et les espaces de Banach de dimension infinie. Nous étendons ensuite la notion d’application différentiable aux espaces de Banach, nous donnons plusieurs exemples, et nous abordons les sujets suivants: le théorème des accroissements finis, le théorème du point fixe de Banach, le théorème d’inversion locale, ainsi que le théorème des fonctions implicites.

I.1 Rappel sur la différentiabilité dans \mathbb{R}^n

Pour une fonction à une variable $f : (a, b) \rightarrow \mathbb{R}$, la dérivée au point x_0 est définie par

$$f'(x_0) = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}. \quad (1.1)$$

De toute évidence, cette définition n’a pas de sens pour des fonctions à plusieurs variables.

Pour une fonction $f : U \rightarrow \mathbb{R}^m$ (U étant un ouvert de \mathbb{R}^n), on dit que $f(x)$ est *différentiable* en $x_0 \in U$, s’il existe une application linéaire $f'(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}^m$, telle que

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + r(x)\|x - x_0\|, \quad (1.2)$$

où la fonction $r : U \rightarrow \mathbb{R}^m$ (qui dépend du paramètre x_0) satisfait $r(x) \rightarrow 0$ pour $x \rightarrow x_0$. On a vu dans [HW95, IV.3] que cette définition ne dépend pas de la norme choisie et que l’application linéaire est donnée par la *matrice jacobienne*

$$f'(x_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_0) & \dots & \frac{\partial f_1}{\partial x_n}(x_0) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(x_0) & \dots & \frac{\partial f_m}{\partial x_n}(x_0) \end{pmatrix}, \quad (1.3)$$

où $x = (x_1, \dots, x_n)^T$ et $f(x) = (f_1(x), \dots, f_m(x))^T$. L’exemple suivant montre qu’il n’est pas toujours avantageux de représenter l’application linéaire $f'(x)$ à l’aide de la matrice jacobienne.

Exemple 1.1 Identifions l'espace $\mathbb{R}^{n \times n}$ avec l'ensemble des matrices carrées de dimension n , et considérons l'application $f : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}$ définie par $f(X) = X^2$. La propriété $(X_0 + H)^2 = X_0^2 + X_0H + HX_0 + H^2$ suggère la définition

$$f'(X_0)H = X_0H + HX_0, \quad r(X) = (X - X_0)^2 / \|X - X_0\|. \quad (1.4)$$

Pour démontrer que l'application linéaire $f'(X_0)$ de (1.4) est vraiment la dérivée de $f(X)$, il faut voir que $r(X) \rightarrow 0$ si $X \rightarrow X_0$. Ceci découle du fait que, pour la norme euclidienne dans $\mathbb{R}^{n \times n}$, on a $\|(X - X_0)^2\| \leq \|X - X_0\|^2$.

I.2 Espaces de Banach et de Hilbert

Essayons d'étendre la définition (1.2) de la différentiabilité à une fonction $f : E \rightarrow F$. Dans les espaces E et F il faut savoir additionner, soustraire, multiplier avec un nombre réel et il faut avoir à disposition une norme. Rappelons qu'une *norme* sur un espace vectoriel E est une application $\|\cdot\| : E \rightarrow \mathbb{R}$ qui vérifie les trois propriétés:

- (N1) $\|x\| \geq 0$ et $\|x\| = 0 \Leftrightarrow x = 0$,
- (N2) $\|\lambda x\| = |\lambda| \cdot \|x\|$,
- (N3) $\|x + y\| \leq \|x\| + \|y\|$ (inégalité du triangle).

Un cas particulier important d'une norme est

$$\|x\| := \sqrt{\langle x, x \rangle}, \quad (2.1)$$

où $\langle \cdot, \cdot \rangle : E \times E \rightarrow \mathbb{R}$ est un *produit scalaire*, c.-à-d.,

- (PS1) $\langle x, x \rangle \geq 0$ et $\langle x, x \rangle = 0 \Leftrightarrow x = 0$ (définie positive),
- (PS2) $\langle x, y \rangle = \langle y, x \rangle$ (symétrique),
- (PS3) $\langle \lambda x_1 + \mu x_2, y \rangle = \lambda \langle x_1, y \rangle + \mu \langle x_2, y \rangle$ (linéaire).

Comme pour la norme euclidienne dans \mathbb{R}^n on vérifie que $\|\cdot\|$, donné par (2.1), satisfait les propriétés (N1), (N2), (N3) d'une norme (Exercice 3).

Dès qu'on a donné une norme sur E , on peut définir la convergence de suites (voir [HW95, IV.1]). On dit qu'une suite $\{x_n\}$ à valeurs dans E *converge vers* $a \in E$, si

$$\forall \varepsilon > 0 \quad \exists N \geq 1 \quad \forall n \geq N \quad \|x_n - a\| < \varepsilon. \quad (2.2)$$

Elle est une *suite de Cauchy*, si

$$\forall \varepsilon > 0 \quad \exists N \geq 1 \quad \forall n, m \geq N \quad \|x_n - x_m\| < \varepsilon. \quad (2.3)$$

Définition 2.1 Un espace vectoriel E muni d'une norme $\|\cdot\|$ s'appelle *espace vectoriel normé*. Il est *complet* si chaque suite de Cauchy dans E est convergente. Un espace vectoriel normé et complet s'appelle un *espace de Banach*.

Un espace de Banach, dont la norme est donnée par un produit scalaire, s'appelle un *espace de Hilbert*.

Le premier exemple d'un espace de Hilbert est l'espace \mathbb{R}^n avec comme norme la norme euclidienne. La complétude est une conséquence du Théorème IV.1.8 de [HW95]. Avec la norme $\|x\|_1 = \sum_{i=1}^n |x_i|$ (ou $\|x\|_\infty = \max_i |x_i|$) l'espace \mathbb{R}^n n'est pas un espace de Hilbert (Exercice 4), mais il est un espace de Banach, car toutes les normes sont équivalentes dans \mathbb{R}^n ([HW95, Théorème IV.2.4]).

Proposition 2.2 Soit $\mathcal{C}([0, 1]) := \{f : [0, 1] \rightarrow \mathbb{R} ; f \text{ est continue}\}$. Avec la norme

$$\|f\|_\infty = \sup_{t \in [0, 1]} |f(t)|, \quad (2.4)$$

$\mathcal{C}([0, 1])$ est un espace de Banach. Par contre, avec une des normes

$$\|f\|_1 = \int_0^1 |f(t)| dt \quad \text{ou} \quad \|f\|_2 = \sqrt{\int_0^1 |f(t)|^2 dt}, \quad (2.5)$$

l'espace $\mathcal{C}([0, 1])$ n'est pas complet.

Remarque. La norme $\|f\|_2$ est définie à partir du produit scalaire $\langle f, g \rangle = \int_0^1 f(t)g(t) dt$, mais $\mathcal{C}([0, 1])$ avec cette norme n'est pas un espace de Hilbert.

Démonstration. Comme dans \mathbb{R}^n , en remplaçant la somme finie par une intégrale, on vérifie que $\|f\|_1$, $\|f\|_2$ et $\|f\|_\infty$ sont des normes sur $\mathcal{C}([0, 1])$.

a) Soit $\{f_n\}_{n \geq 1}$ une suite de Cauchy pour la norme $\|f\|_\infty$. Alors, la propriété

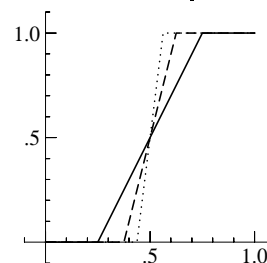
$$|f_n(t) - f_m(t)| \leq \|f_n - f_m\|_\infty < \varepsilon \quad \text{pour } n, m \geq N \quad (2.6)$$

implique que $\{f_n(t)\}_{n \geq 1}$ est une suite de Cauchy dans \mathbb{R} . Comme \mathbb{R} est complet ([HW95, Théorème III.1.8]), la suite $\{f_n(t)\}_{n \geq 1}$ converge vers un élément de \mathbb{R} , qu'on dénote par $f(t)$. Il reste donc à démontrer que la fonction $f(t)$, définie de cette manière, est continue. En passant à la limite $m \rightarrow \infty$ dans (2.6), nous obtenons

$$|f_n(t) - f(t)| \leq \varepsilon \quad \text{pour } n \geq N,$$

où N dépend de $\varepsilon > 0$ mais pas de $t \in [0, 1]$. Alors, la suite $\{f_n\}$ converge uniformément vers f , et la continuité de f est une conséquence de [HW95, Théorème III.4.2].

b) Considérons la suite $\{f_n\}$ dans $\mathcal{C}([0, 1])$, où $f_n(t)$ est linéaire par morceaux, 0 sur $[0, 1/2 - 1/n]$ et 1 sur $[1/2 + 1/n, 1]$ (voir le dessin, $n = 4, 8, 16$). Pour $m \geq n$ on a que $\|f_n - f_m\|_1 < 1/n$ et $\|f_n - f_m\|_2 < 1/\sqrt{n}$. La suite $\{f_n\}$ est alors une suite de Cauchy. Comme la fonction limite $f(t)$ n'est pas continue sur $[0, 1]$, cette suite ne converge pas dans $\mathcal{C}([0, 1])$. \square



Proposition 2.3 Pour un ensemble arbitraire A considérons l'espace

$$\mathcal{B}(A) := \{f : A \rightarrow \mathbb{R} ; f \text{ est bornée}\} \quad \text{avec} \quad \|f\|_\infty = \sup_{t \in A} |f(t)|. \quad (2.7)$$

$\mathcal{B}(A)$ est un espace de Banach.

La démonstration de cette proposition est presque la même que pour la partie (a) de la Proposition 2.2. Nous omettons les détails. \square

Toutes les notions et définitions des paragraphes IV.1 et IV.2 de [HW95] peuvent être étendues aux espaces de Banach:

- Deux normes $\|\cdot\|_p$ et $\|\cdot\|_q$ sont *équivalentes*, s'il existe des constantes positives C_1 et C_2 pour lesquelles $C_1 \|x\|_p \leq \|x\|_q \leq C_2 \|x\|_p$ pour tout $x \in E$.
- Une *boule* de centre a et de rayon r est l'ensemble $B_r(a) = \{x \in E ; \|x - a\| < r\}$.
- Un ensemble $V \subset E$ est un *voisinage* de $a \in E$, s'il existe un $\varepsilon > 0$ tel que $B_\varepsilon(a) \subset V$.
- Un ensemble $U \subset E$ est *ouvert*, si U est voisinage de chacun de ses éléments, c.-à-d., $\forall x \in U \exists \varepsilon > 0 B_\varepsilon(x) \subset U$.
- Un ensemble $V \subset E$ est *fermé*, si la limite de chaque suite convergente $\{x_n\}$ avec $x_n \in V$, est dans V .
- Un ensemble $K \subset E$ est *compact*, si chaque suite $\{x_n\}$ avec $x_n \in K$ possède une sous-suite qui converge vers un élément de K .
- Soient E et F deux espaces normés et $U \subset E$. Une fonction $f : U \rightarrow F$ est *continue* en $x_0 \in U$ si $\forall \varepsilon > 0 \exists \delta > 0 \forall x \in U : \|x - x_0\| < \delta \implies \|f(x) - f(x_0)\| < \varepsilon$.
- Soient E un espace normé. L'application $x \mapsto \|x\|$ est continue (car $|\|x\| - \|x_0\|| \leq \|x - x_0\|$).

La majorité des résultats de [HW95, IV.1 et IV.2] reste vraie si l'on remplace \mathbb{R}^n et \mathbb{R}^m par des espaces de Banach. Néanmoins, il faut faire attention, car certaines propriétés sont perdues si la dimension de l'espace de Banach est infinie (comme c'est le cas pour les exemples des Propositions 2.2 et 2.3). A l'aide de contre-exemples, nous démontrons que:

- La boule fermée $\{x \in E ; \|x\| \leq 1\}$ n'est pas forcément compacte.
- Deux normes dans le même espace ne sont pas toujours équivalentes.
- Le théorème de Bolzano-Weierstrass "*chaque suite bornée possède une sous-suite convergente*" n'est plus vrai.
- La caractérisation " *K compact $\Leftrightarrow K$ fermé et borné*" n'est plus valable.

Considérons l'espace $\mathcal{B}(\mathbb{R})$ de (2.7) et définissons $f_n \in \mathcal{B}(\mathbb{R})$ par $f_n(t) = 1$ si $t \in [n, n+1)$ et par $f_n(t) = 0$ si $t \notin [n, n+1)$. La suite $\{f_n\}$ est bornée (on a $\|f_n\|_\infty \leq 1$) et elle satisfait $\|f_n - f_m\|_\infty = 1$ pour $n \neq m$. Par conséquent, cette suite ne peut pas avoir une sous-suite convergente. Ce contre-exemple montre en même temps que la boule fermée n'est pas compacte, que le théorème de Bolzano-Weierstrass n'est pas vrai sur $\mathcal{B}(\mathbb{R})$, et que la caractérisation ci-dessus de la compacité n'est pas valable.

La Proposition 2.2 montre que dans l'espace $\mathcal{C}([0, 1])$ les normes $\|f\|_\infty$ et $\|f\|_1$ ne sont pas équivalentes. Au cas contraire, la suite $\{f_n\}$ de la partie (b) de la démonstration serait aussi une suite de Cauchy pour la norme $\|\cdot\|_\infty$, ce qui impliquerait la convergence de $\{f_n\}$ vers une fonction continue $f \in \mathcal{C}([0, 1])$.

I.3 Applications linéaires

Les applications linéaires jouent un rôle important dans la définition (1.2) de la différentiabilité. Dans \mathbb{R}^n , elles sont toujours continues, même uniformément continues ([HW95, Théorème IV.2.6]). Nous verrons dans ce paragraphe que ceci n'est pas toujours le cas si la dimension de l'espace vectoriel est infinie.

Proposition 3.1 *Pour une application linéaire $A : E \rightarrow F$ (E et F sont des espaces vectoriels normés), les conditions suivantes sont équivalentes :*

- (a) A est continue en tout point de E ;
- (b) A est continue à l'origine $0 \in E$;
- (c) $\|A(x)\|$ est bornée sur la boule-unité $\{x \in E ; \|x\| \leq 1\}$.

Démonstration. Il est clair que (a) \Rightarrow (b). Montrons que (b) \Rightarrow (c): la continuité de $A(x)$ à l'origine implique que pour $\varepsilon = 1$ il existe un $\delta > 0$ tel que $\|A(y)\| \leq 1$ pour $\|y\| \leq \delta$. En utilisant la linéarité de A , on obtient alors

$$\|A(x)\| = \left\| A\left(\frac{1}{\delta} \delta x\right) \right\| = \frac{1}{\delta} \|A(\delta x)\| \leq \frac{1}{\delta} \quad \text{pour } \|x\| \leq 1.$$

On vient donc de prouver que $\|A(x)\|$ est bornée sur la boule-unité.

Pour démontrer (c) \Rightarrow (a), nous fixons arbitrairement un $x_0 \in E$ et nous supposons que $\|A(x)\|$ est bornée par M sur la boule-unité. Alors, on a pour $x \neq x_0$

$$\|A(x) - A(x_0)\| = \|A(x - x_0)\| = \|x - x_0\| \cdot \left\| A\left(\frac{x - x_0}{\|x - x_0\|}\right) \right\| \leq M \|x - x_0\|,$$

ce qui implique la continuité de A en x_0 . □

Surtout dans les espaces de dimension infinie, nous appelons une application linéaire aussi *opérateur linéaire*. Nous écrivons souvent Ax au lieu de $A(x)$ et nous disons aussi *application bornée* pour une application continue (à cause de la propriété (c) dans la proposition précédente). Une application linéaire A de \mathbb{R}^n dans \mathbb{R}^m est représentée par une matrice, qu'on dénote par la même lettre A . L'expression Ax peut donc être interprétée comme la valeur de l'application A au point x , ou bien comme le produit de la matrice A avec le vecteur x .

Exemple 3.2 Il existe des opérateurs linéaires non-continus. Considérons l'espace $\mathcal{C}([0, 1])$ avec la norme $\|f\|_1$ de (2.5) et les fonctions $f_n \in \mathcal{C}([0, 1])$ données par $f_n(t) = n - n^2 t/2$ si $t \in [0, 2/n]$ et $f_n(t) = 0$ si $t \geq 2/n$. L'application $A : \mathcal{C}([0, 1]) \rightarrow \mathbb{R}$ définie par $A(f) = f(0)$ est linéaire mais non-bornée, car $|A(f_n)| = n$ et $\|f_n\|_1 = 1$ pour tout n .

Définition 3.3 Soient E et F des espaces vectoriels normés. On dénote par $\mathcal{L}(E, F)$ l'ensemble des applications linéaires continues de E dans F . Pour un élément $A \in \mathcal{L}(E, F)$, on définit

$$\|A\| = \sup_{\|x\| \leq 1} \|Ax\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}. \quad (3.1)$$

Pour $E = F$, on écrit aussi $\mathcal{L}(E)$ à la place de $\mathcal{L}(E, E)$.

Cette définition signifie que $\|A\|$ est le plus petit nombre réel tel que

$$\|Ax\| \leq \|A\| \cdot \|x\| \quad \text{pour tout } x \in E. \quad (3.2)$$

L'inégalité (3.2) est fondamentale pour tous les calculs avec des applications linéaires.

Proposition 3.4 *L'espace $\mathcal{L}(E, F)$ muni de (3.1) est un espace vectoriel normé. Si F est complet, alors $\mathcal{L}(E, F)$ est aussi complet.*

Démonstration. Les propriétés (N1) et (N2) d'une norme sont faciles à vérifier. Démontrons (N3): pour $A, B \in \mathcal{L}(E, F)$ nous avons

$$\|(A + B)x\| \leq \|Ax\| + \|Bx\| \leq (\|A\| + \|B\|)\|x\|.$$

En divisant cette relation par $\|x\|$ et en prenant le supremum, nous obtenons l'inégalité du triangle $\|A + B\| \leq \|A\| + \|B\|$.

La démonstration de la complétude est similaire à celle de la partie (a) de la Proposition 2.2. Soit $\{A_n\}$ une suite de Cauchy dans $\mathcal{L}(E, F)$. Pour $\|x\| \leq 1$ on obtient

$$\|A_n x - A_m x\| \leq \|A_n - A_m\| \cdot \|x\| \leq \varepsilon \cdot \|x\| \leq \varepsilon \quad \text{pour } n, m \geq N, \quad (3.3)$$

ce qui implique que $\{A_n x\}$ est une suite de Cauchy dans F . Cet espace étant complet, la suite $\{A_n x\}$ possède une limite qu'on denote par Ax . De cette manière, on obtient une application linéaire $A : E \rightarrow F$. En passant à la limite $m \rightarrow \infty$ dans (3.3) et en divisant par $\|x\|$, on voit que $A_n - A$ (et alors aussi A) est bornée et que A est la limite de $\{A_n\}$. \square

Proposition 3.5 Soit $I \in \mathcal{L}(E)$ l'identité (c.-à-d., $Ix = x$) et considérons des applications linéaires $A \in \mathcal{L}(E, F)$ et $B \in \mathcal{L}(G, E)$. Alors, on a

$$\|I\| = 1, \quad \|AB\| \leq \|A\| \cdot \|B\|. \quad (3.4)$$

Démonstration. La propriété $\|I\| = 1$ est évidente. Pour démontrer l'estimation de $\|AB\|$, nous appliquons deux fois l'inégalité fondamentale (3.2),

$$\|(AB)x\| \leq \|A\| \cdot \|Bx\| \leq \|A\| \cdot \|B\| \cdot \|x\|.$$

Ensuite, nous divisons cette relation par $\|x\|$ et nous prenons le supremum sur $x \neq 0$. \square

Exemple 3.6 Soit A une matrice $m \times n$, c.-à-d., $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, et notons par

$$\|A\|_p = \sup_{x \neq 0} \frac{\|Ax\|_p}{\|x\|_p} \quad (3.5)$$

l'expression (3.1) si l'on utilise la même norme dans les deux espaces \mathbb{R}^n et \mathbb{R}^m . Alors, on a les formules explicites

$$\begin{aligned} \|A\|_1 &= \max_{j=1, \dots, n} \left(\sum_{i=1}^m |a_{ij}| \right), & \|A\|_\infty &= \max_{i=1, \dots, m} \left(\sum_{j=1}^n |a_{ij}| \right), \\ \|A\|_2 &= \sqrt{\text{plus grande valeur propre de } A^T A}. \end{aligned}$$

Démonstration. Pour la norme $\|x\|_1$, on a

$$\|Ax\|_1 = \sum_{i=1}^m \left| \sum_{j=1}^n a_{ij} x_j \right| \leq \sum_{i=1}^m \sum_{j=1}^n |a_{ij}| \cdot |x_j| = \sum_{j=1}^n \left(\sum_{i=1}^m |a_{ij}| \right) |x_j| \leq \max_{j=1, \dots, n} \left(\sum_{i=1}^m |a_{ij}| \right) \cdot \|x\|_1.$$

On en déduit que $\|A\|_1 \leq \max_j (\sum_i |a_{ij}|)$. Pour montrer l'égalité, on choisit un j_0 avec $\max_j (\sum_i |a_{ij}|) = \sum_i |a_{ij_0}|$ et on pose $x = (0, \dots, 0, 1, 0, \dots, 0)^T$, où 1 est à la position j_0 . Avec ce choix de x , on a égalité dans l'estimation ci-dessus, ce qui démontre que $\|A\|_1$ ne

peut pas être plus petit que $\max_j(\sum_i |a_{ij}|)$. La formule pour la norme $\|x\|_\infty$ se démontre de la même manière.

La matrice $A^T A$ étant symétrique et semi-définie positive ($x^T A^T A x = \|Ax\|_2^2 \geq 0$), il existe une matrice orthogonale U ($U^T U = I$) telle que $U^T A^T A U = \text{diag}(\lambda_1, \dots, \lambda_n)$, où $\lambda_i \geq 0$ sont les valeurs propres de $A^T A$. On obtient alors avec la transformation $x = Uy$ et en utilisant $\|x\|_2 = \|y\|_2$ que

$$\|Ax\|_2^2 = x^T A^T A x = y^T U^T A^T A U y = \sum_{i=1}^n \lambda_i |y_i|^2 \leq \lambda_{\max} \|y\|_2^2 = \lambda_{\max} \|x\|_2^2.$$

Ceci implique $\|A\|_2 \leq \sqrt{\lambda_{\max}}$. Pour montrer l'égalité, on pose x égal au vecteur propre de $A^T A$ qui correspond à la valeur propre λ_{\max} . \square

Exemple 3.7 Pour la matrice

$$A = \begin{pmatrix} 4 & 3 \\ -2 & 5 \\ 4 & 1 \end{pmatrix} \quad \text{on a} \quad \begin{aligned} \|A\|_1 &= \max(10, 9) = 10 \\ \|A\|_2 &= \dots = \sqrt{(71 + \sqrt{145})}/2 \approx 6.4437 \\ \|A\|_\infty &= \max(7, 7, 5) = 7. \end{aligned}$$

Exemple 3.8 Soit $k : [0, 1] \times [0, 1] \rightarrow \mathbb{R}$ une fonction continue à deux variables. Nous considérons l'opérateur linéaire $A : \mathcal{C}([0, 1]) \rightarrow \mathcal{C}([0, 1])$, défini par

$$(Af)(t) = \int_0^1 k(t, s) f(s) ds, \tag{3.6}$$

et les normes de la Proposition 2.2. Avec la notation (3.5) nous avons

$$\begin{aligned} \|A\|_1 &= \max_{s \in [0, 1]} \int_0^1 |k(t, s)| dt, & \|A\|_\infty &= \max_{t \in [0, 1]} \int_0^1 |k(t, s)| ds, \\ \|A\|_2 &\leq \sqrt{\int_0^1 \int_0^1 |k(t, s)|^2 ds dt}. \end{aligned}$$

Les formules pour $\|A\|_1$ et $\|A\|_\infty$ sont obtenues comme dans l'Exemple 3.6, l'estimation pour $\|A\|_2$ comme dans [HW95, Théorème IV.2.6].

Proposition 3.9 Soit E un espace de Banach et supposons que l'opérateur $A \in \mathcal{L}(E)$ satisfasse $\|A\| < 1$. Alors, $I - A$ est inversible, $(I - A)^{-1}$ est continue, et on a

$$(I - A)^{-1} = I + A + A^2 + A^3 + \dots \tag{3.7}$$

(série géométrique ou "série de Neumann").

Démonstration. Montrons d'abord que $A_n := I + A + A^2 + \dots + A^n$ est une suite de Cauchy dans $\mathcal{L}(E)$. En utilisant $\|A^n\| \leq \|A\|^n$, on obtient pour $m > n$ que

$$\|A_n - A_m\| = \|A^{n+1} + \dots + A^m\| \leq \|A^{n+1}\| + \dots + \|A^m\| \leq \|A\|^{n+1} + \dots + \|A\|^m \leq \frac{\|A\|^{n+1}}{1 - \|A\|}$$

(observons que $\|A\|^{n+1} \rightarrow 0$ si $n \rightarrow \infty$). Comme $\mathcal{L}(E)$ est complet (voir la Proposition 3.4), la suite $\{A_n\}$ converge vers un $B \in \mathcal{L}(E)$. En passant à la limite $n \rightarrow \infty$ dans l'identité $A_n(I - A) = (I - A)A_n = I - A^{n+1}$, nous obtenons $B(I - A) = (I - A)B = I$, ce qui démontre l'inversibilité de $I - A$ et la formule (3.7). \square

I.4 Différentiabilité dans les espaces normés

Nous donnons maintenant une généralisation directe de la définition (1.2).

Définition 4.1 Soient E, F deux espaces vectoriels normés et $U \subset E$ un ouvert. On dit que $f : U \rightarrow F$ est *différentiable* en $a \in U$ s'il existe une application linéaire continue $f'(a) \in \mathcal{L}(E, F)$ telle que

$$f(x) = f(a) + f'(a)(x - a) + r(x)\|x - a\|, \quad (4.1)$$

où la fonction $r : U \rightarrow F$ satisfait $r(x) \rightarrow 0$ pour $x \rightarrow a$.¹

Par rapport à la définition usuelle dans \mathbb{R}^n , on demande en plus que l'application linéaire $f'(a)$ soit continue. Ceci est important, car sinon une fonction différentiable pourrait ne pas être continue (voir les exemples ci-dessous).

Si $f : E \rightarrow F$ est linéaire et continue, on a $f'(a) = f$ quelque soit $a \in E$. Une application linéaire non-bornée n'est pas différentiable (e.g., l'opérateur de l'Exemple 3.2).

Exemple 4.2 Soient $E = \mathbb{R}^3$, $F = \mathbb{R}^2$ et $f : E \rightarrow F$ donnée par

$$f(x) = \begin{pmatrix} f_1(x_1, x_2, x_3) \\ f_2(x_1, x_2, x_3) \end{pmatrix} = \begin{pmatrix} x_1^2 + x_2^2 + x_3^2 \\ x_1 x_2 x_3 \end{pmatrix}.$$

Sa dérivée au point $a = (a_1, a_2, a_3)^T$ est l'application linéaire $f'(a) \in \mathcal{L}(\mathbb{R}^3, \mathbb{R}^2)$ définie par la matrice

$$f'(a) = \begin{pmatrix} 2a_1 & 2a_2 & 2a_3 \\ a_2 a_3 & a_1 a_3 & a_1 a_2 \end{pmatrix}.$$

On vérifie aisément que l'application $r(x)$ de (4.1) satisfait $r(x) \rightarrow 0$ si $x \rightarrow a$. Comme toutes les normes sont équivalentes sur \mathbb{R}^n , il ne faut pas préciser la norme avec laquelle on travaille. De plus, une application linéaire de \mathbb{R}^n dans \mathbb{R}^m est automatiquement continue.

Exemple 4.3 Considérons l'espace $\mathcal{C}([0, 1])$ avec la norme $\|\cdot\|_\infty$, notons ses éléments par $x(t)$ ou $a(t)$, et étudions l'application $f : \mathcal{C}([0, 1]) \rightarrow \mathcal{C}([0, 1])$,

$$f(x)(t) = \int_0^1 k(t, s)g(x(s)) ds, \quad (4.2)$$

où $k : [0, 1] \times [0, 1] \rightarrow \mathbb{R}$ est continue et $g : \mathbb{R} \rightarrow \mathbb{R}$ est 2 fois continûment différentiable. Pour une fonction $a \in \mathcal{C}([0, 1])$ donnée, cherchons la dérivée $f'(a)$. Avec $h \in \mathcal{C}([0, 1])$ on a

$$\begin{aligned} (f(a+h) - f(a))(t) &= \int_0^1 k(t, s)(g(a(s) + h(s)) - g(a(s))) ds \\ &= \int_0^1 k(t, s)(g'(a(s)) + \frac{1}{2}g''(\alpha(s))h(s))h(s) ds, \end{aligned}$$

où $\alpha(s)$ est une valeur entre $a(s)$ et $a(s) + h(s)$. Ce calcul montre que l'application linéaire $f'(a) : \mathcal{C}([0, 1]) \rightarrow \mathcal{C}([0, 1])$, définie par

$$(f'(a)h)(t) := \int_0^1 k(t, s)g'(a(s))h(s) ds,$$

est un bon candidat pour la dérivée de f . Cette application est continue (voir Exemple 3.8). De plus, le reste $\|f(a+h) - f(a) - f'(a)h\|_\infty$ peut être estimé par $Const\|h\|_\infty^2$, car $g''(x)$ est bornée dans l'intervalle contenant $a(s)$ et $a(s) + h(s)$ pour tout $s \in [0, 1]$ et pour toute h avec $\|h\|_\infty \leq 1$. Ceci démontre que la fonction non-linéaire (4.2) est différentiable en a .

1. Il est possible d'étendre la caractérisation de Carathéodory [HW95, Lemma IV.3.5], mais elle nécessite la connaissance du Théorème de Hahn-Banach de l'analyse fonctionnelle.

Définition 4.4 Une application $A : E \rightarrow F$ (où E et F sont des espaces vectoriels normés) est un *isomorphisme*, si A est linéaire, bijective, continue et si l'inverse A^{-1} est continu.² On dénote

$$\text{GL}(E, F) := \{A \in \mathcal{L}(E, F) ; A \text{ est un isomorphisme}\}. \quad (4.3)$$

Pour $E = F$, on écrit aussi $\text{GL}(E)$ à la place de $\text{GL}(E, E)$.³

Lemme 4.5 Si E et F sont des espaces de Banach, l'ensemble $\text{GL}(E, F)$ est ouvert dans $\mathcal{L}(E, F)$.

Démonstration. Pour $A \in \text{GL}(E, F)$ l'application $A + H$ est un isomorphisme pour tout H avec $\|H\| \leq 1/\|A^{-1}\|$. Ceci découle de $A + H = A(I + A^{-1}H)$, de $\|A^{-1}H\| \leq \|A^{-1}\| \|H\| < 1$, et de la Proposition 3.9. \square

Exemple 4.6 Considérons l'application $f(X) = X^{-1}$ de $\text{GL}(E)$ dans $\mathcal{L}(E)$ (où E est un espace de Banach). Avec la série de Neumann (Proposition 3.9) nous obtenons

$$(A + H)^{-1} = (I + A^{-1}H)^{-1}A^{-1} = A^{-1} - A^{-1}HA^{-1} + \sum_{j \geq 2} (-A^{-1}H)^j A^{-1},$$

dont la partie linéaire est

$$f'(A)H = -A^{-1}HA^{-1}. \quad (4.4)$$

La continuité de $f'(A)$ est une conséquence de $\|f'(A)H\| \leq \|A^{-1}\|^2 \|H\|$ (en utilisant la Proposition 3.1). Le reste $\sum_{j \geq 2} (-A^{-1}H)^j A^{-1}$ peut être estimé par $\|H\|^2 \|A^{-1}\|^3 / (1 - \|A^{-1}H\|)$. Divisé par $\|H\|$, il tend vers 0 si $\|H\| \rightarrow 0$. Ceci démontre la différentiabilité de $f(X) = X^{-1}$ ainsi que la formule (4.4).

Pour deux espaces vectoriels E_1, E_2 de normes respectives $\|\cdot\|_1, \|\cdot\|_2$, considérons le *produit cartésien* $E_1 \times E_2$. Si on le munit de la norme

$$\|(x_1, x_2)\| := \|x_1\|_1 + \|x_2\|_2, \quad (4.5)$$

on obtient un espace vectoriel normé. C'est un espace de Banach si les espaces E_1 et E_2 sont des espaces de Banach.

Exemple 4.7 Soit $B : E_1 \times E_2 \rightarrow F$ une *application bilinéaire bornée*, c.-à-d., elle satisfait

$$\|B(x_1, x_2)\| \leq M \|x_1\|_1 \|x_2\|_2 \quad \text{pour tout } (x_1, x_2) \in E_1 \times E_2. \quad (4.6)$$

Une telle application est différentiable et sa dérivée est donnée par

$$B'(a_1, a_2)(h_1, h_2) = B(a_1, h_2) + B(h_1, a_2). \quad (4.7)$$

La continuité de $B'(a_1, a_2)$ résulte de $\|B'(a_1, a_2)(h_1, h_2)\| \leq M(\|a_1\|_1 \|h_2\|_2 + \|h_1\|_1 \|a_2\|_2) \leq M \max(\|a_1\|_1, \|a_2\|_2) \|(h_1, h_2)\|$. La différentiabilité de B est alors une conséquence de

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

car $\|B(h_1, h_2)\| \leq M \|h_1\|_1 \|h_2\|_2 \leq M(\|h_1\|_1 + \|h_2\|_2)^2 / 4 = M \|(h_1, h_2)\|^2 / 4$.

2. La continuité de A^{-1} est une conséquence des hypothèses sur A . La démonstration de ce résultat est difficile et utilise le "théorème sur les applications ouvertes" (analyse fonctionnelle).

3. GL est une abréviation pour "general linear group".

Résumons quelques règles du calcul différentiel:

- *Linéarité de la dérivée.* Soient f, g deux applications $U \rightarrow F$ (U un ouvert de E) différentiables en $a \in U$, et soit $\lambda \in \mathbb{R}$. Alors, on a

$$(f + g)'(a) = f'(a) + g'(a), \quad (\lambda f)'(a) = \lambda f'(a). \quad (4.8)$$

- *Dérivée d'une fonction composée.* Soient E, F, G trois espaces vectoriels normés, soit U un ouvert de E , et soit V un ouvert de F . On considère deux applications, $f : U \rightarrow F$ différentiable en $a \in U$ et $g : V \rightarrow G$ différentiable en $b = f(a) \in V$ (on suppose $f(U) \subset V$). Alors, $g \circ f$ est différentiable et on a

$$(g \circ f)'(a) = g'(f(a)) \circ f'(a). \quad (4.9)$$

- *Formule de Leibniz.* Soient E, F_1, F_2, G des espaces vectoriels normés, U un ouvert de E , $f : U \rightarrow F_1$ et $g : U \rightarrow F_2$ des applications différentiables en $a \in U$, et $B : F_1 \times F_2 \rightarrow G$ une application bilinéaire bornée (c.-à-d., elle satisfait (4.6)). Alors, la fonction $p(x) := B(f(x), g(x))$ est différentiable en a et on a

$$p'(a)h = B(f'(a)h, g(a)) + B(f(a), g'(a)h). \quad (4.10)$$

La démonstration de (4.8) est triviale et celle de (4.9) est la même que pour des fonctions dans \mathbb{R}^n . La formule (4.10) est une conséquence de (4.7) et de (4.9), car $p = B \circ d$ où $d : U \rightarrow F_1 \times F_2$ est donnée par

$$d(x) = (f(x), g(x)) \quad \text{avec pour dérivée} \quad d'(x)h = (f'(x)h, g'(x)h).$$

I.5 Théorème des accroissements finis

La terminologie des «accroissements finis» s'explique par des raisons historiques: la notion d'accroissements «finis» s'oppose à celle d'accroissements «infinitésimaux» ... (H. Cartan 1967)

Pour une fonction $f : [a, b] \rightarrow \mathbb{R}$, le théorème des accroissements finis (ou théorème de Lagrange) affirme qu'il existe un $\xi \in (a, b)$ tel que

$$f(b) - f(a) = f'(\xi)(b - a),$$

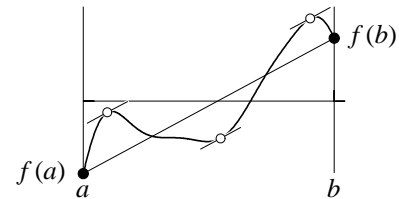
si f est continue sur $[a, b]$ et différentiable sur (a, b) .

Pour une fonction $f : [a, b] \rightarrow \mathbb{R}^m$, l'affirmation sous cette forme n'est plus vraie. Un contre-exemple (déjà donné dans [HW95, IV.3]) est la fonction $f(x) = (f_1(x), f_2(x))^T$ où $f_1(x) = \cos x$, $f_2(x) = \sin x$, et $[a, b] = [0, 2\pi]$. On a que $f(a) = f(b)$, mais il n'y existe pas de ξ avec $f'(\xi) = 0$. Par contre, on voit que l'inégalité $\|f(b) - f(a)\| \leq \|f'(\xi)\| \cdot \|b - a\|$ reste vraie dans cet exemple.

Dans un premier théorème, nous considérons des fonctions $f : [a, b] \rightarrow F$ où F est un espace vectoriel normé. La démonstration va être différente de celle de [HW95, Théorème IV.3.7], car nous n'avons pas de produit scalaire à disposition dans F . Ensuite, nous étendons le résultat à des fonctions $f : U \rightarrow F$, où $U \subset E$, et E est un autre espace vectoriel normé.

Lemme 5.1 Soient $a < b$ deux réels, F un espace vectoriel normé, $f : [a, b] \rightarrow F$ et $g : [a, b] \rightarrow \mathbb{R}$ deux applications continues sur $[a, b]$ et différentiables sur (a, b) . Supposons que $\|f'(t)\| \leq g'(t)$ pour $a < t < b$. Alors,

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



Démonstration. Pour un $\varepsilon > 0$ donné, considérons l'inégalité

$$\|f(t) - f(a)\| \leq g(t) - g(a) + \varepsilon(t - a) + \varepsilon. \quad (5.1)$$

Elle est satisfaite pour $t = a$ et aussi, par continuité de f et g , dans un voisinage de a . Donc, le suprémum

$$c := \sup \{t \in [a, b] ; (5.1) \text{ est satisfaite} \}$$

existe et on a que $c > a$.

Montrons que l'inégalité (5.1) est satisfaite pour c . Par définition du suprémum, il existe une suite $\{t_n\}$, $t_n \rightarrow c$, telle que $\|f(t_n) - f(a)\| \leq g(t_n) - g(a) + \varepsilon(t_n - a) + \varepsilon$. En passant à la limite $n \rightarrow \infty$, la continuité de f et g montre l'inégalité (5.1) pour $t = c$.

Supposons que $c < b$. La différentiabilité de f et g au point c implique qu'il existe un $\eta > 0$ tel que

$$\begin{aligned} \|f(t) - f(c)\| &\leq \|f'(c)\|(t - c) + \frac{\varepsilon}{2}(t - c) \\ g(t) - g(c) &\geq g'(c)(t - c) - \frac{\varepsilon}{2}(t - c) \end{aligned}$$

pour tout $t \in [c, c + \eta]$. L'hypothèse $\|f'(c)\| \leq g'(c)$ implique alors que

$$\|f(t) - f(c)\| \leq g'(c)(t - c) + \frac{\varepsilon}{2}(t - c) \leq g(t) - g(c) + \varepsilon(t - c).$$

Ceci, en plus de l'inégalité (5.1) pour $t = c$, implique que

$$\begin{aligned} \|f(t) - f(a)\| &\leq \|f(t) - f(c)\| + \|f(c) - f(a)\| \\ &\leq g(t) - g(c) + \varepsilon(t - c) + g(c) - g(a) + \varepsilon(c - a) + \varepsilon \\ &= g(t) - g(a) + \varepsilon(t - a) + \varepsilon \end{aligned}$$

pour tout $t \in [c, c + \eta]$, ce qui contredit la définition de c . Alors, on a $c = b$. Si l'on laisse tendre ε vers 0 dans (5.1) avec $t = b$, on obtient l'affirmation du théorème. \square

Considérons maintenant la situation où f est définie sur un ouvert U d'un espace vectoriel normé E , qui n'est plus nécessairement \mathbb{R} . Pour $a, b \in E$, on appelle *segment* d'extrémités a et b l'ensemble des points $x \in E$ de la forme $x = a + t(b - a)$ avec $0 \leq t \leq 1$.

Théorème 5.2 (Théorème des accroissements finis) *Soient E, F des espaces vectoriels normés et $U \subset E$ un ouvert. Si $f : U \rightarrow F$ est différentiable dans U , et si le segment d'extrémités a et b est contenu dans U , on a*

$$\|f(b) - f(a)\| \leq \sup_{0 < t < 1} \|f'(a + t(b - a))\| \cdot \|b - a\|. \quad (5.2)$$

Démonstration. L'application $h(t) := f(a + t(b - a))$ est une application différentiable de $[0, 1]$ dans F et on a $h'(t) = f'(a + t(b - a))(b - a)$, donc

$$\|h'(t)\| \leq \|f'(a + t(b - a))\| \|b - a\|.$$

Il suffit alors d'appliquer le Lemme 5.1, en remplaçant a par 0, b par 1, f par h , et $g(t)$ par Mt où $M = \sup_{0 < s < 1} \|f'(a + s(b - a))\| \|b - a\|$. \square

On dit qu'un sous-ensemble D d'un espace vectoriel est *convexe* si, quels que soient $a, b \in D$, le segment $\{a + t(b - a); t \in [0, 1]\}$ est dans D .

Corollaire 5.3 Soient E, F des espaces vectoriels normés, $U \subset E$ un ouvert, $f : U \rightarrow F$ différentiable dans U , et $D \subset U$ un ensemble convexe. Alors,

$$\|f(x) - f(y)\| \leq \sup_{z \in D} \|f'(z)\| \cdot \|x - y\| \quad \text{pour } x, y \in D.$$

Démonstration. Conséquence immédiate du Théorème 5.2. □

Un ouvert U d'un espace vectoriel normé est dit *connexe*, si deux points quelconques de U peuvent être joints par une ligne brisée dans U . Rappelons qu'une *ligne brisée* est l'union fini de segments $\cup_{i=1}^m \{a_{i-1} + t(a_i - a_{i-1}); t \in [0, 1]\}$.

Corollaire 5.4 Soient E, F des espaces vectoriels normés, U un ouvert connexe, et $f : U \rightarrow F$ une application différentiable dans U . Si $f'(x) = 0$ pour tout $x \in U$, alors f est constante.

Démonstration. Fixons $a \in U$ et prenons $x \in U$ arbitraire. Comme U est connexe, il existe une suite finie $a = a_0, a_1, \dots, a_m = x$ telle que les segments $\{a_{i-1} + t(a_i - a_{i-1}); t \in [0, 1]\}$ sont dans U . Le Théorème 5.2 implique alors que

$$\|f(a_i) - f(a_{i-1})\| \leq \sup_{0 < t < 1} \|f'(a_{i-1} + t(a_i - a_{i-1}))\| \cdot \|a_i - a_{i-1}\| = 0.$$

Donc, $f(a_i) = f(a_{i-1})$ et par conséquent aussi $f(x) = f(a)$. □

I.6 Théorème du point fixe de Banach

Considérons le problème de résoudre une équation nonlinéaire dans un espace de Banach. Pour des fonctions f et g de E dans E , le problème peut être formulé de la manière suivante:

- $g(x) = 0$, on cherche un zéro de g , ou
- $f(x) = x$, on cherche un point fixe de f .

En posant $f(x) = x + g(x)$ ou $f(x) = x + Ag(x)$ avec $A \in \text{GL}(E)$, on peut passer d'une formulation à l'autre. Le théorème suivant donne une condition suffisante pour l'existence et l'unicité d'une solution à ce problème.

Théorème 6.1 (Théorème du point fixe de Banach, 1922) Soient E un espace de Banach, $D \subset E$ fermé, et $f : D \rightarrow E$ une application satisfaisant

- (a) $f(D) \subset D$,
- (b) f est une contraction sur D , c.-à-d., il existe un $\alpha < 1$ tel que

$$\|f(x) - f(y)\| \leq \alpha \|x - y\| \quad \text{pour } x, y \in D. \quad (6.1)$$

Alors, f possède un unique point fixe dans D .

Remarques. La condition (6.1) implique que f est uniformément continue sur D . Inversément, si f est différentiable dans un voisinage de D , si D est convexe, et si $\sup_{x \in D} \|f'(x)\| < 1$, alors f est une contraction. Ceci est une conséquence du Corollaire 5.3.

Démonstration. Unicité. Soient x et y deux points fixes, c.-à-d., $f(x) = x$ et $f(y) = y$. La contractivité implique que

$$\|x - y\| = \|f(x) - f(y)\| \leq \alpha \|x - y\|$$

avec $\alpha < 1$, ce qui est possible seulement si $x = y$.

Existence. Prenons $x_0 \in D$ arbitraire et considérons l'itération $x_{n+1} = f(x_n)$. L'hypothèse $f(D) \subset D$ implique que $x_n \in D$ pour tout $n \geq 0$. Montrons que $\{x_n\}$ est une suite de Cauchy. On a que $\|x_{n+1} - x_n\| = \|f(x_n) - f(x_{n-1})\| \leq \alpha \|x_n - x_{n-1}\|$ et, en appliquant cette inégalité itérativement, on obtient que

$$\|x_{n+1} - x_n\| \leq \alpha^n \|x_1 - x_0\|.$$

Pour $m \geq n$ on en déduit que

$$\begin{aligned} \|x_m - x_n\| &\leq \|x_m - x_{m-1}\| + \|x_{m-1} - x_{m-2}\| + \dots + \|x_{n+1} - x_n\| \\ &\leq (\alpha^{m-1} + \alpha^{m-2} + \dots + \alpha^n) \|x_1 - x_0\| \leq \frac{\alpha^n}{1 - \alpha} \|x_1 - x_0\|. \end{aligned}$$

Alors, $\{x_n\}$ est une suite de Cauchy (observons que $\alpha^n \rightarrow 0$). Comme l'espace E est complet, cette suite converge vers un $x \in E$. La limite x est dans D , car D est fermé. En prenant la limite $n \rightarrow \infty$ dans $x_{n+1} = f(x_n)$ et en utilisant la continuité de f , on obtient $x = f(x)$, c.-à-d., x est un point fixe de f . \square

La démonstration du théorème du point fixe de Banach est constructive et nous conduit à l'algorithme suivant.

Méthode des approximations successives Pour résoudre un problème $x = f(x)$ dans un espace de Banach, cette méthode est définie par:

- on choisit x_0 arbitrairement,
- on applique l'itération $x_{n+1} = f(x_n)$.

Sous les hypothèses du théorème, cet algorithme converge vers la solution unique du problème. Souvent, les hypothèses sont difficiles à vérifier, mais on peut quand même appliquer cet algorithme. Si on a convergence, on est sûr d'avoir trouvé une solution (si f est continue). Elle n'est pas nécessairement unique.

Exemple 6.2 Prenons la fonction $f(x) = \cos x$ sur $D = [0, 1]$. Cette fonction est une contraction, car $|f'(x)| = |\sin x| \leq \sin 1 < 1$ pour $x \in D$. Un autre exemple est la fonction $f(x) = e^x/4$ sur $D = [0, 1.1]$. Pour cette fonction on a $|f'(x)| = e^x/4 \leq e^{1.1}/4 < 1$ sur D . Les itérations sont illustrées dans la Fig. 6.1.

Exemple 6.3 Considérons le système de deux équations nonlinéaires à deux inconnues

$$x = 2 + (x^2 + y^2)/20, \quad y = 1 - xy^3/10.$$

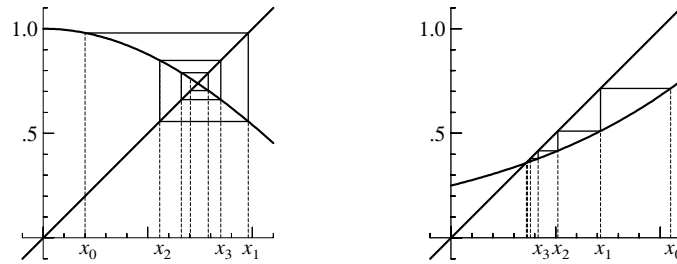


FIG. 6.1 – Méthode des approximations successives pour $f(x) = \cos x$ et $f(x) = e^x/4$

Sans vérifier les hypothèses du théorème du point fixe de Banach, appliquons la méthode itérative avec comme valeurs initiales $x_0 = 2, y_0 = 1$. On obtient alors

$$\begin{aligned} x_1 &= 2.2500000, & y_1 &= 0.8000000, \\ x_2 &= 2.2851250, & y_2 &= 0.8848000, \\ x_3 &= 2.3002333, & y_3 &= 0.8417129, \\ x_4 &= 2.2999777, & y_4 &= 0.8628284, \end{aligned}$$

et on observe convergence vers la solution $x = 2.3014505, y = 0.8557662$ du système.

Exemple 6.4 Considérons l'espace de Banach $\mathcal{C}([0, 1])$ avec la norme $\|\cdot\|_\infty$, et le problème à point fixe

$$y(t) = f(t) + \lambda \int_0^1 k(t, s)y(s) ds, \quad (6.2)$$

où f et k sont données et λ est tel que $|\lambda| \cdot \max_{0 \leq t \leq 1} \int_0^1 |k(t, s)| ds < 1$ (voir l'Exemple 3.8). La méthode des approximations successives s'écrit comme suit: on prend une approximation initiale, par exemple $y_0(t) = f(t)$, et on itère selon

$$y_{n+1}(t) = f(t) + \lambda \int_0^1 k(t, s)y_n(s) ds.$$

La suite $\{y_n(t)\}$ converge vers la solution unique de (6.2).

Considérons maintenant le problème de résoudre $f(x) = y$ pour un vecteur donné y , et cherchons des résultats sur l'existence et l'unicité (locale) d'une solution, c.-à-d., nous cherchons des résultats sur la bijectivité de la fonction f . De plus, on aimerait savoir si la solution dépend continûment du paramètre y . Ceci est résumé dans la définition suivante.

Définition 6.5 Soient E, F des espaces vectoriels normés, $U \subset E$ et $V \subset F$. Une application $f : U \rightarrow V$ est un *homéomorphisme* de U sur V , si f est bijective, continue, et si l'inverse $f^{-1} : V \rightarrow U$ est continue.

Méthode de Newton simplifiée Supposons $f(a) = b$ et considérons le problème $f(x) = y$ pour un y donné proche de b . Soit x_0 une approximation de la solution cherchée (par exemple $x_0 = a$). L'idée est de linéariser le problème autour de x_0 et de résoudre le problème linéarisé pour obtenir une meilleure approximation x_1 :

$$f(x_0) + f'(x_0)(x_1 - x_0) = y \quad \text{ou} \quad x_1 = x_0 - f'(x_0)^{-1}(f(x_0) - y).$$

En itérant cette procédure, on obtient la *méthode de Newton*

$$x_{n+1} = x_n - f'(x_n)^{-1}(f(x_n) - y)$$

(voir Fig.6.2). Elle peut être interprétée comme la méthode des approximations successives appliquée à $x = g(x)$ où $g(x) = x - f'(x)^{-1}(f(x) - y)$.

En pratique, on remplace souvent la dérivée $f'(x_n)$ par une approximation A qui ne dépend pas de x_n . Ceci simplifie et le calcul numérique et la théorie (étude de convergence, voir le théorème suivant). Nous considérons alors la *méthode de Newton simplifiée*, qui s'écrit comme $x_{n+1} = g(x_n)$ où

$$g(x) := x - A^{-1}(f(x) - y). \quad (6.3)$$

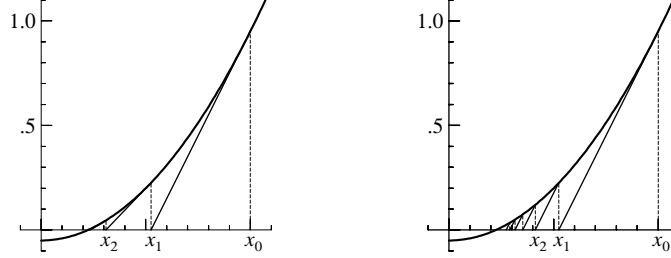


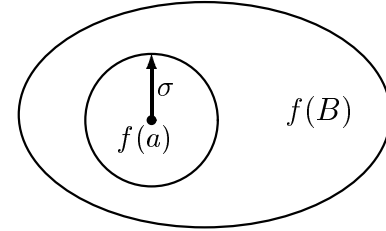
FIG. 6.2 – Méthode de Newton (gauche) et méthode de Newton simplifiée (droite)

Proposition 6.6 Soient E et F des espaces de Banach, $B := \{x \in E; \|x - a\| \leq \rho\}$, $A : E \rightarrow F$ un isomorphisme, et supposons que $f : B \rightarrow F$ satisfasse

$$\|x - z - A^{-1}(f(x) - f(z))\| \leq \alpha \|x - z\| \quad \text{pour } x, z \in B \quad (6.4)$$

avec $\alpha < 1$. Alors, on a

- f est un homéomorphisme de B sur $f(B)$,
- avec $\sigma := \rho(1 - \alpha)/\|A^{-1}\|$ on a $\{y \in F; \|y - f(a)\| \leq \sigma\} \subset f(B)$,
- pour $y \in F$ satisfaisant $\|y - f(a)\| \leq \sigma$, l'itération $x_0 = a$, $x_{n+1} = g(x_n)$ (méthode de Newton simplifiée) converge vers la solution de $f(x) = y$.



Démonstration. a) L'application $g(x)$ de (6.3) est une contraction sur B par notre hypothèse. La continuité de $f(x)$ est donc une conséquence de

$$\|f(x) - f(z)\| = \|A(x - z) - A(g(x) - g(z))\| \leq \|A\|(1 + \alpha)\|x - z\|.$$

De la définition de $g(x)$, nous déduisons $x - z = g(x) - g(z) + A^{-1}(f(x) - f(z))$ et nous obtenons l'estimation $\|x - z\| \leq \alpha \|x - z\| + \|A^{-1}\| \|f(x) - f(z)\|$. On a donc

$$\|x - z\| \leq \frac{\|A^{-1}\|}{1 - \alpha} \|f(x) - f(z)\| \quad \text{pour } x, z \in B. \quad (6.5)$$

La fonction $f : B \rightarrow f(B)$ est surjective par définition. La propriété (6.5) démontre qu'elle est injective et donc aussi bijective. En posant $u = f(x)$ et $v = f(z)$ dans (6.5), on obtient

$$\|f^{-1}(u) - f^{-1}(v)\| \leq \frac{\|A^{-1}\|}{1 - \alpha} \|u - v\|, \quad (6.6)$$

ce qui démontre la continuité de f^{-1} sur $f(B)$.

b) Considérons maintenant un $y \in F$ avec $\|y - f(a)\| \leq \sigma$. Montrons que $g(B) \subset B$. Ceci découle de $g(x) - a = g(x) - g(a) + g(a) - a = g(x) - g(a) - A^{-1}(f(a) - y)$, car

$$\|g(x) - a\| \leq \alpha \|x - a\| + \|A^{-1}\| \cdot \|f(a) - y\| \leq \alpha \rho + \|A^{-1}\| \sigma = \rho$$

pour $x \in B$. Comme B est fermé, on peut appliquer le théorème du point fixe de Banach. Pour un tel y il existe donc un $x \in B$ satisfaisant $g(x) = x$, c.-à-d., $f(x) = y$. \square

Si $f(x)$ est différentiable en a (le centre de la boule B), il est naturel de choisir $A = f'(a)$ pourvu que cette application soit un isomorphisme. Mais la différentiabilité au point a ne garantit pas que l'estimation (6.4) soit vérifiée avec un $\alpha < 1$. Nous considérons alors une condition plus forte que la différentiabilité.

Définition 6.7 Soient E, F des espaces vectoriels normés, $U \subset E$ un ouvert et $a \in U$. On dit que $f : U \rightarrow F$ est *strictement différentiable* en a s'il existe une application linéaire continue $A : E \rightarrow F$ telle que

$$f(x) - f(z) = A(x - z) + r(x, z)\|x - z\|, \quad (6.7)$$

où la fonction $r : U \times U \rightarrow F$ satisfait $r(x, z) \rightarrow 0$ si $(x, z) \rightarrow (a, a)$.

En posant $z = a$ dans (6.7) on retrouve (4.1) avec $A = f'(a)$. Ainsi "strictement différentiable en a " entraîne "différentiable en a ", ce qui est heureux pour la terminologie choisie. On voit aussi de (6.7) qu'une fonction strictement différentiable en a est continue dans tout un voisinage de a .

Exemple 6.8 Il existe des fonctions qui sont différentiables en a , mais pas strictement différentiables en a . On peut prendre une fonction qui est différentiable en a , mais qui n'est continue dans aucun voisinage de a (exemple 5 de [HW95, III.6]). Un autre exemple est la fonction $f(x) = x^2 \cos(1/x)$ en $a = 0$. Les suites $x_n = (2n\pi)^{-1}$ et $z_n = ((2n + 1)\pi)^{-1}$ convergent vers 0, mais $(f(x_n) - f(z_n))/(x_n - z_n) = (x_n^2 + z_n^2)/(x_n - z_n)$ ne converge pas vers $f'(0) = 0$ si $n \rightarrow \infty$.

Proposition 6.9 Soient E, F des espaces vectoriels normés, $U \subset E$ un ouvert et $a \in U$. Si $f : U \rightarrow F$ est différentiable dans U et si l'application $f' : U \rightarrow \mathcal{L}(E, F)$ est continue au point a , alors f est strictement différentiable au point a .

Démonstration. Appliquons le théorème des accroissements finis à $g(x) := f(x) - f'(a)x$. Comme la dérivée de $g(x)$ est $g'(x) = f'(x) - f'(a)$, on a

$$\|f(x) - f(z) - f'(a)(x - z)\| \leq \sup_{0 < t < 1} \|f'(z + t(x - z)) - f'(a)\| \cdot \|x - z\|.$$

La continuité de $f'(x)$ au point a implique que l'expression $\sup_{0 < t < 1} \|f'(z + t(x - z)) - f'(a)\|$ tend vers 0 si $(x, z) \rightarrow (a, a)$. \square

Exemple 6.10 Reprenons la fonction $f(X) = X^{-1}$ de l'Exemple 4.6, pour laquelle la dérivée est $f'(X)H = -X^{-1}HX^{-1}$. Montrons que cette fonction est strictement différentiable dans $\text{GL}(E)$. Comme $f(X)$ est continue en $A \in \text{GL}(E)$, on a que pour tout $\varepsilon > 0$ il existe un $\delta > 0$ tel que $\|X^{-1} - A^{-1}\| < \varepsilon$ si $\|X - A\| < \delta$. En écrivant

$$f'(X)H - f'(A)H = -X^{-1}HX^{-1} + A^{-1}HA^{-1} = -X^{-1}H(X^{-1} - A^{-1}) - (X^{-1} - A^{-1})HA^{-1},$$

on obtient $\|(f'(X) - f'(A))H\| \leq \varepsilon(\|X^{-1}\| + \|A^{-1}\|)\|H\|$. On en déduit $\|f'(X) - f'(A)\| \leq \varepsilon(\|X^{-1}\| + \|A^{-1}\|)$ et donc la continuité de $f'(X)$ au point A . Ceci nous permet d'appliquer la Proposition 6.9.

Pour toutes les fonctions des exemples du paragraphe I.4, on peut démontrer sans difficulté que $f'(x)$, considérée comme fonction de x , est continue. Alors, la Proposition 6.9 implique qu'elles sont strictement différentiables.

I.7 Théorème d'inversion locale

Nous poursuivons l'étude de la résolution d'équations nonlinéaires $f(x) = 0$ (ou $f(x) = y$ pour un y donné) dans un espace de Banach.

Exemple 7.1 Considérons le système nonlinéaire

$$x_1^2 + x_2^3 - 3x_1 = y_1, \quad x_1^4 - x_1^2 x_2 + 2 = y_2. \quad (7.1)$$

Pour $(x_1, x_2) = (1, 2)$ on obtient $(y_1, y_2) = (6, 1)$. La question à laquelle on aimerait trouver une réponse est la suivante: *pour (y_1, y_2) proche de $(6, 1)$, existe-t-il une solution (x_1, x_2) de (7.1) qui est proche de $(1, 2)$? Est-elle localement unique?*

Nous disons qu'une application $f : U \rightarrow F$ (où E, F sont des espaces vectoriels normés et $U \subset E$ est un ouvert) est un *homéomorphisme local près de $a \in U$* , s'il existe un voisinage ouvert $U' \subset U$ de a et un voisinage ouvert V' de $f(a)$ tels que la restriction $f|_{U'}$ est un homéomorphisme de U' sur V' (voir la Définition 6.5).

Lemme 7.2 *Soient E, F des espaces de Banach, $U \subset E$ un ouvert et $a \in U$. Si $f : U \rightarrow F$ est strictement différentiable en a et si $f'(a)$ est un isomorphisme de E sur F , alors f est un homéomorphisme local près de a . De plus, l'application inverse f^{-1} est strictement différentiable en $b = f(a)$ et on a*

$$(f^{-1})'(b) = f'(a)^{-1}. \quad (7.2)$$

Démonstration. L'idée est d'appliquer la Proposition 6.6 avec $A = f'(a)$. La fonction f est strictement différentiable en a . Ceci signifie que pour tout $\varepsilon > 0$ il existe un $\delta > 0$ tel que

$$\|f(x) - f(z) - f'(a)(x - z)\| \leq \varepsilon \|x - z\| \quad \text{pour } x, z \in B, \quad (7.3)$$

où $B = \{x \in E ; \|x - a\| \leq \delta\}$. Alors, on a aussi

$$\|x - z - f'(a)^{-1}(f(x) - f(z))\| \leq \varepsilon \|f'(a)^{-1}\| \cdot \|x - z\| \quad \text{pour } x, z \in B. \quad (7.4)$$

On fixe $\varepsilon > 0$ tel que $\varepsilon \|f'(a)^{-1}\| \leq 1/2$. Pour le δ correspondant, f est un homéomorphisme de B sur $f(B)$ (Proposition 6.6) et $f(B)$ contient l'ouvert $V' = \{y \in F ; \|y - b\| < \sigma\}$ où $\sigma := \delta / (2\|f'(a)^{-1}\|)$. Donc, f est aussi un homéomorphisme de l'ouvert $U' := f^{-1}(V')$ sur V' .

Pour montrer que f^{-1} est strictement différentiable en $b = f(a)$, nous posons $x = f^{-1}(u)$ et $z = f^{-1}(v)$ dans (7.4). Ceci donne pour $u, v \in V'$ que

$$\|f^{-1}(u) - f^{-1}(v) - f'(a)^{-1}(u - v)\| \leq \varepsilon \|f'(a)^{-1}\| \cdot \|f^{-1}(u) - f^{-1}(v)\| \leq 2\varepsilon \|f'(a)^{-1}\|^2 \|u - v\|$$

(ici on a utilisé l'estimation (6.6) avec $A = f'(a)$ et $\alpha = 1/2$). Donc f^{-1} est strictement différentiable en b et la dérivée est donnée par (7.2). \square

Exemple 7.3 Pour le problème de l'Exemple 7.1 nous avons

$$f'(a) = \begin{pmatrix} 2a_1 - 3 & 3a_2^2 \\ 4a_1^3 - 2a_1a_2 & -a_1^2 \end{pmatrix} = \begin{pmatrix} -1 & 12 \\ 0 & -1 \end{pmatrix}.$$

Pour la norme euclidienne, on calcule $\|f'(a)^{-1}\| \approx 12.1$. Si l'on pose $\varepsilon = 0.04$, on peut trouver un $\delta > 0$ tel que (7.3) est vérifié. Pour tout $y \in \mathbb{R}^2$ avec $\|y - b\| \leq 2\delta/12.1$ le système (7.1) possède alors une solution x qui satisfait $\|x - a\| \leq \delta$. Dans cette boule de rayon δ , il n'y a pas d'autre solution. De plus, la démonstration du Lemme 7.2 montre que la méthode de Newton simplifiée converge vers cette solution.

Le but suivant est d'étudier si la solution de $f(x) = y$ dépend de manière différentiable de y . Plus précisément, nous étudions les conditions sous lesquelles la fonction inverse est différentiable dans tout un voisinage de $b = f(a)$.

Définition 7.4 Soient E, F des espaces vectoriels normés, $U \subset E$ et $V \subset F$ des ouverts.

- $f : U \rightarrow F$ est *continûment différentiable* (ou de classe \mathcal{C}^1) sur U , si f est différentiable en tout point de U et si $f' : U \rightarrow \mathcal{L}(E, F)$ est continue.
- $f : U \rightarrow V$ est un *difféomorphisme* (de classe \mathcal{C}^1) de U sur V , si f est bijective, continûment différentiable, et si l'inverse $f^{-1} : V \rightarrow U$ est continûment différentiable.
- $f : U \rightarrow F$ est un *difféomorphisme local près de $a \in U$* , s'il existe un voisinage ouvert $U' \subset U$ de a et un voisinage ouvert V' de $f(a)$ tels que f est un difféomorphisme de U' sur V' .

Remarquons qu'un homéomorphisme (voir la Définition 6.5) qui est continûment différentiable n'est pas toujours un difféomorphisme. La fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ définie par $f(x) = x^3$ nous sert comme contre-exemple, car l'inverse $f^{-1}(y) = \sqrt[3]{y}$ est continu mais pas différentiable à l'origine.

Une fonction qui est continûment différentiable sur U est strictement différentiable sur U . Ceci est une conséquence de la Proposition 6.9.

Théorème 7.5 (Théorème d'inversion locale) Soient E, F des espaces de Banach, $U \subset E$ un ouvert et $a \in U$. Une application $f : U \rightarrow F$ de classe \mathcal{C}^1 est un difféomorphisme local près de a , si et seulement si $f'(a)$ est un isomorphisme de E sur F . De plus, on a

$$(f^{-1})'(y) = f'(x)^{-1} \quad \text{pour } y = f(x) \text{ dans un voisinage de } b = f(a).$$

Démonstration. “ \Rightarrow ” : Si $f : U \rightarrow F$ est un difféomorphisme local près de a , on peut dériver l'identité $f^{-1}(f(x)) = x$. Ceci donne

$$(f^{-1})'(y) f'(x) = I \quad \text{avec } y = f(x) \tag{7.5}$$

dans un voisinage de a . Par conséquent, $f'(a)$ est inversible. L'inverse $f'(a)^{-1} = (f^{-1})'(b)$ est une application bornée car f^{-1} est supposée différentiable en b .

“ \Leftarrow ” : Le Lemme 7.2 implique que f est un homéomorphisme local près de a . Il reste à démontrer que f^{-1} est continûment différentiable dans un voisinage de $b = f(a)$. Comme $f'(x)$ est proche de $f'(a)$ (continuité de $f' : U \rightarrow \mathcal{L}(E, F)$) et $\text{GL}(E, F)$ est ouvert (voir le Lemme 4.5), $f'(x)$ est un isomorphisme pour tout x dans un voisinage U' de a . On peut donc appliquer le Lemme 7.2 à chaque point $x \in U'$, ce qui implique la différentiabilité de f^{-1} dans $V' := f(U')$. En dérivant l'identité $f^{-1}(f(x)) = x$, on obtient (7.5) et donc aussi $(f^{-1})'(y) = f'(f^{-1}(y))^{-1}$ pour $y \in V'$. La fonction $(f^{-1})' : V' \rightarrow \mathcal{L}(F, E)$, étant la composition des applications continues f^{-1} , f' et $(\cdot)^{-1}$, est par conséquent continue. \square

Un résultat global est le suivant.

Corollaire 7.6 Soient E, F des espaces de Banach, $U \subset E$ un ouvert et $f : U \rightarrow F$ continûment différentiable sur U . Alors, f est un difféomorphisme de U sur $f(U)$, si et seulement si (i) f est injective sur U et (ii) $f'(x)$ est un isomorphisme pour tout $x \in U$.

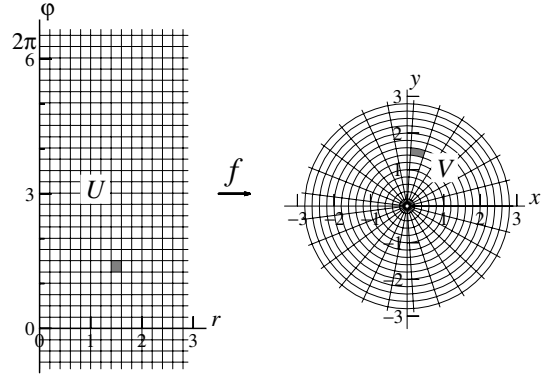
Démonstration. Ce résultat est une conséquence du Théorème 7.5, car la différentiabilité est une propriété locale. □

Exemple 7.7 (Coordonnées polaires) Soient $U = \{(r, \varphi) \in \mathbb{R}^2; r > 0\}$, $V = \mathbb{R}^2 \setminus \{(0, 0)\}$ et $f : U \rightarrow V$ donnée par

$$(r, \varphi) \mapsto (r \cos \varphi, r \sin \varphi).$$

Cette application est un difféomorphisme local près de chaque point de U , car

$$f'(r, \varphi) = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix}$$



et $\det(f'(r, \varphi)) = r > 0$. Elle n'est pas injective (on a $f(r, \varphi + 2\pi) = f(r, \varphi)$). Donc, elle n'est pas un difféomorphisme de U sur V . Si l'on restreint les ensembles U et V à

$$U_0 = \{(r, \varphi); r > 0, -\pi < \varphi < \pi\}, \quad V_0 = \mathbb{R}^2 \setminus \{(x, 0); x \leq 0\},$$

f devient un difféomorphisme de U_0 sur V_0 .

Exemple 7.8 (Coordonnées sphériques) L'application

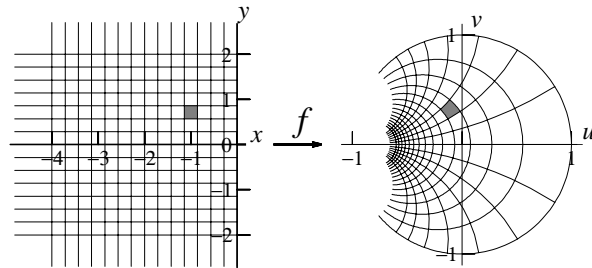
$$(r, \varphi, \theta) \mapsto (r \cos \varphi \sin \theta, r \sin \varphi \sin \theta, r \cos \theta)$$

est un difféomorphisme de $U = \{(r, \varphi, \theta); r > 0, -\pi < \varphi < \pi, 0 < \theta < \pi\}$ sur l'ensemble $V = \mathbb{R}^3 \setminus \{(x, 0, z); x \leq 0, z \in \mathbb{R}\}$, car le déterminant de la matrice Jacobienne est $-r^2 \sin \theta \neq 0$.

Exemple 7.9 (Transformation de Cayley) L'application

$$(x, y) \mapsto \left(\frac{1 - x^2 - y^2}{(1 - x)^2 + y^2}, \frac{2y}{(1 - x)^2 + y^2} \right) \tag{7.6}$$

est un difféomorphisme du demi-plan gauche $\{(x, y); x < 0\}$ sur le disque ouvert $\{(u, v); u^2 + v^2 < 1\}$. Pour démontrer ceci, nous identifions \mathbb{R}^2 avec le plan complexe (en posant $z = x + iy$ et $w = u + iv$) et nous remarquons que l'application (7.6) est équivalente à



$$w = \frac{1 + z}{1 - z} \quad \text{avec inverse} \quad z = \frac{w - 1}{w + 1}.$$

Donc, l'application (7.6) est bijective et elle est continûment différentiable (comme fonction rationnelle), ainsi que son inverse.

I.8 Théorème des fonctions implicites

Dans le paragraphe précédent nous avons considéré le problème de résoudre $f(y) = x$ et nous avons trouvé des conditions suffisantes permettant d'écrire la solution sous la forme $y = g(x)$ (attention: nous avons inversé les rôles de x et y). Le but de ce paragraphe est d'étendre ce résultat au problème

$$f(x, y) = 0 \quad (8.1)$$

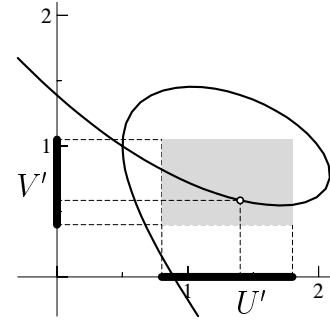
où x, y et $f(x, y)$ sont dans des espaces de Banach. On cherche à savoir si l'équation (8.1) peut être résolue pour obtenir $y = g(x)$ tel que (du moins localement)

$$f(x, y) = 0 \quad \iff \quad y = g(x). \quad (8.2)$$

Exemple 8.1 Le cas où x, y et $f(x, y)$ sont dans \mathbb{R} a été traité au cours d'Analyse I (voir [HW95, Théorème IV.3.8]). Considérons par exemple la fonction

$$f(x, y) = 16x^3 - 84x^2 + 162x - 89 + 27y^3 + 54xy^2 - 108y^2 + 36x^2y - 180xy + 162y.$$

Pour un point (a, b) satisfaisant $f(a, b) = 0$ et $\frac{\partial f}{\partial y}(a, b) \neq 0$, il existe des voisinages U' de a , V' de b et une fonction différentiable $g : U' \rightarrow V'$ tels que (8.2) est vraie pour $(x, y) \in U' \times V'$. Les points sur la courbe ayant une tangente verticale ou horizontale peuvent être trouvés par la condition $\frac{\partial f}{\partial y}(a, b) = 0$ et $\frac{\partial f}{\partial x}(a, b) = 0$ respectivement. Au point du croisement on a nécessairement $f(a, b) = 0$, $\frac{\partial f}{\partial y}(a, b) = 0$ et $\frac{\partial f}{\partial x}(a, b) = 0$ (3 conditions pour 2 inconnues).



Exemple 8.2 Une fonction $f(x_1, x_2, y) = 0$ représente une surface dans \mathbb{R}^3 . Quand peut-on écrire cette équation sous la forme $y = g(x_1, x_2)$?

Le système de deux fonctions

$$f_1(x, y_1, y_2) = 0, \quad f_2(x, y_1, y_2) = 0$$

représente l'intersection de deux surfaces dans \mathbb{R}^3 . Sous quelle condition peut-on écrire cette intersection sous la forme $y_1 = g_1(x), y_2 = g_2(x)$, ce qui représenterait une courbe dans \mathbb{R}^3 ?

Pour une fonction $f : U \rightarrow G$, $U \subset E \times F$ (où E, F, G sont des espaces de Banach) on définit la *dérivée partielle* $\frac{\partial f}{\partial y}(a, b)$ comme la dérivée de l'application $h(y) = f(a, y)$, où a est considéré comme un paramètre fixé, c.-à-d., $\frac{\partial f}{\partial y}(a, b) := h'(b)$. Si f est de classe \mathcal{C}^1 , h est aussi de classe \mathcal{C}^1 , car $h = f \circ \lambda$ où l'injection $\lambda : F \rightarrow E \times F$, définie par $\lambda(y) = (a, y)$, est continûment différentiable.

Dans le cas où $x \in \mathbb{R}^m$, $y \in \mathbb{R}^n$ et $f : \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ est donnée par

$$f(x, y) = \begin{pmatrix} f_1(x_1, \dots, x_m, y_1, \dots, y_n) \\ \vdots \\ f_n(x_1, \dots, x_m, y_1, \dots, y_n) \end{pmatrix} \quad \text{on a} \quad \frac{\partial f}{\partial y}(x, y) = \begin{pmatrix} \frac{\partial f_1}{\partial y_1}(x, y) & \dots & \frac{\partial f_1}{\partial y_n}(x, y) \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial y_1}(x, y) & \dots & \frac{\partial f_n}{\partial y_n}(x, y) \end{pmatrix}.$$

Théorème 8.3 (Théorème des fonctions implicites) Soient E, F et G des espaces de Banach, $U \subset E$ et $V \subset F$ des ouverts et $f : U \times V \rightarrow G$ une application de classe \mathcal{C}^1 . Supposons qu'en $(a, b) \in U \times V$

$$f(a, b) = 0 \quad \text{et} \quad \frac{\partial f}{\partial y}(a, b) \quad \text{est un isomorphisme de } F \text{ sur } G.$$

- (i) Il existe alors un voisinage U' de a , un voisinage V' de b et une application unique $g : U' \rightarrow V'$ tels que $f(x, g(x)) = 0$ pour $x \in U'$.
(ii) L'application $g : U' \rightarrow V'$ est de classe \mathcal{C}^1 et on a

$$g'(x) = -\left(\frac{\partial f}{\partial y}(x, g(x))\right)^{-1} \frac{\partial f}{\partial x}(x, g(x)). \quad (8.3)$$

Démonstration. L'idée est de considérer l'application

$$F : U \times V \rightarrow E \times G \quad \text{définie par} \quad F(x, y) = (x, f(x, y))$$

et d'appliquer le théorème d'inversion locale. Cette application est de classe \mathcal{C}^1 et a pour dérivée

$$F'(a, b)(h, k) = \left(h, \frac{\partial f}{\partial x}(a, b)h + \frac{\partial f}{\partial y}(a, b)k\right).$$

De plus, elle satisfait $F(a, b) = (a, 0)$. Puisque $\frac{\partial f}{\partial y}(a, b)$ est un isomorphisme, $F'(a, b)$ est inversible avec pour inverse

$$F'(a, b)^{-1}(\hat{h}, \hat{k}) = \left(\hat{h}, \frac{\partial f}{\partial y}(a, b)^{-1}\left(\hat{k} - \frac{\partial f}{\partial x}(a, b)\hat{h}\right)\right).$$

Cet inverse est continu car $\frac{\partial f}{\partial x}(a, b)$ et $\frac{\partial f}{\partial y}(a, b)^{-1}$ le sont. D'après le théorème d'inversion locale (Théorème 7.5), F est un difféomorphisme de classe \mathcal{C}^1 d'un voisinage de (a, b) sur un voisinage de $(a, 0)$. On peut supposer qu'ils contiennent $U' \times V'$ et $U' \times W'$, respectivement, où U', V' et W' sont des voisinages de a, b et $0 \in G$. Quitte à réduire U' , on peut aussi supposer que $F^{-1}(U' \times \{0\}) \subset U' \times V'$. Le difféomorphisme inverse F^{-1} est de la forme $F^{-1}(x, z) = (x, \hat{g}(x, z))$ et on a donc $f(x, \hat{g}(x, z)) = z$. L'application $g(x) := \hat{g}(x, 0)$ est l'application cherchée.

Comme $F^{-1}(x, z)$ est de classe \mathcal{C}^1 , les applications $\hat{g}(x, z)$ et $g(x) = \hat{g}(x, 0)$ sont aussi de classe \mathcal{C}^1 . On obtient finalement la formule (8.3) en dérivant l'identité $f(x, g(x)) = 0$ par rapport à x . \square

Exemple 8.4 Soit $p_a(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n$ un polynôme à coefficients réels $a = (a_0, a_1, \dots, a_n)$. Est-ce que la racine x_0 de $p_a(x) = 0$ est une fonction différentiable de a_0, \dots, a_n ? Pour répondre à cette question, nous considérons la fonction

$$f(x, a_0, \dots, a_n) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n = p_a(x)$$

à $n + 2$ variables pour laquelle

$$\frac{\partial f}{\partial x}(x, a_0, \dots, a_n) = a_1 + 2a_2x + \dots + na_nx^{n-1} = p'_a(x).$$

Si pour des coefficients $a^* = (a_0^*, \dots, a_n^*)$ on a $p_{a^*}(x_0^*) = 0$ et $p'_{a^*}(x_0^*) \neq 0$, l'équation $p_a(x) = 0$ possède, pour a proche de a^* , une solution $x_0(a)$ qui est proche de x_0^* et qui est une fonction différentiable de a_0, \dots, a_n .

En particulier, le polynôme $p_\varepsilon(x) = x^6 - x^5 + x^3 - 1 + \varepsilon$ possède un zéro proche de $x_0^* = 1$ qui dépend différentiablement de ε (on a $x_0(\varepsilon) = 1 - \varepsilon/4 + \mathcal{O}(\varepsilon^2)$). Par contre, les zéros du polynôme $p_\varepsilon(x) = x^2 - \varepsilon$ satisfont $x_0(\varepsilon) = \pm\sqrt{\varepsilon}$ pour $\varepsilon > 0$ (une fonction non-différentiable à l'origine) et pour $\varepsilon < 0$ le polynôme n'a même pas de zéro réel.

I.9 Applications bilinéaires et multilinéaires

Les dérivées d'ordre supérieur sont des applications bilinéaires (pour la 2ème dérivée), trilineaires (pour la 3ème dérivée) et multilinéaires (pour la n ème dérivée). C'est la raison pour laquelle nous allons étudier ces applications un peu plus en détail.

Rappelons qu'une application

$$A : E_1 \times \dots \times E_n \rightarrow F$$

(où E_1, \dots, E_n, F sont des espaces vectoriels normés) est dite *multilinéaire* (bilinéaire si $n = 2$) si, pour chaque $i \in \{1, \dots, n\}$ et pour chaque $a_i \in E_i$ l'application partielle $x_i \mapsto f(a_1, \dots, a_{i-1}, x_i, a_{i+1}, \dots, a_n)$ est linéaire. L'espace $E_1 \times \dots \times E_n$, muni de la norme

$$\|(x_1, \dots, x_n)\| := \max \{\|x_1\|_1, \dots, \|x_n\|_n\}, \quad (9.1)$$

est un espace vectoriel normé. C'est un espace de Banach si les espaces E_1, \dots, E_n sont des espaces de Banach. Cela a donc un sens d'étudier la continuité d'une application multilinéaire (voir l'exemple 4.7).

Proposition 9.1 *Pour une application multilinéaire $A : E_1 \times \dots \times E_n \rightarrow F$ (E_1, \dots, E_n et F sont des espaces vectoriels normés), les conditions suivantes sont équivalentes:*

- (a) A est continue en tout point de E_1, \dots, E_n ;
- (b) A est continue à l'origine $(0, \dots, 0) \in E_1 \times \dots \times E_n$;
- (c) $\|A(x_1, \dots, x_n)\|$ est bornée sur la boule-unité de $E_1 \times \dots \times E_n$.

Démonstration. La démonstration procède comme celle de la Proposition 3.1. Pour démontrer (c) \Rightarrow (a), nous écrivons

$$\begin{aligned} A(x_1, \dots, x_n) - A(a_1, \dots, a_n) &= A(x_1 - a_1, x_2, \dots, x_n) \\ &\quad + A(a_1, x_2 - a_2, x_3, \dots, x_n) + \dots + A(a_1, \dots, a_{n-1}, x_n - a_n) \end{aligned}$$

et estimons chaque terme séparément. □

Définition 9.2 Soient E_1, \dots, E_n et F des espaces vectoriels normés. On dénote par $\mathcal{L}(E_1, \dots, E_n; F)$ l'ensemble des applications multilinéaires continues de $E_1 \times \dots \times E_n$ dans F . Pour un élément $A \in \mathcal{L}(E_1, \dots, E_n; F)$, on définit

$$\|A\| = \sup_{\|(x_1, \dots, x_n)\| \leq 1} \|A(x_1, \dots, x_n)\| = \sup_{x_1 \neq 0, \dots, x_n \neq 0} \frac{\|A(x_1, \dots, x_n)\|}{\|x_1\|_1 \cdot \dots \cdot \|x_n\|_n}. \quad (9.2)$$

Si $E_1 = \dots = E_n = E$, on écrit aussi $\mathcal{L}^n(E; F)$ au lieu de $\mathcal{L}(E, \dots, E; F)$.

L'espace $\mathcal{L}(E_1, \dots, E_n; F)$ muni de (9.2) est un espace vectoriel normé. Si F est complet, alors $\mathcal{L}(E_1, \dots, E_n; F)$ est aussi complet. Cette affirmation est démontrée exactement comme pour la Proposition 3.4. Le résultat de la proposition suivante est la base pour l'interprétation de la deuxième dérivée d'une fonction comme application bilinéaire.

Proposition 9.3 (Isométrie naturelle) *Soient E, F, G des espaces vectoriels normés. Alors, l'application*

$$\psi : \mathcal{L}(E, \mathcal{L}(F, G)) \rightarrow \mathcal{L}(E, F; G),$$

définie par $\psi(A) = B$ où $B(x, y) = (Ax)y$, est un isomorphisme qui est également une isométrie, c.-à-d., on a $\|\psi(A)\| = \|A\|$ pour tout $A \in \mathcal{L}(E, \mathcal{L}(F, G))$.

Démonstration. L'application ψ est linéaire et bijective. Son inverse est donné par $\psi^{-1}(B) = A$ où, pour $x \in E$, $Ax \in \mathcal{L}(F, G)$ est l'application $y \mapsto B(x, y)$. Pour démontrer la continuité de ψ et de ψ^{-1} , il suffit de voir que $\|\psi(A)\| = \|A\|$.

En utilisant l'estimation (3.2) une fois pour Ax et une deuxième fois pour A , nous obtenons

$$\|B(x, y)\| = \|(Ax)y\| \leq \|Ax\| \cdot \|y\| \leq \|A\| \cdot \|x\| \cdot \|y\|.$$

Ceci implique que $B = \psi(A)$ satisfait $\|B\| \leq \|A\|$.

D'autre part, la Définition 9.2 montre que

$$\|(Ax)y\| = \|B(x, y)\| \leq \|B\| \cdot \|x\| \cdot \|y\|.$$

La norme de l'application linéaire Ax satisfait donc $\|Ax\| \leq \|B\| \cdot \|x\|$ et, par conséquent, $\|A\| \leq \|B\|$. Ces deux inégalités montrent que l'application ψ est une isométrie. \square

On peut donc identifier les applications linéaires continues de E dans $\mathcal{L}(F, G)$ avec les applications bilinéaires de $E \times F$ dans G .

Exemple 9.4 Une matrice C (avec n colonnes et m lignes) peut être identifiée avec l'application bilinéaire $B : \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}$ définie par $B(x, y) = x^T C y$. Elle peut également être identifiée avec un élément de $\mathcal{L}(\mathbb{R}^m, \mathcal{L}(\mathbb{R}^n, \mathbb{R}))$ de la manière suivante: pour $x \in \mathbb{R}^m$, le vecteur $x^T C \in \mathcal{L}(\mathbb{R}^n, \mathbb{R})$ définit l'application $x^T C : y \mapsto x^T C y$.

L'affirmation de la proposition précédente peut être généralisée à des applications multilinéaires. Si E_1, \dots, E_n et F sont des espaces vectoriels normés, l'application

$$\psi : \mathcal{L}(E_1, \mathcal{L}(E_2, \dots, E_n; F)) \rightarrow \mathcal{L}(E_1, \dots, E_n; F), \quad (9.3)$$

définie par $\psi(A) = B$ où $B(x_1, \dots, x_n) = (Ax_1)(x_2, \dots, x_n)$, est un isomorphisme et une isométrie. La démonstration de ce fait est identique à celle de la Proposition 9.3.

I.10 Dérivées d'ordre supérieur

La dérivée d'une fonction $f : U \rightarrow F$ (avec $U \subset E$) est une application $f' : U \rightarrow \mathcal{L}(E, F)$. Comme $\mathcal{L}(E, F)$ est un espace vectoriel normé, rien ne nous empêche de considérer la différentiabilité de f' .

Définition 10.1 Soient E, F deux espaces vectoriels normés, $U \subset E$ un ouvert et $f : U \rightarrow F$ différentiable dans un voisinage de $a \in U$. On dit que f est *deux fois différentiable* en a si

$f' : U \rightarrow \mathcal{L}(E, F)$ est différentiable en a . La dérivée de f' satisfait $(f')'(a) \in \mathcal{L}(E, \mathcal{L}(E, F))$. En utilisant l'identification de la Proposition 9.3, on définit la *deuxième dérivée* de f en a comme l'application bilinéaire

$$f''(a)(h, k) := \left((f')'(a)h \right) k.$$

Dans la situation de la définition précédente, considérons l'application $g_k : U \rightarrow F$ définie par $g_k(x) := f'(x)k$ (avec un $k \in E$ fixé). Alors, par la formule (4.10) de Leibniz (on a $g_k(x) = B(f'(x), k)$ où $B : \mathcal{L}(E, F) \times E \rightarrow F$ est l'application bilinéaire $B(A, v) = Av$) on obtient

$$g'_k(a)h = \left((f')'(a)h \right) k = f''(a)(h, k).$$

Exemple 10.2 Considérons une fonction $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ et notons $x = (x_1, \dots, x_n)^T$. Pour $k \in \mathbb{R}^n$ fixé, la dérivée de la fonction $g_k(x) := f'(x)k = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(x) k_j$ est donnée par $g'_k(x)h = \sum_{i=1}^n \left(\sum_{j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(x) k_j \right) h_i$. La deuxième dérivée de f est donc

$$f''(x)(h, k) = \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(x) h_i k_j.$$

Pour la fonction $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ de l'exemple 4.2, la deuxième dérivée est

$$f''(a)(h, k) = \begin{pmatrix} 2h_1k_1 + 2h_2k_2 + 2h_3k_3 \\ a_1(h_2k_3 + h_3k_2) + a_2(h_1k_3 + h_3k_1) + a_3(h_1k_2 + h_2k_1) \end{pmatrix}.$$

Exemple 10.3 Pour la fonction $f(X) = X^{-1}$ de $\text{GL}(E)$ dans $\mathcal{L}(E)$ l'Exemple 4.6 montre que $g_K(X) := f'(X)K = -X^{-1}KX^{-1}$ pour $K \in \mathcal{L}(E)$. La formule de Leibniz nous donne alors

$$f''(A)(H, K) = A^{-1}HA^{-1}KA^{-1} + A^{-1}KA^{-1}HA^{-1}.$$

Théorème 10.4 Soient E, F deux espaces vectoriels normés, $U \subset E$ un ouvert, et supposons que $f : U \rightarrow F$ soit deux fois différentiable en $a \in U$. Alors, l'application bilinéaire $f''(a) : E \times E \rightarrow F$ est symétrique, c.-à-d.,

$$f''(a)(h, k) = f''(a)(k, h) \quad \text{pour } h, k \in E.$$

Démonstration. Nous allons montrer que

$$f''(a)(h, k) = \lim_{\varepsilon \rightarrow 0} \frac{1}{\varepsilon^2} \left(f(a + \varepsilon h + \varepsilon k) - f(a + \varepsilon h) - f(a + \varepsilon k) + f(a) \right). \quad (10.1)$$

Puisque l'expression de droite est symétrique en h et k , il est de même de $f''(a)(h, k)$.

Pour démontrer (10.1), nous considérons l'application

$$g_u(v) := f(a + u + v) - f(a + u) - f(a + v) + f(a) - f''(a)(u, v)$$

où u et v sont suffisamment petits en norme. Comme $g_u(0) = 0$, le théorème des accroissements finis (Théorème 5.2) implique que

$$\|g_u(v)\| \leq \sup_{0 < t < 1} \|g'_u(tv)\| \cdot \|v\|. \quad (10.2)$$

Il reste donc à estimer la dérivée

$$g'_u(v) = f'(a + u + v) - f'(a + v) - f''(a)(u, \cdot)$$

(la notation $f''(a)(u, \cdot)$ est utilisée pour l'application $h \mapsto f''(a)(u, h)$). Le fait que f est deux fois différentiable en a (c.-à-d., $x \mapsto f'(x)$ est différentiable en a) implique que

$$\begin{aligned} f'(a + u + v) &= f'(a) + f''(a)(u + v, \cdot) + r(u + v)\|u + v\| \\ f'(a + v) &= f'(a) + f''(a)(v, \cdot) + r(v)\|v\| \end{aligned}$$

où $r(v) \rightarrow 0$ si $v \rightarrow 0$. La soustraction de ces deux équations donne pour $g'_u(v)$ la formule $g'_u(v) = r(u + v)\|u + v\| - r(v)\|v\|$ et on obtient pour $0 < t < 1$

$$\|g'_u(tv)\| \leq (\|r(u + tv)\| + \|r(tv)\|)(\|u\| + \|v\|). \quad (10.3)$$

On déduit des estimations (10.2) et (10.3) que $\|g_u(v)\|/(\|u\| + \|v\|)^2 \rightarrow 0$ si $\|u\| + \|v\| \rightarrow 0$. En posant $u = \varepsilon h$ et $v = \varepsilon k$ on obtient donc l'affirmation (10.1). \square

La définition de la troisième dérivée est analogue à Définition 10.1. Supposons qu'une fonction $f : U \rightarrow F$ (E, F des espaces vectoriels normés et $U \subset E$ un ouvert) soit 2 fois différentiable dans un voisinage de $a \in U$. On dit que f est 3 fois différentiable en a si $f'' : U \rightarrow \mathcal{L}^2(E; F)$ est différentiable en a . La dérivée satisfait $(f'')'(a) \in \mathcal{L}(E, \mathcal{L}^2(E; F))$ et, en utilisant l'isométrie entre $\mathcal{L}(E, \mathcal{L}^2(E; F))$ et $\mathcal{L}^3(E; F)$ on définit la 3^{ème} dérivée de f en a comme l'application trilinéaire

$$f'''(a)(h, k, l) := ((f'')'(a)h)(k, l).$$

De manière évidente on définit par récurrence la p ème dérivée $f^{(p)}(a)$ comme application multilinéaire $f^{(p)}(a) \in \mathcal{L}^p(E; F)$.

Pour un calcul pratique de la 3^{ème} dérivée, on peut utiliser la formule

$$g'_{kl}(a)h = ((f'')'(a)h)(k, l) = f'''(a)(h, k, l) \quad (10.4)$$

où $g_{kl} : U \rightarrow F$ est définie par $g_{kl}(x) := f''(x)(k, l)$ et $k, l \in E$ sont des vecteurs fixés. La 3^{ème} dérivée de f peut également être interprétée comme la deuxième dérivée de $g_l(x) := f'(x)l$, c.-à-d.,

$$g''_l(a)(h, k) = ((f')''(a)(h, k))l = f'''(a)(h, k, l). \quad (10.5)$$

Les deux formules (10.4) et (10.5) montrent qu'on peut échanger k et l ainsi que h et k sans changer la valeur de $f'''(a)(h, k, l)$. On a donc le résultat suivant.

Corollaire 10.5 *Si, sous les hypothèses du Théorème 10.4, la fonction $f : U \rightarrow F$ est p fois différentiable en $a \in U$, l'application multilinéaire $f^{(p)}(a)$ est symétrique, c.-à-d.,*

$$f^{(p)}(a)(h_1, \dots, h_p) = f^{(p)}(a)(h_{\sigma(1)}, \dots, h_{\sigma(p)})$$

où σ est une permutation quelconque de $\{1, \dots, p\}$. \square

Exemple 10.6 La 3^{ème} dérivée d'une application $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ est donnée par

$$f'''(x)(h, k, l) = \sum_{i=1}^n \sum_{j=1}^n \sum_{m=1}^n \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_m}(x) h_i k_j l_m.$$

Définition 10.7 Soient E, F des espaces vectoriels normés, $U \subset E$ et $V \subset F$ des ouverts.

- $f : U \rightarrow F$ est p fois continûment différentiable (ou de classe \mathcal{C}^p) sur U , si f est p fois différentiable en tout point de U et si $f^{(p)} : U \rightarrow \mathcal{L}^p(E; F)$ est continue.
- $f : U \rightarrow F$ est indéfiniment différentiable (ou de classe \mathcal{C}^∞) sur U , si f est de classe \mathcal{C}^p pour tout p .
- $f : U \rightarrow V$ est un *difféomorphisme* (de classe \mathcal{C}^p) de U sur V , si f est bijective, p fois continûment différentiable, et si l'inverse $f^{-1} : V \rightarrow U$ est p fois continûment différentiable.

Exemple 10.8 La fonction $f(X) = X^{-1}$ de $\text{GL}(E)$ dans $\mathcal{L}(E)$ est indéfiniment différentiable. En effet, en dérivant la formule de l'Exemple 10.3 on obtient

$$\begin{aligned} f'''(A)(H, K, L) = & -A^{-1}HA^{-1}KA^{-1}LA^{-1} - A^{-1}HA^{-1}LA^{-1}KA^{-1} \\ & -A^{-1}LA^{-1}HA^{-1}KA^{-1} - A^{-1}KA^{-1}HA^{-1}LA^{-1} \\ & -A^{-1}KA^{-1}LA^{-1}HA^{-1} - A^{-1}LA^{-1}KA^{-1}HA^{-1}. \end{aligned}$$

On voit par récurrence que chaque dérivée est une combinaison linéaire d'expressions qui sont un produit alterné de A^{-1} avec des applications linéaires constantes.

Proposition 10.9 Soient E, F des espaces vectoriels normés, $U \subset E$, $V \subset F$ des ouverts et $f : U \rightarrow V$ un difféomorphisme de classe \mathcal{C}^1 . Si f est de classe \mathcal{C}^p , alors la fonction inverse $f^{-1} : V \rightarrow U$ est aussi de classe \mathcal{C}^p .

Démonstration. Par le Théorème 7.5 on sait que

$$(f^{-1})'(y) = f'(f^{-1}(y))^{-1},$$

c.-à-d., $(f^{-1})'$ est la composition de trois applications, notamment $y \mapsto f^{-1}(y)$, $x \mapsto f'(x)$ et $A \mapsto A^{-1}$. Toutes les trois applications sont continûment différentiables. Alors, f^{-1} est de classe \mathcal{C}^2 . La démonstration pour $p \geq 3$ est analogue. \square

Une conséquence de cette proposition est la suivante: si dans le théorème d'inversion locale (Théorème 7.5) la fonction f est de classe \mathcal{C}^p , alors son inverse est automatiquement de classe \mathcal{C}^p . Similairement, si dans le théorème des fonctions implicites (Théorème 8.3) la fonction f est de classe \mathcal{C}^p , alors l'application g est aussi de classe \mathcal{C}^p .

I.11 Exercices

1. Soit $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ donnée par

$$f(x) = \begin{pmatrix} x_1x_2^2 - 3x_1x_2x_3 \\ x_1^2 + x_2^2 + x_3^2 \end{pmatrix},$$

et $g : \mathbb{R}^n \rightarrow \mathbb{R}$ donnée par $g(x) = x^T Ax + b^T x$, où A est une matrice $n \times n$ et $b \in \mathbb{R}^n$. Calculer les deux premières dérivées de f et g .

2. Montrer que l'application $N : \mathbb{R}^2 \rightarrow \mathbb{R}$ donnée par $N(x) = \sqrt{x_1^2 - x_1x_2 + 4x_2^2}$, définit une norme sur \mathbb{R}^2 et montrer (explicitement) que cette norme est équivalente à la norme $\|\cdot\|_\infty$.
3. Démontrer que (2.1) est une norme.

4. Si la norme découle d'un produit scalaire (2.1), alors on a

$$\|x + y\| + \|x - y\| = 2(\|x\| + \|y\|) \quad (\text{identité du parallélogramme})$$

Montrer à l'aide de contre-exemples que \mathbb{R}^n munit de la norme $\|\cdot\|_1$ ou $\|\cdot\|_\infty$ n'est pas un espace de Hilbert.

5. On considère les ensembles suivants de \mathbb{R}^2 :

$$\begin{aligned} A &= \{(x, y) \in \mathbb{R}^2, 0 < x + y \leq 1\}, \\ B &= \{(x, y) \in \mathbb{R}^2, \max(|x|, |y|) < 3\}, \\ C &= \{(x, y) \in \mathbb{R}^2, (x^2 + y^2)^2 - 2x^2 + 2y^2 = 0\}, \\ D &= \{(x, y) \in \mathbb{R}^2, y = 0, x \in \{1/n, n = 1, 2, \dots\}\}. \end{aligned}$$

Quels ensembles sont-ils ouverts, fermés, bornés, compacts? Justifier.

6. Pour $\alpha \geq 0$ considérons la fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ donnée par

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

- (a) Montrer que $\partial f / \partial x$ et $\partial f / \partial y$ existent partout et sont telles que les applications $x \mapsto \frac{\partial f}{\partial x}(x, b)$, $y \mapsto \frac{\partial f}{\partial y}(a, y)$ sont continues pour tout $a, b \in \mathbb{R}$.
- (b) Si $\alpha \geq 1/2$, les applications $\partial f / \partial x$ et $\partial f / \partial y$ sont discontinues en $(0, 0)$.
- (c) $f(x, y)$ est continue à l'origine si et seulement si $\alpha < 1$.
- (d) $f(x, y)$ est dérivable à l'origine si et seulement si $\alpha < 1/2$.
- (e) Pour $\alpha = 1$, trouver un ouvert $U \subset \mathbb{R}$ et un fermé $F \subset \mathbb{R}$ tels que $f^{-1}(U)$ ne soit pas ouvert et $f^{-1}(F)$ ne soit pas fermé.
7. Soit E un espace de Hilbert et soit $A \in \mathcal{L}(E)$. Montrer que

$$\|A\| = \sup_{u \neq 0, v \neq 0} \frac{\langle u, Av \rangle}{\|u\| \cdot \|v\|}$$

où $\langle \cdot, \cdot \rangle$ est le produit scalaire et $\|\cdot\|$ est la norme induite par le produit scalaire.

8. Démontrer que $\|f\|_2$ est une norme sur $\mathcal{C}([0, 1])$.
9. Démontrer que, pour un ensemble arbitraire A ,

$$\mathcal{B}(A) = \{f : A \rightarrow \mathbb{R}; f \text{ est bornée}\} \quad \text{avec} \quad \|f\|_\infty = \sup_{t \in A} |f(t)|,$$

est un espace de Banach (Proposition 2.3).

10. Le rayon spectral $\rho(A)$ d'une matrice $n \times n$, A , est défini comme suit:

$$\rho(A) = \max \{|\lambda|; \lambda \text{ valeur propre de } A\}.$$

Soit $E = \mathbb{R}^n$ muni d'une norme et soit $A \in \mathcal{L}(E)$. Montrer que pour toute norme sur \mathbb{R}^n on a l'inégalité suivante $\|A\| \geq \rho(A)$, et que si A est symétrique on l'égalité $\|A\|_2 = \rho(A)$.

11. Soit

$$A = \begin{pmatrix} 0.999 & 1000 \\ 0 & 0.999 \end{pmatrix}.$$

- (a) Calculer le rayon spectral $\rho(A)$ ainsi que $\|A\|_1$, $\|A\|_2$ et $\|A\|_\infty$.
- (b) Trouver une norme $\|\cdot\|$ sur \mathbb{R}^2 telle que, pour cette norme, on ait $\|A\| \leq 1$.
12. Soit l'ensemble des matrices $n \times n$ que l'on identifie avec $\mathbb{R}^{n \times n}$ et que l'on muni de la norme $\|A\| = \sup_{\|x\| \leq 1} \|Ax\|$. Considérons l'application linéaire $l : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}$ donnée par $l(H) = X_0 H + H X_0$ où X_0 est une matrice fixée (voir Exemple 1.1). Calculer la norme de l'application l .

13. Dans la situation de l'Exemple 3.8 démontrer que

$$\|A\|_\infty = \max_{t \in [0,1]} \int_0^1 |k(t,s)| ds.$$

Indication. Considérer en premier que la fonction $k(t,s)$ possède qu'un nombre fini de zéros pour tout t .

14. On considère l'équation intégrale suivante, dite équation intégrale de Fredholm (2ème espèce):

$$y(t) = f(t) + \int_0^1 k(t,s)y(s)ds, \quad (11.1)$$

où $f \in \mathcal{C}([0,1])$ et $k \in \mathcal{C}([0,1] \times [0,1])$ sont données et y est l'inconnue. On suppose que

$$\max_{t \in [0,1]} \int_0^1 |k(t,s)| ds < 1$$

- (a) Démontrer que (11.1) possède une unique solution dans $\mathcal{C}([0,1])$.
 (b) En utilisant la série de Neumann, écrire la solution sous la forme

$$y(t) = f(t) + \int_0^1 r(t,s)f(s)ds$$

15. Résoudre

$$y(t) = f(t) + \frac{1}{2} \int_0^1 e^{t-s} y(s) ds. \quad (11.2)$$

- (a) On pose $a \equiv \int_0^1 e^{-s} y(s) ds$. En multipliant (11.2) par e^{-t} et en intégrant, on trouve a et donc aussi la solution $y(t) = f(t) + ae^t/2$.
 (b) En utilisant la formule de l'exercice précédent.
16. Soit E un espace de Banach, on considère l'application $f : \mathcal{L}(E) \rightarrow \mathcal{L}(E)$ donnée par $f(X) = X^3$. Calculer la dérivée de f . Montrer par des exemples, qu'en général, $f'(X) \neq 3X^2$.
17. Soit E un espace de Banach, on considère l'application $\exp : \mathcal{L}(E) \rightarrow \mathcal{L}(E)$ donnée par la série suivante:

$$\exp(X) = \sum_{n=0}^{\infty} \frac{1}{n!} X^n.$$

- (a) Montrer que l'application \exp est bien définie.
 (b) Calculer la dérivée, \exp' de l'application \exp .
 (c) Montrer par des exemples que, en général, $\exp'(X) \neq \exp(X)$.

Indication. Si cela vous aide, supposez que $E = \mathbb{R}^n$.

18. Calculer la dérivée de la fonction $f : \text{GL}(\mathbb{R}^n) \rightarrow \mathcal{L}(\mathbb{R}^n)$ donnée par $f(X) = (X^T X)^{-1}$.
19. Soit $A : \mathbb{R} \rightarrow \mathbb{R}^{n \times n}$ une fonction différentiable à valeur dans les matrices $n \times n$. Montrer que

$$\frac{d}{dx} (A(x))^{-1} = -A(x)^{-1} A'(x) A(x)^{-1}$$

Indication. Utiliser l'Exemple 4.6 ainsi que la formule pour la dérivée de fonctions composées.

20. (a) Soit $M : E_1 \times \dots \times E_n \rightarrow F$ une application multilinéaire où E_1, \dots, E_n, F sont des espaces vectoriels normés. Calculer la dérivée de M .
 (b) Soit $A : \mathbb{R} \rightarrow \mathbb{R}^{3 \times 3}$ une fonction matricielle différentiable dont les colonnes sont $A_1(x)$, $A_2(x)$ et $A_3(x)$. Montrer que

$$\frac{d}{dx} \det(A(x)) = \det(A'_1, A_2, A_3) + \det(A_1, A'_2, A_3) + \det(A_1, A_2, A'_3)$$

21. Soit $f(x,y) = \ln(x^2 + y^2)$ pour $x^2 + y^2 > 0$ et $g(x,y) = (xy, \sqrt{x}/y)^T$ pour $x > 0, y > 0$. Calculer la dérivée de $h(x,y) = (f \circ g)(x,y) = \ln(x^2 y^2 + x/y^2)$ directement et ensuite avec la formule de différentiation des fonctions composées.
22. Soit $f : [a,b] \rightarrow \mathbb{R}$ continue, $x_1, x_2, \dots, x_n \in (a,b)$ et f différentiable sur l'ensemble $D = (a,b) \setminus \{x_1, x_2, \dots, x_n\}$. Démontrer ou donner des contre-exemples aux affirmations suivantes:
 (a) Il existe un $\xi \in D$ tel que $f(b) - f(a) = f'(\xi)(b - a)$.

(b) On a l'estimation $|f(b) - f(a)| \leq \sup_{t \in D} |f'(t)|(b - a)$.

23. Considérons la fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ définie par

$$f(x) = \begin{cases} x + e^{-x/2} & \text{si } x \geq 0 \\ e^{x/2} & \text{si } x \leq 0 \end{cases} .$$

(a) Montrer que $|f(x) - f(y)| < |x - y|$ pour $x \neq y$.

(b) Montrer que la fonction $f(x)$ ne possède pas de point fixe.

Est-ce une contradiction au théorème du point fixe?

24. On considère $\mathcal{C}([0, 1])$ avec la norme $\|\cdot\|_\infty$ et l'application $f : \mathcal{C}([0, 1]) \rightarrow \mathcal{C}([0, 1])$ donnée par

$$f(x)(t) = \lambda \int_0^1 k(t, s)g(x(s))ds$$

où $k \in \mathcal{C}([0, 1] \times [0, 1])$ et $g : \mathbb{R} \rightarrow \mathbb{R}$ est suffisamment différentiable. Trouver une condition sur λ sous laquelle f est une contraction sur le fermé

$$D = \{x \in \mathcal{C}([0, 1]) ; \|x\|_\infty \leq 1\} .$$

25. Montrer que l'hypothèse “ D fermé” ne peut en général pas être omise dans le théorème du point fixe de Banach: trouver un D non fermé et une application $f : D \rightarrow E$ tels que $f(D) \subset D$ et f est une contraction sur D , mais f ne possède pas de point fixe dans D .

26. Soit U un ouvert connexe de \mathbb{R} et E un espace vectoriel normé. Soit $f : U \rightarrow E$ une application qui satisfait

$$\|f(x) - f(y)\| \leq C\|x - y\|^2 \quad \text{pour tout } x, y \in U ,$$

où $C > 0$. Montrer que f est constante sur U .

27. Soit la fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ donnée par

$$f(x_1, x_2) = \begin{pmatrix} x_1^2 + x_2^2 + 5 \\ 3x_1x_2 + 2 \\ x_1^3 + x_2^3 \end{pmatrix} ,$$

et soit $K = \{(x_1, x_2) \in \mathbb{R}^2 \mid |x_1| \leq 1, |x_2| \leq 2\}$. Trouver une constante L telle que

$$\|f(x) - f(y)\|_\infty \leq L\|x - y\|_\infty \quad \text{pour } x, y \in K .$$

Remarque. L s'appelle *constante de Lipschitz*.

28. Soit $A \in \mathcal{L}(E)$ avec $\|A\| < 1$ et considérons le problème

$$x = Ax + b .$$

Montrer que la k -ième approximation, x_k , obtenue par la méthode des approximations successives (avec $x_0 = b$) est égale à l'approximation obtenue en tronquant la série de Neumann pour $(I - A)^{-1}$ après k termes. Trouver ainsi une connection entre l'Exemple 6.4 et l'Exercice 14.

29. Estimation de l'erreur pour la méthode des approximations successives. Soit x^* un point fixe donné par cette méthode. Démontrer

$$\|x^* - x_k\| \leq \frac{\alpha}{1 - \alpha} \|x_k - x_{k-1}\| . \quad (11.3)$$

Pour les fonctions de l'Exemple 6.2, faire quelques itérations et utiliser (11.3) pour obtenir une estimation rigoureuse de l'erreur.

30. Soit E, F des espaces vectoriels normés et $U \subset E$ un fermé. On se pose la question suivante: soit $f : U \rightarrow F$ injective et continue, est-ce que f est un homéomorphisme de U sur $f(U)$?

(a) Démontrer que ceci est vrai si $E = \mathbb{R}^n$ et U borné (utiliser la démonstration de [HW95, Théorème III.3.9]).

(b) Montrer par un contre-exemple que ce résultat n'est plus vrai dans les espaces vectoriels de dimension infinie (considérer par exemple l'identité dans un espace muni de 2 normes non-équivalentes).

31. Considérons la fonction $f : [-1, 1] \rightarrow \mathbb{R}$ donnée par

$$f(x) = \begin{cases} \frac{1}{n+1} \left(2|x| - \frac{1}{n} \right) & \text{si } \frac{1}{n} \leq |x| \leq \frac{1}{n-1} \\ 0 & \text{si } x = 0 \end{cases}.$$

Montrer que f est strictement différentiable à l'origine, mais qu'elle n'est différentiable dans aucun voisinage de 0.

Remarque. $f(x)$ satisfait $f(1/n) = 1/n(n+1)$ et elle est linéaire sur $[1/n, 1/(n+1)]$.

32. Soit $A \in \mathcal{L}(E)$ un isomorphisme. Pour calculer son inverse, on applique la méthode de Newton à l'équation $X^{-1} - A = 0$. Montrer que l'on obtient ainsi l'itération

$$X_{k+1} = X_k + X_k(I - AX_k)$$

et vérifier que la suite $\{X_k\}$ converge vers A^{-1} si $\|I - AX_0\| < 1$.

Indication. $E_k := I - AX_k$ satisfait $E_{k+1} = E_k E_k$.

33. Soit E, F, G des espaces vectoriels normés. Soit $U \subset E$ un ouvert et $a \in U$ et soit $f : U \rightarrow G$ strictement différentiable en a . Soit $V \subset F$ un ouvert avec $b = f(a) \in V$ et soit $g : V \rightarrow G$ strictement différentiable en $b = f(a)$. Démontrer que $g \circ f$ est strictement différentiable en a et que $(g \circ f)'(a) = g'(f(a))f'(a)$.

34. Considérons la fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ donnée par

$$f_1(x_1, x_2) = x_1, \quad f_2(x_1, x_2) = \begin{cases} x_2 - x_1^2 & \text{si } x_1^2 \leq x_2 \\ (x_2^2 - x_1^2 x_2)/x_1^2 & \text{si } 0 \leq x_2 \leq x_1^2 \\ -f_2(x_1, -x_2) & \text{si } x_2 \leq 0 \end{cases}$$

Montrer que

(a) f est partout différentiable.

(b) $f'(0)$ est inversible (calculer $f'(0)$).

(c) f n'est pas strictement différentiable à l'origine.

(d) il n'existe pas de voisinage de l'origine où f est injective (i.e. f n'est pas un homéomorphisme près de l'origine).

35. Donner un difféomorphisme (a) de l'intervalle $(0, 1)$ avec \mathbb{R} , (b) de l'intervalle $(0, 1)$ avec la demi-droite \mathbb{R}_+^* .

36. Soient le "cube" $C = \{x \in \mathbb{R}^n; 0 < x_i < 1, i = 1, \dots, n\}$ et le "simplexe" $S = \{y \in \mathbb{R}^n; y_i > 0, \sum_{i=1}^n y_i < 1\}$ et on considère l'application $\varphi : C \rightarrow \mathbb{R}^n$ donnée par $\varphi(x) = y$ avec $y_k = (1 - x_1) \cdots (1 - x_{k-1})x_k$ pour $k = 1, \dots, n$.

(a) Pour $n = 2$ dessiner les images par φ des droites parallèles aux axes.

(b) Vérifier que $\sum_{i=1}^n y_i = 1 - \prod_{i=1}^n (1 - x_j)$.

(c) Montrer que φ est un difféomorphisme de C sur S et trouver son inverse.

37. Les coordonnées sphériques de \mathbb{R}^n sont données par $\psi(r, \theta_1, \dots, \theta_{n-1}) = x$ où

$$\begin{aligned} x_1 &= r \cos \theta_1 \\ x_2 &= r \sin \theta_1 \cos \theta_2 \\ x_3 &= r \sin \theta_1 \sin \theta_2 \cos \theta_3 \\ &\vdots \\ x_{n-1} &= r \sin \theta_1 \sin \theta_2 \dots \sin \theta_{n-2} \cos \theta_{n-1} \\ x_n &= r \sin \theta_1 \sin \theta_2 \dots \sin \theta_{n-2} \sin \theta_{n-1}. \end{aligned}$$

Montrer que

$$\det \psi' = r^{n-1} \sin^{n-2} \theta_1 \sin^{n-3} \theta_2 \cdots \sin \theta_{n-2},$$

et montrer que la restriction de l'application ψ à $(0, \infty) \times (0, \pi)^{n-2} \times (-\pi, \pi)$ est un difféomorphisme.

38. Soit α_n le volume de la sphère $\{x \in \mathbb{R}^n; \|x\|_2 \leq 1\}$ dans \mathbb{R}^n . En utilisant les résultats de l'exercice 37, le théorème sur les changements de coordonnées ([HW95, Théorème IV.5.7]) et la définition de la fonction Gamma $\Gamma(x)$ ([HW95, Définition III.8.10]) montrer que

$$\alpha_n = \frac{\pi^{n/2}}{\Gamma(n/2 + 1)}.$$

Montrer également que $\alpha_n \rightarrow 0$ si $n \rightarrow \infty$.

Indication. Pour $I_k \equiv \int_0^\pi \sin^k t dt$ on a $(k+1)I_{k+1} = kI_{k-1}$.

39. Dans le plan avec coordonnées cartésiennes (x, y) on considère les deux courbes

$$\begin{aligned} C_1 &= \{(x, y); x^2 - y = 0\} \\ C_2 &= \{(x, y); x - y^2 = 0\}. \end{aligned}$$

Trouver un difféomorphisme local près de $(0, 0)$ tel que

$$\begin{aligned} f(C_1) &= \{(u, v); v = 0\} \\ f(C_2) &= \{(u, v); u = 0\}. \end{aligned}$$

40. Pour tout $a \in \mathbb{R}$ on considère la fonction $f_a : \mathbb{R} \rightarrow \mathbb{R}$ donnée par $f_a(x) = ax - \sin x$. Pour quelles valeurs de a la fonction f_a est-elle un homéomorphisme? un difféomorphisme? un difféomorphisme local près de l'origine?

41. Soit $f : \mathbb{R}^4 \rightarrow \mathbb{R}^2$ l'application définie par

$$\begin{aligned} f_1(x, y, u, v) &= u^3 + vx + y \\ f_2(x, y, u, v) &= uy + v^3 - x \end{aligned}$$

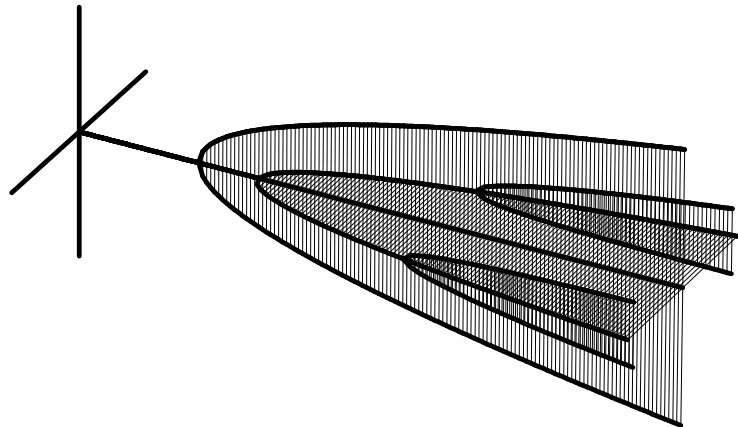
et soit $P_0 = (x_0, y_0, u_0, v_0)$ un point de $f^{-1}(\{0\})$.

- (a) Ecrire une condition sur P_0 qui permet d'obtenir localement près de ce point l'ensemble $f^{-1}(\{0\})$ comme le graphe d'une fonction g de la forme $u = g_1(x, y)$, $v = g_2(x, y)$.
 (b) Calculer $\frac{\partial g_1}{\partial x}(x_0, y_0)$ si $P_0 = (0, -1, 1, 1)$.

42. On considère le système

$$\begin{aligned} \lambda c_1 &= 3c_1 + \frac{9}{4}c_1^3 + \frac{9}{2}c_1c_2^2 \\ \lambda c_2 &= 2c_2 + 3c_1^2c_2 + \frac{3}{2}c_2^2. \end{aligned}$$

Trouver les points de bifurcations des courbes $c_1(\lambda)$ et $c_2(\lambda)$. Une condition nécessaire pour un point de bifurcation est que le théorème de fonctions implicites ne peut pas être appliqué.



43. Démontrer que le théorème d'inversion locale (Théorème 7.5) peut être déduit du théorème des fonctions implicites.
 44. Soit $S = \{(x, y, z); xz + \sin(xy) + \cos(xz) = 1\}$. Déterminer si dans un voisinage de $(0, 1, 1)$ l'ensemble S peut être écrit sous la forme

$$z = f(x, y) \quad \text{ou} \quad y = g(x, z) \quad \text{ou} \quad x = h(y, z).$$

45. Calculer la norme du déterminant vu comme application multilinéaire

$$\det : \mathbb{R}^n \times \cdots \times \mathbb{R}^n \rightarrow \mathbb{R} \quad (n \text{ facteurs}),$$

où \mathbb{R}^n est munie de la norme euclidienne. En déduire l'inégalité $|\det A| \leq n^{n/2}$ (de Hadamard) pour une matrice $A = (a_{ij})_{1 \leq i, j \leq n}$ satisfaisant $|a_{ij}| \leq 1$.

46. Soient \mathbb{R}^n et \mathbb{R}^m munis de la norme euclidienne. Soit $b : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}$ l'application bilinéaire donnée par $b(x, y) = x^T B y$, où B est une matrice arbitraire. Montrer que la norme de b satisfait $\|b\| = \|B\|_2$.
47. Soient $f : E \rightarrow F$ et $g : F \rightarrow G$ (où E , F et G sont des espaces vectoriels normés) deux applications deux fois différentiables. Calculer $(g \circ f)''(x)(h, k)$.
48. Soit $b : E \times F \rightarrow G$ (où E , F et G sont des espaces vectoriels normés) une application bilinéaire continue. Calculer sa deuxième dérivée.
49. Calculer la deuxième dérivée de $f : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ donnée par
- $$f(x) = \left(x_1^3 + x_1 x_2^3, e^{(x_1 - x_2)}, \cos(x_1 x_2) \right)^T.$$
50. Soit $X \in \text{GL}(\mathbb{R}^n)$, calculer toutes les dérivées de la fonction $f(X) = X^3$.

Chapitre II

Maxima et minima relatifs et calcul de variations

Ce chapitre est consacré à l'étude des extrema d'une fonction à valeurs réelles. Il est bien connu que pour une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ la condition $f'(a) = 0$ est nécessaire et les conditions $f'(a) = 0, f''(a) > 0$ sont suffisantes pour que a soit un minimum relatif. Nous allons étendre ce résultat à des fonctions $f : U \rightarrow \mathbb{R}$ où U est un ouvert d'un espace vectoriel normé E . Nous étudierons également des conditions pour que la restriction $f|_{\mathcal{M}}$, où \mathcal{M} est une sous-variété de la forme $\mathcal{M} = \{x \in U ; g(x) = 0\}$, possède un extremum en a . Une application importante est la situation où E est un espace de fonctions. Les conditions nécessaires de l'extremum s'expriment alors par une équation différentielle, l'équation d'Euler-Lagrange.

II.1 Maxima et minima relatifs

Soit U un ouvert d'un espace vectoriel normé E et $f : U \rightarrow \mathbb{R}$. On dit que f possède un *minimum relatif* au point $a \in U$, s'il existe un voisinage V de a ($V \subset U$) tel que

$$f(x) \geq f(a) \quad \text{pour tout } x \in V.$$

Un minimum relatif est dit *strict*, si

$$f(x) > f(a) \quad \text{pour tout } x \in V \setminus \{a\}.$$

De la même manière on définit un maximum relatif ou un maximum relatif strict. On dit que f possède un *extremum relatif* en a , si elle possède un minimum ou un maximum relatif en a .

Théorème 1.1 (condition nécessaire) *Soit U un ouvert d'un espace vectoriel normé E et supposons que $f : U \rightarrow \mathbb{R}$ soit différentiable en $a \in U$. Si f admet un extremum relatif en a , alors $f'(a) = 0$.*

Démonstration. Supposons que f admette un minimum relatif en a (le cas d'un maximum relatif est traité de la même manière). Les hypothèses impliquent que pour un $h \in E$ fixé et pour $|\varepsilon|$ suffisamment petit

$$0 \leq f(a + \varepsilon h) - f(a) = f'(a)\varepsilon h + r(a + \varepsilon h) \cdot |\varepsilon| \cdot \|h\|$$

où $r(a + \varepsilon h) \rightarrow 0$ si $\varepsilon \rightarrow 0$. En divisant cette relation par ε et en considérant la limite $\varepsilon \rightarrow 0$ on obtient $f'(a)h \geq 0$. On en déduit $f'(a) = 0$, car $h \in E$ est arbitraire. \square

Le lemme suivant représente un cas particulier de la série de Taylor avec reste.

Lemme 1.2 Soit $f : U \rightarrow \mathbb{R}$ (U un ouvert d'un espace vectoriel normé E) deux fois différentiable en $a \in U$. Alors, on a

$$f(a+h) = f(a) + f'(a)h + \frac{1}{2!} f''(a)(h, h) + s(h)\|h\|^2, \quad (1.1)$$

où $s(h) \rightarrow 0$ si $h \rightarrow 0$.

Démonstration. Considérons la fonction

$$g(h) := f(a+h) - f(a) - f'(a)h - \frac{1}{2!} f''(a)(h, h),$$

définie dans un voisinage de $0 \in E$. On a $g(0) = 0$ et la dérivée de $g(h)$ est donnée par

$$g'(h) = f'(a+h) - f'(a) - f''(a)(h, \cdot) = r(h)\|h\|$$

avec $r(h) \rightarrow 0$ si $h \rightarrow 0$, car $f'(x)$ est différentiable en a . Par le théorème des accroissements finis, on obtient alors

$$|g(h)| \leq \sup_{0 < t < 1} \|g'(th)\| \cdot \|h\| \leq \sup_{0 < t < 1} \|r(th)\| \cdot \|h\|^2.$$

Ceci implique que la fonction $s(h)$, définie par (1.1), satisfait $s(h) \rightarrow 0$ si $h \rightarrow 0$. □

Théorème 1.3 (condition nécessaire) Soit U un ouvert d'un espace vectoriel normé et $f : U \rightarrow \mathbb{R}$ deux fois différentiable en $a \in U$. Si f possède un minimum relatif en a , alors

$$f''(a)(h, h) \geq 0 \quad \text{pour tout } h \in E. \quad (1.2)$$

Démonstration. On remplace h par εh dans (1.1), on divise la formule obtenue par ε^2 et on considère la limite $\varepsilon \rightarrow 0$. □

On dit que $a \in U$ est un *point critique* de $f : U \rightarrow \mathbb{R}$ si $f'(a) = 0$. La condition (1.2) ne suffit pas pour qu'un point critique soit un minimum relatif de f . La fonction $f(x) = x^3$ nous sert comme contre-exemple.

Théorème 1.4 (condition suffisante) Soit U un ouvert d'un espace vectoriel normé et $f : U \rightarrow \mathbb{R}$ deux fois différentiable en $a \in U$. Si $f'(a) = 0$ et si $f''(a)$ satisfait

$$f''(a)(h, h) \geq \gamma \|h\|^2 \quad \text{pour tout } h \in E \quad (1.3)$$

avec un $\gamma > 0$, alors f admet un minimum relatif strict au point a .

Démonstration. Les hypothèses sur $f'(a)$ et $f''(a)$ et la formule (1.1) impliquent que

$$f(a+h) - f(a) \geq \left(\frac{\gamma}{2} - \|s(h)\| \right) \|h\|^2.$$

Comme $s(h) \rightarrow 0$ si $h \rightarrow 0$, on en déduit que $f(a+h) > f(a)$ pour $h \neq 0$ suffisamment petit. □

Exemple 1.5 Pour le cas particulier $U \subset \mathbb{R}^n$, où $f : U \rightarrow \mathbb{R}$ est une fonction à n variables réelles, la condition $f'(a) = 0$ devient

$$\frac{\partial f}{\partial x_1}(a_1, \dots, a_n) = 0, \quad \dots, \quad \frac{\partial f}{\partial x_n}(a_1, \dots, a_n) = 0,$$

ce qui constitue un système de n équations pour les n inconnues a_1, \dots, a_n . La condition (1.3) est équivalente au fait que la matrice Hessienne

$$H(a) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(a) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(a) \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(a) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(a) \end{pmatrix}$$

est une matrice définie positive, c.-à-d., toutes les valeurs propres sont positives. La valeur de γ correspond à la plus petite valeur propre de $H(a)$.

II.2 Multiplicateurs de Lagrange

Soit U un ouvert d'un espace vectoriel normé E , $f : U \rightarrow \mathbb{R}$, et considérons la sous-variété $\mathcal{M} = \{x \in U ; g(x) = 0\}$ où $g : U \rightarrow \mathbb{R}^m$. Le problème consiste à trouver les minima (ou maxima) de la restriction $f|_{\mathcal{M}}$, c.-à-d., on cherche $a \in \mathcal{M}$ tel que

$$f(x) \geq f(a) \quad \text{pour tout } x \in V \cap \mathcal{M},$$

où V est un voisinage de a dans U .

Exemple 2.1 Soit $U = \mathbb{R}^2$, $f(x_1, x_2) = x_1 x_2$ et $g(x_1, x_2) = x_1^2 + x_2^2 - 1$. On cherche alors les extrema de la fonction $f(x)$ sur le cercle de rayon 1. Par un raisonnement géométrique (voir [HW95, page 325]) on trouve la condition nécessaire

$$\text{grad } f(a_1, a_2) = \lambda \text{grad } g(a_1, a_2)$$

pour un extremum ($\text{grad } f(a)$, qui est orthogonale aux courbes de niveau de f , doit avoir la même direction que $\text{grad } g(a)$). Dans notre exemple on obtient

$$a_2 = 2\lambda a_1, \quad a_1 = 2\lambda a_2, \quad a_1^2 + a_2^2 = 1$$

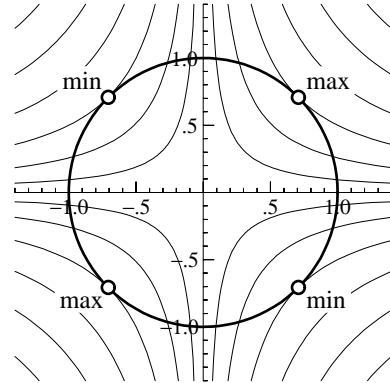
et on trouve sans difficulté les quatre solutions $(\pm\sqrt{2}/2, \pm\sqrt{2}/2)$.

Nous étendons ce résultat à la situation où U est un ouvert d'un espace vectoriel normé (éventuellement de dimension infinie) et où la sous-variété \mathcal{M} est donnée par m équations.

Théorème 2.2 (condition nécessaire) Soit $U \subset E$ un ouvert d'un espace de Banach E , et supposons que $f : U \rightarrow \mathbb{R}$ et $g : U \rightarrow \mathbb{R}^m$ soient continûment différentiables dans un voisinage de $a \in \mathcal{M} = \{x \in U ; g(x) = 0\}$. Si $g'(a)$ est surjective et si $f|_{\mathcal{M}}$ possède un extremum relatif en a , alors il existe des nombres $\lambda_1, \dots, \lambda_m$ (multiplicateurs de Lagrange) tels que

$$f'(a) - \sum_{i=1}^m \lambda_i g'_i(a) = 0 \tag{2.1}$$

($g_i(x)$ est la i ème composante du vecteur $g(x)$).



Remarque. Avec la fonction de Lagrange

$$\mathcal{L}(x, \lambda) := f(x) - \sum_{i=1}^m \lambda_i g_i(x) \quad (2.2)$$

les conditions $a \in \mathcal{M}$ et (2.1) sont équivalentes à $\mathcal{L}'(a, \lambda) = 0$.

Démonstration. Par la surjectivité de $g'(a)$ on peut trouver des vecteurs $h_1, \dots, h_m \in E$ tels que $g'(a)h_j = e_j = (0, \dots, 0, 1, 0, \dots, 0)^T$, c.-à-d., $g'_i(a)h_j = \delta_{ij}$ où $\delta_{ii} = 1$ et $\delta_{ij} = 0$ pour $i \neq j$. Considérons alors la fonction

$$G(x, t) := g(x + t_1 h_1 + \dots + t_m h_m)$$

avec $t = (t_1, \dots, t_m)^T \in \mathbb{R}^m$. On a $G(a, 0) = 0$ et

$$\frac{\partial G}{\partial t}(a, 0) = (g'(a)h_1, \dots, g'(a)h_m) = I \quad (\text{l'identité}).$$

Le théorème des fonctions implicites implique que, proche de $(a, 0)$, il existe une fonction différentiable $t(x) = (t_1(x), \dots, t_m(x))^T$ telle que $G(x, t(x)) = 0$. Le vecteur $y := x + h_1 t_1(x) + \dots + h_m t_m(x)$ satisfait donc $g(y) = 0$ pour tout x dans un voisinage de a . La dérivée de $g_i(y) = 0$ par rapport à x donne alors au point a

$$0 = g'_i(a)(I + h_1 t'_1(a) + \dots + h_m t'_m(a)) = g'_i(a) + t'_i(a).$$

Comme $f|_{\mathcal{M}}$ possède un extremum relatif en a , la condition nécessaire du Théorème 1.1 implique que la dérivée de $f(x + h_1 t_1(x) + \dots + h_m t_m(x))$ est nulle au point $x = a$, c.-à-d.,

$$f'(a)(I + h_1 t'_1(a) + \dots + h_m t'_m(a)) = 0.$$

En posant $\lambda_i := f'(a)h_i$ et en utilisant la relation $t'_i(a) = -g'_i(a)$, on obtient l'affirmation (2.1). \square

Pour le cas particulier $E = \mathbb{R}^n$, la surjectivité de $g'(a)$ s'exprime par le fait que la matrice

$$g'(a) = \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(a) & \cdots & \frac{\partial g_1}{\partial x_n}(a) \\ \vdots & & \vdots \\ \frac{\partial g_m}{\partial x_1}(a) & \cdots & \frac{\partial g_m}{\partial x_n}(a) \end{pmatrix}$$

possède le rang m .

Exemple 2.3 Soient

$$\begin{aligned} f(x_1, x_2, x_3) &= x_1 - x_2 - x_3 \\ g_1(x_1, x_2, x_3) &= x_1^2 + 2x_2^2 - 1 \\ g_2(x_1, x_2, x_3) &= 3x_1 - 4x_3 \end{aligned}$$

et cherchons les extrema de $f|_{\mathcal{M}}$ où $\mathcal{M} := \{x \in \mathbb{R}^3 ; g_1(x) = g_2(x) = 0\}$. L'ensemble \mathcal{M} représente l'intersection d'un cylindre elliptique avec un plan. On vérifie aisément que $g'(x)$ possède le rang 2 pour tout $x \in \mathcal{M}$. Considérons alors la fonction de Lagrange

$$\mathcal{L}(x, \lambda) = x_1 - x_2 - x_3 - \lambda_1(x_1^2 + 2x_2^2 - 1) - \lambda_2(3x_1 - 4x_3).$$

Les conditions nécessaires (2.1) nous donnent

$$1 - 2\lambda_1 x_1 - 3\lambda_2 = 0, \quad -1 - 4\lambda_1 x_2 = 0, \quad -1 + 4\lambda_2 = 0.$$

On obtient donc $\lambda_2 = 1/4$ et, en éliminant λ_1 de deux autres équations, on trouve $x_2 + 2x_1 = 0$. Les solutions du système (2.1) sont alors $(1/3, -2/3, 1/4)$ et $(-1/3, 2/3, -1/4)$.

Théorème 2.4 (condition suffisante) Soit $U \subset E$ un ouvert d'un espace de Banach E . Supposons de plus que $f : U \rightarrow \mathbb{R}$ et $g : U \rightarrow \mathbb{R}^m$ soient deux fois différentiables en $a \in \mathcal{M} := \{x \in U ; g(x) = 0\}$ et que $g'(a)$ soit surjective. Si (2.1) est satisfait et si

$$\left(f''(a) - \sum_{i=1}^m \lambda_i g_i''(a)\right)(h, h) \geq \gamma \|h\|^2 \quad \text{pour } h \text{ satisfaisant } g'(a)h = 0, \quad (2.3)$$

avec un $\gamma > 0$, alors $f|_{\mathcal{M}}$ admet un minimum relatif strict au point a .

Démonstration. En appliquant le Lemme 1.2 à la fonction $f(x) - \sum_{i=1}^m \lambda_i g_i(x)$ et en utilisant (2.1) nous obtenons pour $a + k \in \mathcal{M}$

$$f(a + k) - f(a) = \frac{1}{2} \left(f''(a) - \sum_{i=1}^m \lambda_i g_i''(a)\right)(k, k) + s(k) \|k\|^2, \quad (2.4)$$

où $s(k) \rightarrow 0$ si $k \rightarrow 0$. Considérons maintenant k tel que $a + k \in \mathcal{M}$. La difficulté est que k ne satisfait pas nécessairement $g'(a)k = 0$ et qu'on ne peut pas directement utiliser (2.3).

Pour le vecteur $k \in E$ nous considérons

$$h = k - \alpha_1 h_1 - \dots - \alpha_m h_m \quad (2.5)$$

où les vecteurs h_1, \dots, h_m sont comme dans la démonstration du Théorème 2.2. Nous choisissons $\alpha_i = g'_i(a)k$ afin que h soit dans $\ker g'(a)$. Si $a + k \in \mathcal{M}$ nous avons (de nouveau par le Lemme 1.2) que

$$0 = g(a + k) - g(a) = g'(a)k + \frac{1}{2} g''(a)(k, k) + r(k) \|k\|^2$$

avec $r(k) \rightarrow 0$ pour $k \rightarrow 0$. Ceci implique $\|g'(a)k\| \leq c_1 \|k\|^2$ pour k suffisamment petit, et en utilisant (2.5) aussi $\|h - k\| \leq c_2 \|k\|^2$. L'inégalité $|\|h\| - \|k\|| \leq \|h - k\|$ montre alors que $c_3 \|k\| \leq \|h\| \leq c_4 \|k\|$ si k est suffisamment petit. L'expression de droite de l'équation (2.4) peut donc être écrite sous la forme

$$\frac{1}{2} \left(f''(a) - \sum_{i=1}^m \lambda_i g_i''(a)\right)(h, h) + \hat{s}(k) \|h\|^2,$$

où $\hat{s}(k) \rightarrow 0$ pour $k \rightarrow 0$. Par la condition (2.3), cette expression est positive pour $k \neq 0$ suffisamment petit. Donc, $f(a + k) - f(a)$ est strictement positive pour $k \neq 0$ suffisamment petit, si $a + k \in \mathcal{M}$. \square

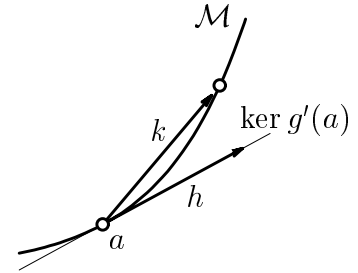
Exemple 2.5 Utilisons la condition (2.3) du Théorème 2.4 pour démontrer que le point $a = (-1/3, 2/3, -1/4)$ est un minimum relatif strict de la fonction $f|_{\mathcal{M}}$ de l'Exemple 2.3. Comme $\lambda_1 = -3/8$ et $\lambda_2 = 1/4$ nous avons pour $h = (h_1, h_2, h_3)^T$ que

$$\left(f''(a) - \lambda_1 g_1''(a) - \lambda_2 g_2''(a)\right)(h, h) = \frac{3}{8}(2h_1^2 + 4h_2^2).$$

Pour h satisfaisant $g'(a)h = 0$ (en particulier $3h_1 - 4h_3 = 0$) on obtient

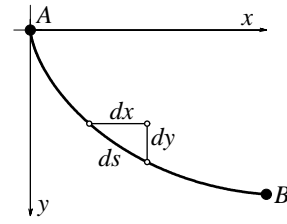
$$\frac{3}{8}(2h_1^2 + 4h_2^2) = \frac{3}{8}\left(h_1^2 + 4h_2^2 + \frac{16}{9}h_3^2\right) \geq \frac{3}{8} \|h\|^2$$

et la condition suffisante pour un minimum strict est vérifiée.



II.3 Calcul de variations

En 1696, Johann Bernoulli posait le problème suivant: *une particule glisse sous l'influence de la gravité sur une courbe reliant A et B (dans un plan vertical). Trouver la courbe pour laquelle le temps qu'il lui faut pour aller de A à B est minimal.* Une telle courbe s'appelle *brachistochrone*. La solution de ce problème, proposée par Jakob Bernoulli, est à l'origine du calcul de variations.



La formulation mathématique part du principe de la conservation de l'énergie mécanique

$$\frac{m}{2} \left(\frac{ds}{dt} \right)^2 - mgy = 0$$

(l'origine est au point A). Ceci implique $ds/dt = \sqrt{2gy}$. Comme $ds = \sqrt{dx^2 + dy^2} = \sqrt{1 + y'(x)^2} dx$, le problème revient à minimiser l'expression

$$T(y) = \int_0^{t_1} dt = \int_0^{t_1} \frac{ds}{\sqrt{2gy}} = \int_0^b \frac{\sqrt{1 + y'(x)^2}}{\sqrt{2gy(x)}} dx, \quad (3.1)$$

où T est un opérateur de $\mathcal{C}^1([0, b])$ dans \mathbb{R} .

Problème Considérons une fonction continue $L(x, y, y')$, définie dans un ouvert de \mathbb{R}^3 et à valeurs dans \mathbb{R} . Le problème consiste à trouver une fonction $y(x)$ telle que

$$T(y) = \int_a^b L(x, y(x), y'(x)) dx \quad (3.2)$$

soit minimale parmi toutes les fonctions de classe \mathcal{C}^1 satisfaisant $y(a) = A$ et $y(b) = B$. La fonction $L(x, y, y')$ s'appelle *fonction de Lagrange* ou *Lagrangien*.

Lemme 3.1 Soit $y \in \mathcal{C}^1([a, b])$ une fonction qui vérifie $y(a) = A$, $y(b) = B$. Soit de plus $L(x, y, y')$ une fonction continûment différentiable dans un voisinage de $\{(x, y(x), y'(x)) \in \mathbb{R}^3; x \in [a, b]\}$. Si l'expression $T(y)$ de (3.2) possède un extremum en y , alors

$$\int_a^b \left(\frac{\partial L}{\partial y}(x, y(x), y'(x)) h(x) + \frac{\partial L}{\partial y'}(x, y(x), y'(x)) h'(x) \right) dx = 0 \quad (3.3)$$

pour tout $h \in \mathcal{C}^1([a, b])$ vérifiant $h(a) = h(b) = 0$.

Démonstration. Considérons la fonction

$$S(\varepsilon) := T(y + \varepsilon h) = \int_a^b L(x, y(x) + \varepsilon h(x), y'(x) + \varepsilon h'(x)) dx,$$

où y satisfait les hypothèses du lemme, $h \in \mathcal{C}^1([a, b])$ avec $h(a) = h(b) = 0$, et ε est un paramètre réel (petit). Si y est un extremum de $T(y)$, 0 en est un de $S(\varepsilon)$, et la condition nécessaire $S'(0) = 0$ doit être satisfaite. L'affirmation du lemme est une conséquence du fait que l'expression de (3.3) est égale à $S'(0)$ (les hypothèses garantissent qu'on peut échanger la dérivation avec l'intégration). \square

Une autre possibilité serait de considérer l'espace vectoriel $E = \{h \in \mathcal{C}^1([a, b]); h(a) = h(b) = 0\}$, de le munir d'une norme, et de considérer l'application $F : E \rightarrow \mathbb{R}$, définie par $F(h) := T(y+h)$. Pour que y soit un extremum de T , $h = 0$ doit être un extremum de F . La condition nécessaire $F'(0) = 0$ (où bien $F'(0)h = 0$ pour tout $h \in E$) du Théorème 1.1 est équivalente à (3.3). Les résultats des Théorèmes 1.3 et 1.4 peuvent également être appliqués.

Définition 3.2 Une fonction $y \in \mathcal{C}^1([a, b])$ qui satisfait $y(a) = A$, $y(b) = B$ et la condition (3.3) pour tout $h \in \mathcal{C}^1([a, b])$ vérifiant $h(a) = h(b) = 0$, s'appelle une *extrémale* du problème (3.2).

Pour trouver la solution du problème variationnel, nous cherchons à éliminer la présence de $h(x)$ dans la condition nécessaire (3.3).

Théorème 3.3 (Equation d'Euler-Lagrange) Soit $L(x, y, y')$ de classe \mathcal{C}^1 . Une fonction $y \in \mathcal{C}^1([a, b])$ est une extrémale du problème (3.2) aux extrémités fixées $y(a) = A$ et $y(b) = B$, si et seulement si

$$\begin{aligned} \frac{\partial L}{\partial y'}(x, y(x), y'(x)) & \text{ est continûment différentiable et} \\ \frac{\partial L}{\partial y}(x, y(x), y'(x)) - \frac{d}{dx} \frac{\partial L}{\partial y'}(x, y(x), y'(x)) & = 0. \end{aligned} \quad (3.4)$$

Démonstration. Suffisance. Notons $A(x) = \frac{\partial L}{\partial y}(x, y(x), y'(x))$ et $B(x) = \frac{\partial L}{\partial y'}(x, y(x), y'(x))$. La condition (3.4), c.-à-d., $A(x) = B'(x)$, implique que

$$\int_a^b (A(x)h(x) + B(x)h'(x)) dx = \int_a^b (B(x)h(x))' dx = B(x)h(x)|_a^b = 0$$

pour $h(a) = h(b) = 0$. Donc, $y(x)$ est une extrémale de (3.2).

Nécessité. La condition (3.3) s'écrit comme $\int_a^b (A(x)h(x) + B(x)h'(x)) dx = 0$. On intègre par parties la première expression de cette intégrale. Si $C(x)$ est une primitive de $A(x)$, c.-à-d., $C(x)$ est continûment différentiable et $C'(x) = A(x)$, on a que

$$0 = \int_a^b (A(x)h(x) + B(x)h'(x)) dx = \int_a^b (B(x) - C'(x))h'(x) dx$$

pour tout $h \in \mathcal{C}^1([a, b])$ satisfaisant $h(a) = h(b) = 0$. En appliquant le Lemme de Du Bois Reymond (Lemme 3.4) on obtient $B(x) - C'(x) = \text{Const}$. Donc, $B(x)$ est continûment différentiable et $B'(x) = C'(x) = A(x)$. \square

Lemme 3.4 (Du Bois Reymond) Soit $d : [a, b] \rightarrow \mathbb{R}$ continue et

$$\int_a^b d(x)h'(x) dx = 0 \quad \text{pour tout } h \in \mathcal{C}^1([a, b]) \text{ vérifiant } h(a) = h(b) = 0,$$

alors $d(x) = \text{Const}$.

Démonstration. Posons

$$m := \frac{1}{b-a} \int_a^b d(x) dx \quad \text{et} \quad h(x) := \int_a^x (d(t) - m) dt.$$

Cette fonction h est dans $\mathcal{C}^1([a, b])$, elle vérifie $h(a) = h(b) = 0$ et on a $h'(x) = d(x) - m$. On obtient donc

$$\int_a^b (d(x) - m)^2 dx = \int_a^b (d(x) - m)h'(x) dx = -m \int_a^b h'(x) dx = -m(h(b) - h(a)) = 0.$$

Comme $d(x) - m$ est continue par hypothèse, ceci est possible seulement si $d(x) - m = 0$. \square

Si $L(x, y, y')$ et $y(x)$ sont de classe \mathcal{C}^2 , l'équation d'Euler-Lagrange peut être écrite sous la forme

$$\frac{\partial L}{\partial y}(x, y, y') - \frac{\partial^2 L}{\partial x \partial y'}(x, y, y') - \frac{\partial^2 L}{\partial y \partial y'}(x, y, y') y' - \frac{\partial^2 L}{\partial y' \partial y'}(x, y, y') y'' = 0, \quad (3.5)$$

qui est une équation différentielle de deuxième ordre. Le théorème suivant nous permet de réduire l'ordre de l'équation différentielle dans la situation où $L(x, y, y')$ ne dépend que de y et y' .

Théorème 3.5 *Soit le Lagrangien L de classe \mathcal{C}^1 et indépendant de la variable x . Alors, toute extrémale $y(x)$ de classe \mathcal{C}^2 vérifie la relation*

$$L(y(x), y'(x)) - \frac{\partial L}{\partial y'}(y(x), y'(x)) y'(x) = \text{Const}. \quad (3.6)$$

Démonstration. La dérivée de $L(y(x), y'(x)) - \frac{\partial L}{\partial y'}(y(x), y'(x)) y'(x)$ par rapport à x donne

$$\frac{\partial L}{\partial y}(\dots) y'(x) + \frac{\partial L}{\partial y'}(\dots) y''(x) - \frac{d}{dx} \left(\frac{\partial L}{\partial y'}(\dots) \right) y'(x) - \frac{\partial L}{\partial y'}(\dots) y''(x),$$

ce qui est nul par (3.4). □

Exemple 3.6 (Brachystochrone) Le problème consiste à minimiser l'intégrale (3.1) parmi toutes les fonctions satisfaisant $y(0) = 0$ et $y(b) = B$ (b et B sont des nombres positifs). La condition nécessaire (3.6) devient

$$\frac{\sqrt{1 + y'^2}}{\sqrt{y}} - \frac{y'}{\sqrt{y} \sqrt{1 + y'^2}} \cdot y' = \text{Const},$$

ce qui est équivalent à $1 + y'^2 - y'^2 = \text{Const} \cdot \sqrt{y} \sqrt{1 + y'^2}$. On en déduit l'équation différentielle $y(1 + y'^2) = C$, dont la solution générale est une cycloïde donnée par (voir [HW95, page 138])

$$x(\varphi) = \frac{C}{2}(\varphi - \sin \varphi) + D, \quad y(\varphi) = \frac{C}{2}(1 - \cos \varphi). \quad (3.7)$$

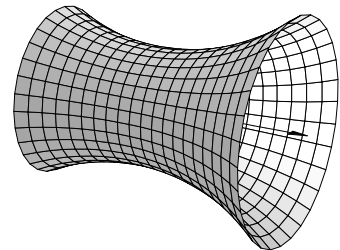
La Fig. 3.1 (le dessin de gauche) montre les solutions pour $0 \leq \varphi \leq 2\pi$ et pour $D = 0$ (pour qu'elles passent par l'origine). On voit que, quelque soit $b > 0$ et $B > 0$, il existe une unique valeur de C telle que la courbe (3.7) passe par (b, B) pour un $\varphi \in (0, 2\pi)$.

Attention. Ce problème n'entre pas exactement dans le cadre de notre théorie, car $L(y, y')$ n'est pas définie pour $y = 0$. Nous ne discutons pas les détails techniques pour justifier la solution.

Exemple 3.7 (Surface de révolution d'aire minimale)

On se donne $a < b$, $A > 0$, $B > 0$, et on cherche une fonction $y(x)$ de classe \mathcal{C}^1 telle que $y(a) = A$, $y(b) = B$ et que la surface de révolution $\{(x, y(x) \cos \varphi, y(x) \sin \varphi) ; x \in [a, b], \varphi \in [0, 2\pi]\}$ possède une aire minimale. On a la formule

$$\text{Aire} = 2\pi \int y ds = 2\pi \int_a^b y(x) \sqrt{1 + y'(x)^2} dx,$$



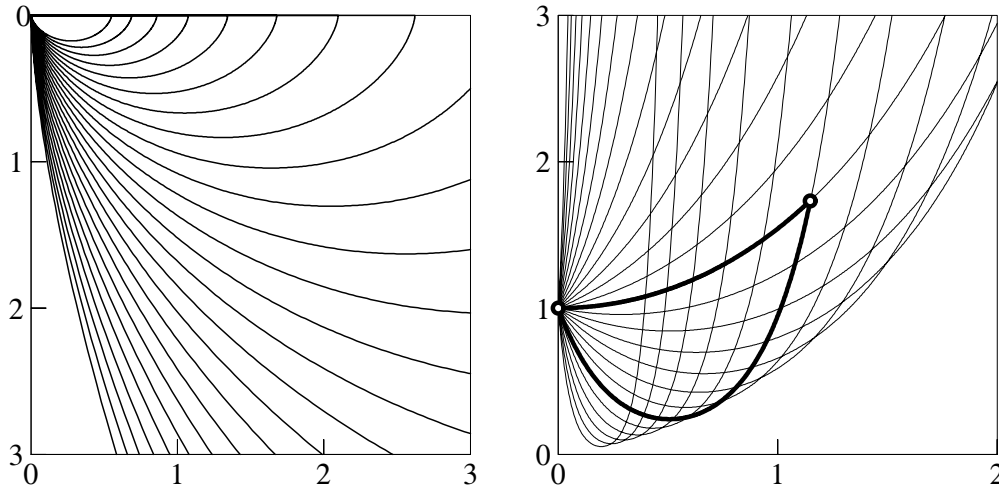


FIG. 3.1 – *Brachystochrone (gauche) et surface de révolution d'aire minimale (droite)*

et l'équation d'Euler-Lagrange sous la forme (3.6) donne

$$y\sqrt{1+y'^2} - \frac{y \cdot y'}{\sqrt{1+y'^2}} \cdot y' = Const.$$

On en déduit

$$y^2 = C^2(1+y'^2) \quad \text{ou} \quad \frac{dy}{dx} = \sqrt{\frac{y^2}{C^2} - 1}.$$

On résout cette équation différentielle en utilisant la séparation des variables et la substitution $y = C \cosh u$. Ceci donne $u = x/C + \lambda$ ou $y(x) = C \cosh(x/C + \lambda)$ comme solution générale. En choisissant $a = 0$ et $A = 1$ (normalisation) la condition $y(0) = 1$ devient $1 = C \cosh \lambda$, et on obtient pour la solution la représentation

$$y(x) = \frac{1}{\cosh \lambda} \cosh(x \cdot \cosh \lambda + \lambda). \tag{3.8}$$

La Fig. 3.1 (le dessin de droite) montre qu'en général ce problème possède deux extrémales (par exemple pour $b = 1.148$ et $B = 1.731$ on obtient $\lambda = -2.1$ et $\lambda = 0$). Pour la courbe supérieure (celle avec le plus grand λ) l'aire de la surface de révolution est minimale (sans démonstration, pour plus de détails voir [Fox, page 48]).

Si le point (b, B) est en-dessous de l'enveloppe des courbes (3.8), il n'y a pas d'extrémale. Dans cette situation, la fonction non-continue $y(x) = 0$ pour $0 < x < b$, $y(0) = 1$ et $y(b) = B$ donne la valeur minimale de l'aire (Aire = $\pi(1+B^2)$). Elle s'appelle solution de Goldschmidt [Tr83, page 214].

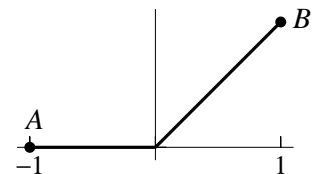
Exemple 3.8 Un exemple plus simple qui n'admet pas d'extrémale est le suivant (du à Weierstraß):

$$\int_{-1}^1 y(x)^2(1-y'(x))^2 dx \rightarrow \min, \quad y(-1) = 0, \quad y(1) = 1.$$

L'équation d'Euler-Lagrange est alors

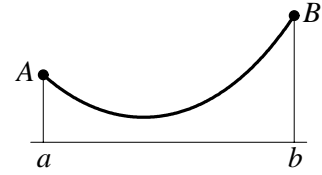
$$y^2(1-y')^2 + 2y^2(1-y')y' = y^2(1-y')(1+y') = Const.$$

Elle n'a pas de solution différentiable qui vérifie les conditions aux bords. L'intégrale est minimale pour la fonction donnée par $y(x) = 0$ pour $-1 \leq x \leq 0$ et par $y(x) = x$ pour $0 \leq x \leq 1$, car elle vaut alors zéro.



II.4 Problèmes isopérimétriques

Considérons le problème suivant, proposé par Euler (1734): *trouver la forme d'une chaîne de longueur ℓ (à l'équilibre), qui est suspendue aux extrémités*. Mathématiquement, ce problème devient



$$\int_a^b y(x) \sqrt{1 + y'(x)^2} dx \rightarrow \min \quad \text{où} \quad \int_a^b \sqrt{1 + y'(x)^2} dx = \ell, \quad (4.1)$$

c.-à-d., l'énergie potentielle est minimale et la longueur est ℓ (condition "isopérimétrique").

Problème général Soient données deux fonctions continues $f(x, y, y')$ et $g(x, y, y')$, définies dans un ouvert de \mathbb{R}^3 et à valeurs dans \mathbb{R} , et $\ell \in \mathbb{R}$. Le *problème isopérimétrique* consiste à trouver une fonction $y(x)$ telle que

$$F(y) := \int_a^b f(x, y(x), y'(x)) dx \quad (4.2)$$

soit minimale parmi toutes les fonctions de classe \mathcal{C}^1 satisfaisant

$$G(y) := \int_a^b g(x, y(x), y'(x)) dx = \ell \quad \text{et} \quad y(a) = A, y(b) = B. \quad (4.3)$$

Nous allons écrire ce problème sous la forme traitée dans le paragraphe II.2. Supposons qu'une solution $y \in \mathcal{C}^1([a, b])$ existe. Comme dans la discussion qui suit le Lemme 3.1, nous introduisons l'espace vectoriel

$$E = \{h \in \mathcal{C}^1([a, b]); h(a) = h(b) = 0\}$$

muni d'une norme et nous considérons les applications $F_0 : E \rightarrow \mathbb{R}$ et $G_0 : E \rightarrow \mathbb{R}$ définies par $F_0(h) = F(y + h)$ et $G_0(h) = G(y + h)$. Le problème consiste alors à montrer que $h = 0$ est un minimum de $F_0(h)$ assujetti à la condition $G_0(h) - \ell = 0$.

Théorème 4.1 Soient $f(x, y, y')$ et $g(x, y, y')$ des fonctions de classe \mathcal{C}^2 , et soit $y \in \mathcal{C}^2([a, b])$ une solution du problème isopérimétrique. Si

$$\frac{\partial g}{\partial y}(x, y(x), y'(x)) - \frac{d}{dx} \frac{\partial g}{\partial y'}(x, y(x), y'(x)) \neq 0, \quad (4.4)$$

alors il existe un $\lambda \in \mathbb{R}$ tel que la fonction

$$L(x, y, y') = f(x, y, y') - \lambda g(x, y, y') \quad (4.5)$$

satisfait l'équation d'Euler-Lagrange (3.4).

Démonstration. Nous allons appliquer le Théorème 2.2 au problème $F_0(h) \rightarrow \min$ assujetti à la restriction $G_0(h) - \ell = 0$. La condition (4.4) implique que $G'_0(0) : E \rightarrow \mathbb{R}$ est surjective. Il existe alors un $\lambda \in \mathbb{R}$ tel que $F'_0(0) - \lambda G'_0(0) = 0$. Mais ceci est exactement la condition (3.3) avec $L(x, y, y')$ pris de (4.5). L'affirmation est alors une conséquence de la démonstration (nécessité) du Théorème 3.3. \square

Exemple 4.2 Pour le problème du début de ce paragraphe nous obtenons

$$L(x, y, y') = (y - \lambda)\sqrt{1 + y'^2}.$$

Après la transformation $y - \lambda = z$, l'équation d'Euler-Lagrange est exactement la même que pour le problème de la surface de révolution d'aire minimale (Exemple 3.7). On obtient donc

$$y(x) = \lambda + C \cosh(x/C + D).$$

Les trois paramètres λ, C, D sont déterminés par $y(a) = A$, $y(b) = B$ et par

$$\ell = \int_a^b \sqrt{1 + y'(x)^2} dx = \int_a^b \cosh(x/C + D) dx = C \sinh(x/C + D) \Big|_a^b.$$

II.5 Géodésiques des surfaces

Avant de discuter le calcul de géodésiques des surfaces (c.-à-d., des courbes de longueur minimale sur une surface donnée), nous généralisons les résultats du paragraphe II.3 au cas de plusieurs fonctions.

Considérons une fonction $L(x, y_1, \dots, y_n, y'_1, \dots, y'_n)$. Le problème consiste à trouver n fonctions $y_1(x), \dots, y_n(x)$ telles que

$$\int_a^b L(x, y_1(x), \dots, y_n(x), y'_1(x), \dots, y'_n(x)) dx \quad (5.1)$$

soit minimale parmi toutes les fonctions de classe \mathcal{C}^1 satisfaisant $y_i(a) = A_i$ et $y_i(b) = B_i$ pour tout $i = 1, \dots, n$. Ce problème correspond à celui du paragraphe II.3 où $y : [a, b] \rightarrow \mathbb{R}^n$. En considérant la fonction $S(\varepsilon) := \int_a^b L(x, y(x) + \varepsilon h(x), y'(x) + \varepsilon h'(x)) dx$, nous obtenons comme dans le Lemme 3.1 la condition nécessaire

$$\int_a^b \sum_{i=1}^n \left(\frac{\partial L}{\partial y_i}(x, y(x), y'(x)) h_i(x) + \frac{\partial L}{\partial y'_i}(x, y(x), y'(x)) h'_i(x) \right) dx = 0 \quad (5.2)$$

pour tout $h_i \in \mathcal{C}^1([a, b])$ satisfaisant $h_i(a) = h_i(b) = 0$. Si l'on pose toutes les fonctions $h_j(x)$ égales à zéro sauf la i ème, on voit que la condition (5.2) est équivalente à

$$\int_a^b \left(\frac{\partial L}{\partial y_i}(x, y(x), y'(x)) h_i(x) + \frac{\partial L}{\partial y'_i}(x, y(x), y'(x)) h'_i(x) \right) dx = 0$$

pour $i = 1, \dots, n$. On est donc de retour à la situation à une fonction, et comme dans le paragraphe II.3 on obtient le résultat suivant.

Théorème 5.1 Soit $L(x, y_1, \dots, y_n, y'_1, \dots, y'_n)$ de classe \mathcal{C}^1 . Une fonction $y : [a, b] \rightarrow \mathbb{R}^n$ de classe \mathcal{C}^1 est une extrémale du problème (5.1) aux extrémités fixées $y_i(a) = A_i$ et $y_i(b) = B_i$, si et seulement si pour tout $i \in \{1, \dots, n\}$

$$\begin{aligned} \frac{\partial L}{\partial y'_i}(x, y(x), y'(x)) & \text{ est continûment différentiable et} \\ \frac{\partial L}{\partial y_i}(x, y(x), y'(x)) - \frac{d}{dx} \frac{\partial L}{\partial y'_i}(x, y(x), y'(x)) & = 0. \end{aligned} \quad (5.3)$$

Revenons au problème mentionné au début de ce paragraphe et considérons une *surface* \mathcal{M} dans \mathbb{R}^3 , donnée par une paramétrisation $f : U \rightarrow \mathbb{R}^3$ où $U \subset \mathbb{R}^2$ est un ouvert. On suppose que f soit suffisamment différentiable et que $f'(u)$ soit de rang 2 pour tout $u \in U$. Des exemples sont

$$f(\varphi, \theta) = (\cos \varphi \sin \theta, \sin \varphi \sin \theta, \cos \theta)^T \quad (\text{sphère de rayon 1}) \quad (5.4)$$

$$f(\varphi, z) = (\cos \varphi, \sin \varphi, z)^T \quad (\text{cylindre}) \quad (5.5)$$

$$f(\varphi, z) = ((1-z) \cos \varphi, (1-z) \sin \varphi, z)^T \quad (\text{cône}). \quad (5.6)$$

Ils sont dessinés dans Fig. 5.1.

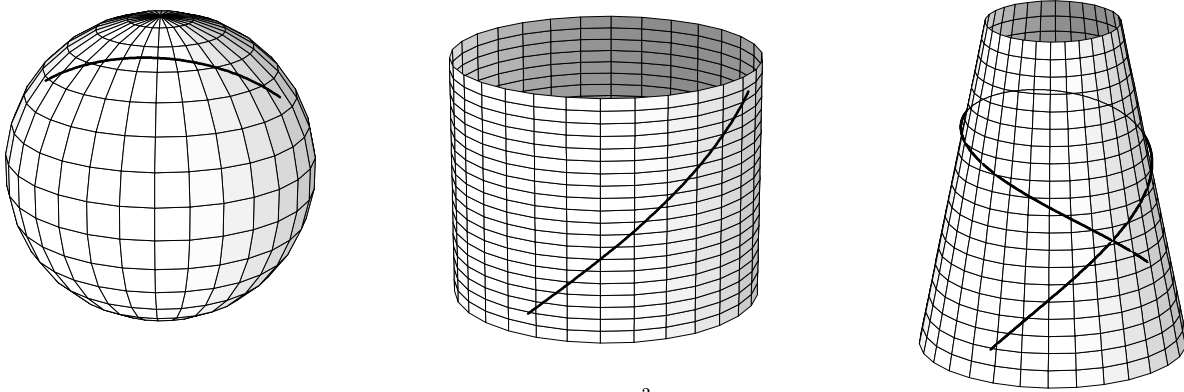


FIG. 5.1 – Surfaces dans \mathbb{R}^3 avec des géodésiques

Si $\beta : [a, b] \rightarrow U$ est une fonction différentiable avec $\beta'(t) \neq 0$ pour tout $t \in [a, b]$, alors $\gamma(t) := f(\beta(t))$ représente une *courbe* sur la surface \mathcal{M} . La longueur de cette courbe est donnée par

$$\int ds = \int_a^b \|\gamma'(t)\| dt = \int_a^b \|(f \circ \beta)'(t)\| dt, \quad (5.7)$$

car $ds \approx \|\gamma(t+dt) - \gamma(t)\| \approx \|\gamma'(t)\| dt$. Comme $(f \circ \beta)'(t) = f'(\beta(t))\beta'(t)$, nous avons $\|(f \circ \beta)'(t)\|^2 = \beta'(t)^T f'(\beta(t))^T f'(\beta(t))\beta'(t)$, et l'intégrale (5.7) devient

$$\int_a^b \sqrt{L(\beta(t), \beta'(t))} dt \quad \text{avec} \quad L(\beta, \beta') = \sum_{i,j=1}^2 g_{ij}(\beta) \beta'_i \beta'_j, \quad (5.8)$$

où $\beta_1(t), \beta_2(t)$ sont les deux composantes de $\beta(t)$, et les $g_{ij}(u)$ sont les éléments de la matrice $G(u) := f'(u)^T f'(u)$, c.-à-d.,

$$g_{ij}(u) = \sum_{k=1}^3 \frac{\partial f_k}{\partial u_i}(u) \cdot \frac{\partial f_k}{\partial u_j}(u). \quad (5.9)$$

Problème Pour une surface \mathcal{M} donnée, calculer les courbes sur \mathcal{M} ayant des extrémités fixées et une longueur minimale. Il s'agit alors de minimiser l'intégrale (5.8). Les extrémales de ce problème variationnel s'appellent *géodésiques* de la surface.

Ce problème est un cas particulier de (5.1) et nous pourrions utiliser le Théorème 5.1 pour calculer les extrémales du problème. Il y a néanmoins une astuce qui simplifie énormément le calcul. Tout d'abord, nous remarquons qu'une reparamétrisation de la courbe $\gamma(t)$ ne

change pas sa longueur. On peut donc se restreindre à des paramétrisations particulières. Une possibilité est d'utiliser la transformation $t \leftrightarrow s$ définie par $s(t) = \int_a^t \|\gamma'(\tau)\| d\tau$. Elle est bijective et continûment différentiable. Pour la nouvelle variable, la courbe $\gamma(s) := \gamma(t(s))$ satisfait $\|\gamma'(s)\| = \|\gamma'(t(s))t'(s)\| = 1$, car $s'(t) = \|\gamma'(t)\|$ et donc $t'(s) = 1/\|\gamma'(t(s))\|$. On dit alors que $\gamma(s)$ est paramétrée par la longueur d'arc.

Théorème 5.2 a) Supposons que la courbe $\gamma(s) = f(\beta(s))$ soit de classe \mathcal{C}^1 et paramétrée avec un paramètre naturel (c.-à-d., $\|\gamma'(s)\| = \text{Const} \neq 0$). Alors, γ est une géodésique de \mathcal{M} , si et seulement si $\beta(s)$ est une extrémale de

$$\int_a^b L(\beta(s), \beta'(s)) ds, \quad (5.10)$$

où $L(\beta, \beta')$ est donné par (5.8).

b) Toutes les extrémales de (5.10) sont paramétrées avec un paramètre naturel.

Démonstration. a) Si $\gamma(s) = f(\beta(s))$ est paramétrée avec un paramètre naturel, on a $L(\beta(s), \beta'(s)) = \|(f \circ \beta)'(s)\|^2 = \|\gamma'(s)\|^2 = \text{Const}$ et, par conséquent,

$$\begin{aligned} \frac{\partial \sqrt{L}}{\partial \beta_i}(\beta(s), \beta'(s)) - \frac{d}{ds} \left(\frac{\partial \sqrt{L}}{\partial \beta'_i}(\beta(s), \beta'(s)) \right) &= \frac{\frac{\partial L}{\partial \beta_i}(\beta(s), \beta'(s))}{2\sqrt{L(\beta(s), \beta'(s))}} - \frac{d}{ds} \left(\frac{\frac{\partial L}{\partial \beta'_i}(\beta(s), \beta'(s))}{2\sqrt{L(\beta(s), \beta'(s))}} \right) \\ &= \frac{1}{2\sqrt{\text{Const}}} \left(\frac{\partial L}{\partial \beta_i}(\beta(s), \beta'(s)) - \frac{d}{ds} \left(\frac{\partial L}{\partial \beta'_i}(\beta(s), \beta'(s)) \right) \right). \end{aligned}$$

Ceci démontre la partie (a).

b) La fonction $L(\beta, \beta')$ est homogène de degré 2 en β' , c.-à-d., $L(\beta, \sigma\beta') = \sigma^2 L(\beta, \beta')$. La dérivée par rapport à σ (pour $\sigma = 1$) donne l'identité d'Euler $\sum_{k=1}^2 \frac{\partial L}{\partial \beta'_k}(\beta, \beta')\beta'_k = 2L(\beta, \beta')$. On a donc

$$\begin{aligned} \frac{dL(\beta(s), \beta'(s))}{ds} &= \frac{d}{ds} \left(\sum_{k=1}^2 \frac{\partial L}{\partial \beta'_k}(\beta(s), \beta'(s))\beta'_k(s) - L(\beta(s), \beta'(s)) \right) \\ &= \sum_{k=1}^2 \left(\frac{d}{ds} \left(\frac{\partial L}{\partial \beta'_k}(\beta(s), \beta'(s)) \right) \beta'_k(s) + \frac{\partial L}{\partial \beta'_k}(\beta(s), \beta'(s))\beta''_k(s) \right) \\ &\quad - \sum_{k=1}^2 \frac{\partial L}{\partial \beta_k}(\beta(s), \beta'(s))\beta'_k(s) - \sum_{k=1}^2 \frac{\partial L}{\partial \beta'_k}(\beta(s), \beta'(s))\beta''_k(s) \\ &= - \sum_{k=1}^2 \left(\frac{\partial L}{\partial \beta_k}(\beta(s), \beta'(s)) - \frac{d}{ds} \left(\frac{\partial L}{\partial \beta'_k}(\beta(s), \beta'(s)) \right) \right) \beta'_k(s) = 0, \end{aligned}$$

et $L(\beta(s), \beta'(s)) = \|\gamma'(s)\|^2$ ne dépend pas de s pour une extrémale $\beta(s)$ de (5.10). \square

Exemple 5.3 (Géodésiques de la sphère) Pour la fonction $f(\varphi, \theta)$ de (5.4), on a

$$f'(\varphi, \theta) = \begin{pmatrix} -\sin \varphi \sin \theta & \cos \varphi \cos \theta \\ \cos \varphi \sin \theta & \sin \varphi \cos \theta \\ 0 & -\sin \theta \end{pmatrix}$$

et donc $g_{11}(\varphi, \theta) = \sin^2 \theta$, $g_{12}(\varphi, \theta) = g_{21}(\varphi, \theta) = 0$ et $g_{22}(\varphi, \theta) = 1$. La fonction L de (5.8) devient

$$L(\varphi, \theta, \varphi', \theta') = \sin^2 \theta \cdot \varphi'^2 + \theta'^2,$$

et les équations d'Euler-Lagrange sont

$$\frac{d}{ds}(2 \sin^2 \theta \cdot \varphi') = 0, \quad 2 \sin \theta \cos \theta \cdot \varphi'^2 - \frac{d}{ds}(2\theta') = 0. \quad (5.11)$$

La première équation implique $\sin^2 \theta(s) \cdot \varphi'(s) = \text{Const}$ et on obtient

$$\varphi(b) = \varphi(a) + \int_a^b \frac{\text{Const}}{\sin^2 \theta(s)} ds.$$

En supposant que $\varphi(a) = \varphi(b)$ mais $\theta(a) \neq \theta(b)$ (après une rotation de la sphère ceci est toujours possible), cette relation implique que la constante Const doit être nulle, d'où $\varphi'(s) = 0$ et $\varphi(s) = \varphi(a)$ pour tout s . De la deuxième équation de (5.11) on déduit alors que $\theta(s) = c_1 + c_2 s$. Les géodésiques de la sphère sont donc les cercles de rayon 1 avec pour centre l'origine de la sphère (grands cercles).

Exemple 5.4 (Géodésiques du cylindre) Pour la paramétrisation (5.5) on obtient le Lagrangien $L(\varphi, z, \varphi', z') = \varphi'^2 + z'^2$ et les équations d'Euler-Lagrange impliquent $\varphi''(s) = 0$ et $z''(s) = 0$. Les géodésiques du cylindre sont les images de droites sous l'application (5.5). Autrement dit, si l'on déroule le cylindre, les géodésiques deviennent des droites.

Considérons encore des surfaces de révolution $f(\varphi, z) = (\rho(z) \cos \varphi, \rho(z) \sin \varphi, z)^T$. La dérivée de f est

$$f'(\varphi, z) = \begin{pmatrix} -\rho(z) \sin \varphi & \rho'(z) \cos \varphi \\ \rho(z) \cos \varphi & \rho'(z) \sin \varphi \\ 0 & 1 \end{pmatrix},$$

et pour le Lagrangien (5.8) nous trouvons

$$L(\varphi, z, \varphi', z') = \rho(z)^2 \varphi'^2 + (1 + \rho'(z)^2) z'^2. \quad (5.12)$$

Théorème 5.5 (Clairaut) Pour une géodésique $(\varphi(s), z(s))$ sur une surface de révolution $f(\varphi, z) = (\rho(z) \cos \varphi, \rho(z) \sin \varphi, z)^T$ on a

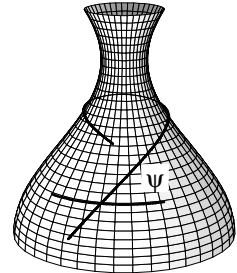
$$\rho(z(s)) \cdot \cos \psi(s) = \text{Const},$$

où $\psi(s)$ est l'angle au point $\gamma(s)$ entre $\gamma'(s)$ (direction de la courbe) et le cercle horizontal de la surface.

Démonstration. L'équation d'Euler-Lagrange pour L de (5.12) donne $\frac{d}{ds}(2\rho(z(s))^2 \varphi'(s)) = 0$, d'où $\rho(z(s))^2 \varphi'(s) = \text{Const}$. On en déduit que

$$\cos \psi(s) = \frac{\langle \gamma'(s), (-\sin \varphi(s), \cos \varphi(s), 0)^T \rangle}{\|\gamma'(s)\|} = \frac{\rho(z(s)) \varphi'(s)}{C} = \frac{\text{Const}}{\rho(z(s))},$$

car $\gamma'(s) = f'(\varphi(s), z(s))(\varphi'(s), z'(s))^T$ et $\|\gamma'(s)\| = C$ par le Théorème 5.2, partie (b). \square



II.6 Exercices

1. Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}$ deux fois différentiable en $a \in \mathbb{R}^n$. Démontrer que la condition du Théorème 1.4

$$f''(a)(h, h) \geq \gamma \|h\|^2 \quad \text{pour tout } h \in \mathbb{R}^n,$$

avec $\gamma > 0$ est équivalente à $f''(a)(h, h) > 0$ pour tout $h \neq 0$.

2. Soit $E := \{f : [1, \infty) \rightarrow \mathbb{R}; f \text{ continue et bornée}\}$ avec pour norme $\|f\|_\infty = \sup_{t \in [1, \infty)} |f(t)|$. On considère l'application $F : E \rightarrow \mathbb{R}$ donnée par

$$F(f) = \int_1^\infty \frac{f^2(t)}{t^4} dt - \int_1^\infty \frac{f^3(t)}{t^2} dt.$$

Démontrer que

- (a) $F'(0) = 0$ et $F''(0)(h, h) > 0$ pour $h \neq 0$.
 (b) $0 \in E$ n'est pas un minimum relatif de F .
3. Montrer que la valeur maximale de l'expression

$$\frac{ax^2 + 2bxy + cy^2}{ex^2 + 2fxy + gy^2}$$

est la plus grande racine de $(eg - f^2)\lambda^2 - (ag - 2bf + ec)\lambda + (ac - b^2) = 0$.

4. Considérons la fonction

$$f(x_1, x_2, x_3) = \alpha x_1^2 e^{x_2} + x_2^2 e^{x_3} + x_3^2 e^{x_1}.$$

Démontrer que l'origine est un point critique. Pour quelle valeur de $\alpha \in \mathbb{R}$, l'origine est-elle un extremum de f ?

5. Soit $m < n$ et A une matrice $m \times n$ de rang m . Parmi toutes les solutions de $Ax = b$ trouver à l'aide de multiplicateurs de Lagrange la solution pour laquelle $\|x\|_2$ est minimale. Donner une interprétation géométrique du résultat.
6. Soit $E = \mathcal{C}([0, 1])$ et soit $F : E \rightarrow \mathbb{R}$ donnée par $F(y) = 2y(0)^3 - 3y(0)^2$. Démontrer les affirmations suivantes:

- (a) $y_0(x) \equiv 1$ est un minimum relatif de F , si l'on munit E de la norme $\|y\|_\infty = \max_{x \in [0, 1]} |y(x)|$.
 (b) $y_0(x) \equiv 1$ n'est pas un minimum relatif de F , si l'on munit E de la norme $\|y\|_1 = \int_0^1 dx |y(x)|$.

7. Calculer géométriquement le minimum de $f|_{\mathcal{M}}$ où $f(x_1, x_2) = x_1^2 + x_2^2$, et

$$\mathcal{M} = \{(x_1, x_2); g(x_1, x_2) = 0\} \quad \text{avec} \quad g(x_1, x_2) = x_2^2 - (x_1 - 1)^3.$$

Montrer que le système

$$f'(x) - \lambda g'(x) = 0$$

ne possède pas de solution. Expliquer pourquoi.

8. Calculer la valeur maximale de

$$x^a y^c z^c \quad (a, b, c, x, y, z > 0)$$

sujette à la condition $x^k + y^k + z^k = 1$, ($k > 0$). En déduire l'inégalité

$$\left(\frac{u}{a}\right)^a \left(\frac{v}{b}\right)^b \left(\frac{w}{c}\right)^c \leq \left(\frac{u+v+w}{a+b+c}\right)^{a+b+c}.$$

9. (Condition nécessaire). Soit E un espace vectoriel normé, et $U \subset E$ un ouvert. Soient $f : U \rightarrow \mathbb{R}$, $g : U \rightarrow \mathbb{R}^m$ deux fois différentiables en $a \in \mathcal{M}$, où $\mathcal{M} = \{x \in U; g(x) = 0\}$ et soit $g'(a)$ surjective. Montrer l'affirmation suivante: Si $f|_{\mathcal{M}}$ possède un minimum relatif en a , alors il existe $\lambda_1, \dots, \lambda_m$ tels que

$$f'(a) - \sum_{i=1}^m \lambda_i g'_i(a) = 0 \quad \text{et} \quad \left(f''(a) - \sum_{i=1}^m \lambda_i g''_i(a)\right)(h, h) \geq 0$$

pour tout les h tels que $g'(a)h = 0$.

Indication. Appliquer le Théorème 1.3 à la situation de la démonstration du Théorème 2.2.

10. Calculer les extrema relatifs de $f(x, y, z) = x^2 + y^2 + z^2$ sur \mathcal{M} où

$$\mathcal{M} = \left\{ (x, y, z); \frac{x^2}{4} + \frac{y^2}{9} + \frac{z^2}{25} = 1 \text{ et } z = x + y \right\}.$$

Vérifier les conditions suffisantes pour un extremum.

11. Calculer l'aire de révolution pour la fonction

$$y(x) = \frac{1}{\cosh \lambda} \cosh(x \cosh \lambda + \lambda)$$

et pour $0 \leq x \leq b$. Avec $b = 1.148$ comparer les valeurs obtenues pour $\lambda = -2.1$ et $\lambda = 0$ et celle de la solution de Goldschmidt (voir la Fig. 3.1).

Résultats numériques. 13.22, 11.33 et 12.55.

12. Considérer le problème de la brachystochrone (Exemple 3.6) avec les valeurs $b = 1$ et $B = 1$ (et utiliser la normalisation $2g = 1$).

(a) Calculer le temps $T_1(y)$ pour la droite $y(x) = x$.

(b) Démontrer que pour le cercle $y(x) = \sqrt{1 - (x - 1)^2}$ on obtient

$$T_2(y) = \int_0^{\pi/2} (\sin \theta)^{-1/2} d\theta,$$

(la solution proposée par Galilée). Démontrer que $T_2(y) < T_1(y)$ (utiliser une table d'intégrales si nécessaire).

(c) Calculer le temps $T_3(y)$ pour la cycloïde (en utilisant les valeurs $C = 1.1458$ et $\varphi_{\text{end}} = 2.4120$).

Résultats numériques. 2.8284, 2.6221, 2.5819.

13. Pour le problème de minimisation de

$$T(y) = \int_a^b (2xy + (x^2 + 3y^2)y') dx$$

montrer que l'équation d'Euler-Lagrange s'annule identiquement. Trouver une fonction $f(x, y)$ telle que $L(x, y, y')dx = df(x, y)$ et montrer que l'intégrale $T(y)$ ne dépend que de $y(a)$ et $y(b)$, mais pas de la forme de $y(x)$.

14. Déterminer les extrémales des problèmes suivants

$$T_1(y) = \int_a^b (2xy(x) - y^2(x) + 3y^2(x)y'(x)) dx$$

$$T_2(y) = \int_a^b x \sqrt{1 + y'(x)^2} dx$$

15. Déterminer les extrémales des problèmes suivants

$$T(y) = \int_0^1 (y'(x)^2 + 12xy(x)) dx \quad \text{avec } y(0) = 2, y(1) = 3,$$

$$T(y) = \int_1^2 y'(x)(1 + x^2 y'(x)) dx \quad \text{avec } y(1) = 3, y(2) = 2.$$

16. Trouver $\varphi : [a, b] \rightarrow \mathbb{R}$, $\varphi > 0$ avec $\varphi(a) = c$, $\varphi(b) = d$ et $c, d > 0$ qui minimise

$$\int_a^b \frac{\sqrt{1 + (\varphi'(x))^2}}{\varphi(x)} dx.$$

Remarque. Cette intégrale représente la longueur hyperbolique d'une courbe Γ dans le demi-plan de Poincaré $\mathcal{H} = \{z \in \mathbb{C}; \Im(z) > 0\}$ paramétrisée par $(t, \varphi(t))$.

17. Déterminer la fonction $\varphi \mapsto \rho(\varphi)$ extrémale du problème de minimisation de

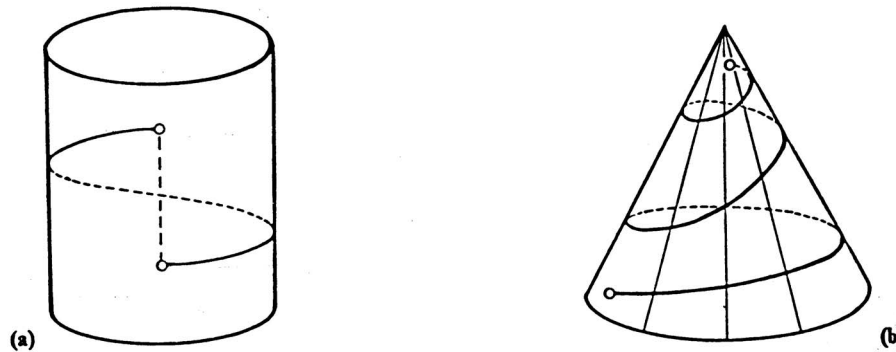
$$\int_{\alpha}^{\beta} \sqrt{\rho^2 + \left(\frac{d\rho}{d\varphi}\right)^2} d\varphi$$

avec $\rho(\alpha) = r$ et $\rho(\beta) = R$. Interpréter géométriquement le problème et sa solution. Même question pour la fonction $\rho \mapsto \varphi(\rho)$ et

$$\int_r^R \sqrt{1 + \rho^2 \left(\frac{d\varphi}{d\rho}\right)^2} d\rho$$

et les conditions $\varphi(r) = \alpha$ et $\varphi(R) = \beta$.

18. La figure ci-dessous a été copiée d'un livre publié en 1996. Les courbes dessinées peuvent-elle être des géodésiques?

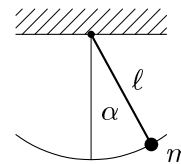


19. Le principe de Hamilton dit que

$$\int L dt \rightarrow \min$$

où $L = T - U$, T est l'énergie cinétique et U est l'énergie potentielle.

(a) Pour le pendule simple, considérons l'angle α comme la coordonnée généralisée. L'énergie cinétique est $T = \frac{m}{2}(\dot{x}^2 + \dot{y}^2) = \frac{m}{2}\ell^2\dot{\alpha}^2$ et l'énergie potentielle est $U = mgy = -mgl \cos \alpha$ (car $x = \ell \sin \alpha$ et $y = -\ell \cos \alpha$). Déterminer les équations d'Euler-Lagrange.



(b) Déterminer les équations du mouvement du pendule sphérique de longueur l . En utilisant les coordonnées sphériques

$$x = l \cos \varphi \sin \theta, \quad y = l \sin \varphi \sin \theta, \quad z = l \cos \theta,$$

le résultat est donné par

$$\begin{aligned} \theta'' &= \frac{g}{l} \sin \theta + \sin \theta \cos \theta (\varphi')^2 \\ \sin \theta \varphi'' &= -2 \cos \theta \theta' \varphi' \end{aligned} \tag{6.1}$$

Indication. $T = m(x'^2 + y'^2 + z'^2)/2$, $U = mgz$.

20. On considère le problème

$$\int_a^b L(x, y(x), y'(x), y''(x)) dx \rightarrow \min,$$

où $y(a) = A_0$, $y'(a) = A_1$, $y(b) = B_0$ and $y'(b) = B_1$. Montrer que si $L(x, y, y', y'')$ et $y(x)$ sont suffisamment différentiables, alors une condition suffisante pour un extremum est

$$\frac{\partial L}{\partial y}(x, y(x), y'(x), y''(x)) - \frac{d}{dx} \frac{\partial L}{\partial y'}(x, y(x), y'(x), y''(x)) + \frac{d^2}{dx^2} \frac{\partial L}{\partial y''}(x, y(x), y'(x), y''(x)) = 0. \tag{6.2}$$

L'équation (6.2) s'appelle équation d'Euler-Poisson.

21. Résoudre le problème

$$\int_{-l}^l \left(\frac{\eta}{2} y''(x)^2 + \rho y(x) \right) dx \rightarrow \min .$$

avec $y(-l) = 0$, $y'(-l) = 0$, $y(l) = 0$ and $y'(l) = 0$ et η et ρ sont des constantes.

Chapitre III

Equations différentielles ordinaires

Dans de nombreuses applications, la dynamique d'un système peut être modélisée par des équations différentielles ordinaires. Par exemple, en mécanique (en utilisant la loi de Newton), en astronomie (le mouvement des planètes ou l'étude des galaxies), en dynamique moléculaire (le comportement des atomes), en chimie (des réactions chimiques), en électronique (circuits intégrés), en biologie (le développement de populations), etc. Au Chapitre II sur le calcul variationnel nous étions aussi confrontés aux équations différentielles ordinaires (les équations d'Euler-Lagrange).

Dans quelques cas très rares on arrive à trouver la solution sous forme analytique (voir [HW95, II.7, II.8]). Mais, en général, on est obligé d'utiliser des méthodes numériques pour obtenir les solutions (voir le cours "Analyse Numérique"). Dans ce chapitre, nous allons traiter surtout des propriétés théoriques de solutions des équations différentielles ordinaires. En particulier, nous étudierons l'existence et l'unicité de solutions, leur sensibilité par rapport à des perturbations et leur comportement sur des longs intervalles (stabilité).

III.1 Terminologie et quelques exemples

Une *équation différentielle ordinaire* est donnée par une relation de la forme

$$F(x, y, y', y'', \dots, y^{(m)}) = 0, \quad (1.1)$$

où $x \in \mathbb{R}$ et $y, y', \dots, y^{(m)} \in \mathbb{R}^n$. La fonction F est définie dans un ouvert de $\mathbb{R} \times \mathbb{R}^{n(m+1)}$. On dit que l'équation différentielle est de *l'ordre* m , si m est la dérivée maximale qui apparaît dans l'équation. Une fonction $y : I \rightarrow \mathbb{R}^n$ (où I est un intervalle dans \mathbb{R}) est une *solution* de (1.1), si elle est de classe \mathcal{C}^m et si

$$F(x, y(x), y'(x), y''(x), \dots, y^{(m)}(x)) = 0 \quad \text{pour tout } x \in I.$$

Exemple 1.1 (Clairaut 1734) Considérons l'équation différentielle

$$y - xy' + f(y') = 0. \quad (1.2)$$

Elle est donnée sous forme *implicite*, c.-à-d., par une équation nonlinéaire en y' . On vérifie sans problème que les droites $y(x) = Cx - f(C)$ sont des solutions (voir Fig. 1.1 pour les cas où $f(t) = 5(t^3 - t)/2$ (gauche) et $f(t) = t^2 + t$ (droite)). L'enveloppe de cette famille de droites est également une solution (voir Exercice 4). Si l'on fixe un point (x_0, y_0) dans \mathbb{R}^2 , on voit que dans un certain domaine il y a plusieurs solutions qui passent par le même

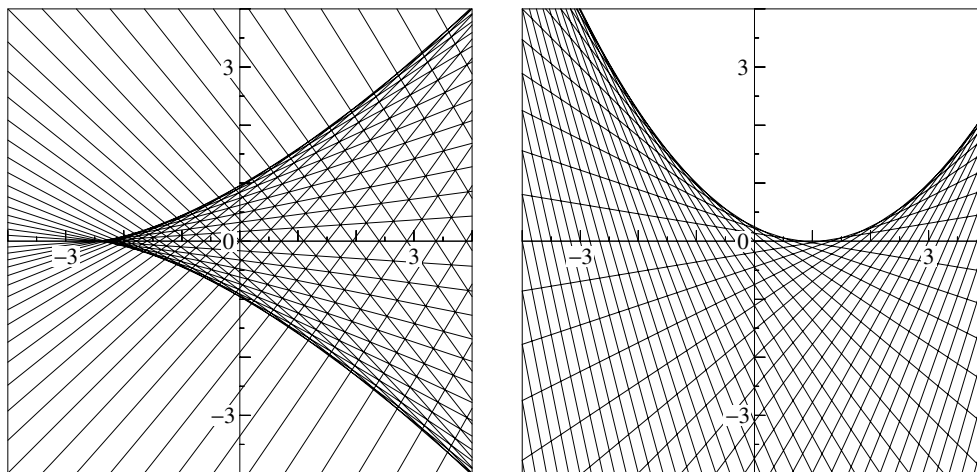


FIG. 1.1 – Solutions de l'équation de Clairaut (Exemple 1.1)¹

(x_0, y_0) , alors que dans une autre partie il y a une seule solution ou pas de solution du tout. L'explication de ce fait est que, pour (x, y) fixé, l'équation (1.2) admet trois, deux, une ou pas de solution pour y' .

Pour une équation différentielle de la forme (1.1) il est en général très difficile d'obtenir des résultats sur l'existence ou l'unicité de solutions (comme on a vu dans l'Exemple 1.1). Nous allons donc nous restreindre à la situation où l'équation (1.1) peut être résolue par rapport à la dérivée maximale $y^{(m)}$. Nous écrivons alors

$$y^{(m)} = g(x, y, y', y'', \dots, y^{(m-1)}). \quad (1.3)$$

Une telle équation différentielle est dite *explicite*. Si l'on introduit des nouvelles variables pour les dérivées dans l'expression de droite de (1.3), par exemple $y_1 := y, y_2 := y', \dots, y_m := y^{(m-1)}$, nous obtenons le système équivalent

$$\begin{aligned} y_1' &= y_2 \\ &\dots \\ y_{m-1}' &= y_m \\ y_m' &= g(x, y_1, \dots, y_m) \end{aligned} \quad (1.4)$$

qui représente une équation différentielle d'ordre 1 pour le super-vecteur $(y_1, \dots, y_m)^T$ (observons que les y_i sont en général aussi des vecteurs). En notant ce super-vecteur de nouveau par y et l'expression de droite par $f(x, y)$, le système (1.4) devient simplement $y' = f(x, y)$. Pour étudier des équations différentielles explicites il suffit alors de considérer le cas $m = 1$ dans (1.3).

Exemple 1.2 (prédateur et proie) L'expansion de deux populations (par exemple, $y(t)$ pour le nombre de lièvres et $z(t)$ pour le nombre de lynx) peut être modélisée par les équations de Volterra-Lotka

$$y' = y(\alpha - \beta z), \quad z' = z(\gamma y - \delta), \quad (1.5)$$

où $\alpha, \beta, \gamma, \delta$ sont des constantes données. Ceci implique que la population y croît si z est plus petit qu'une certaine valeur de seuil (α/β) alors que sinon elle décroît. Pour l'autre population, c'est le contraire. Pour étudier les solutions, divisons les deux équations et considérons

1. Le dessin de gauche de la Fig. 1.1 a été pris du livre [HNW93].

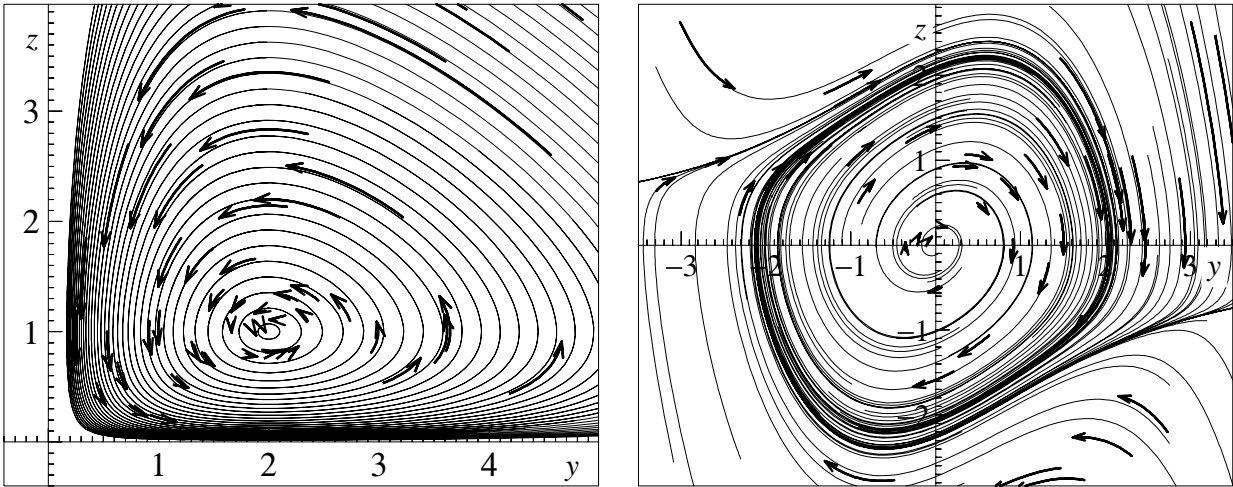


FIG. 1.2 – Equations de Volterra-Lotka et de van der Pol

y comme fonction de z (ou z comme fonction de y). On obtient ainsi (par séparation de variables)

$$\frac{dy}{dz} = \frac{y(\alpha - \beta z)}{z(\gamma y - \delta)} \quad \text{ou} \quad \frac{(\gamma y - \delta)}{y} dy = \frac{(\alpha - \beta z)}{z} dz.$$

Une intégration de la dernière équation donne

$$\gamma y - \delta \ln y = \alpha \ln z - \beta z + \text{Const.} \quad (1.6)$$

Dans la Fig. 1.2 (gauche) sont dessinées des courbes de niveaux de la fonction (1.6) pour $\alpha = 1, \beta = 1, \gamma = 1, \delta = 2$. Chaque solution $(y(t), z(t))$ reste sur une courbe de niveau pour tout $t \in \mathbb{R}$. Ceci suggère que les solutions de (1.5) sont périodiques.

Exemple 1.3 (van der Pol) L'équation van der Pol

$$y'' = \varepsilon(1 - y^2)y' - y, \quad \varepsilon = 0.4 \quad (1.7)$$

est une perturbation de l'oscillateur harmonique $y'' + y = 0$, qui possède comme solution $y(t) = A \cos(t - \varphi)$, $z(t) = y'(t) = -A \sin(t - \varphi)$ (cercles). Les solutions de (1.7) dans le plan (y, z) sont dessinées dans la Fig. 1.2 (droite). On observe que toutes les solutions s'approchent d'une solution périodique qui est une déformation du cercle de rayon 2.

Dans les deux derniers exemples on constate que les courbes de solutions dans l'espace (y, z) ne s'intersectent jamais. Ceci signifie qu'il n'y a pas plusieurs solutions différentes passant par le même point. Notre premier but dans ce chapitre est d'étudier des conditions suffisantes pour qu'un problème

$$y' = f(x, y), \quad y(x_0) = y_0 \quad (1.8)$$

possède une solution unique. On appelle x_0 et y_0 *valeurs initiales*, et le problème (1.8) avec valeurs initiales données s'appelle un *problème de Cauchy* ou un *problème aux valeurs initiales*.

III.2 Existence et unicité du problème de Cauchy

Considérons le problème de Cauchy

$$y' = f(x, y), \quad y(x_0) = y_0 \quad (2.1)$$

où $f : U \rightarrow \mathbb{R}^n$ (avec $U \subset \mathbb{R} \times \mathbb{R}^n$ ouvert) est une fonction continue. En intégrant l'équation différentielle entre x_0 et x on obtient l'équation intégrale

$$y(x) = y_0 + \int_{x_0}^x f(t, y(t)) dt. \quad (2.2)$$

Chaque solution de (2.1) est donc solution de (2.2). Le contraire est vrai aussi. Si une fonction continue $y(x)$ vérifie (2.2) sur un intervalle I , alors elle est automatiquement de classe \mathcal{C}^1 et elle vérifie (2.1).

Itération de Picard-Lindelöf L'équation (2.2) peut être considérée comme un problème à point fixe dans $\mathcal{C}(I)$. L'idée est d'appliquer la méthode des approximations successives (voir I.6). Elle s'écrit sous la forme

$$\begin{aligned} y_0(x) &= y_0 \quad (\text{ou une fonction arbitraire}) \\ y_{k+1}(x) &= y_0 + \int_{x_0}^x f(t, y_k(t)) dt. \end{aligned} \quad (2.3)$$

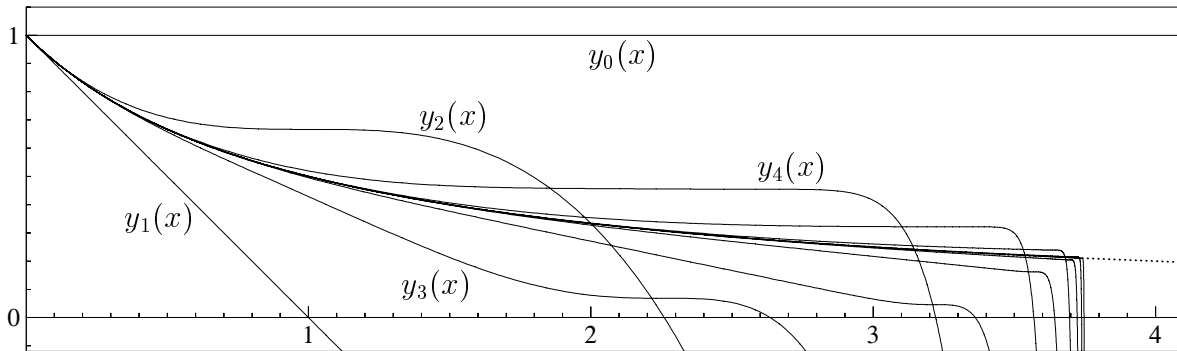


FIG. 2.1 – Itération de Picard-Lindelöf pour le problème de l'Exemple 2.1

Exemple 2.1 Considérons le problème

$$y' = -y^2, \quad y(0) = 1$$

avec comme solution exacte $y(x) = 1/(1+x)$. Les premières approximations obtenues par l'itération de Picard-Lindelöf sont $y_0(x) = 1$, $y_1(x) = 1 - x$ et $y_2(x) = 1 - x + x^2 - x^3/3$ (voir la Fig. 2.1). On observe une convergence rapide vers la solution exacte dans l'intervalle $[0, 3.75]$. Pour x trop grand, l'itération diverge.

Lemme 2.2 Soit $A = \{(x, y) \in \mathbb{R} \times \mathbb{R}^n; |x - x_0| \leq a, \|y - y_0\| \leq b\}$, $f : A \rightarrow \mathbb{R}^n$ une fonction continue et $M = \max_{(x,y) \in A} \|f(x, y)\|$. Pour $\alpha := \min(a, b/M)$, l'opérateur

$$(Ty)(x) := y_0 + \int_{x_0}^x f(t, y(t)) dt \quad (2.4)$$

est bien défini sur $B := \{y : [x_0 - \alpha, x_0 + \alpha] \rightarrow \mathbb{R}^n; y \text{ continue et } \|y(x) - y_0\| \leq b\}$ et il satisfait $T(B) \subset B$.

Démonstration. L'affirmation est une conséquence de

$$\|(Ty)(x) - y_0\| = \left\| \int_{x_0}^x f(t, y(t)) dt \right\| \leq \int_{x_0}^x \|f(t, y(t))\| dt \leq M|x - x_0| \leq M\alpha \leq b. \quad \square$$

On dit qu'une fonction $f : A \rightarrow \mathbb{R}^n$ (avec A comme dans le lemme précédent) satisfait une *condition de Lipschitz* si

$$\|f(x, y) - f(x, z)\| \leq L \|y - z\| \quad \text{pour } (x, y), (x, z) \in A. \quad (2.5)$$

La constante L s'appelle *constante de Lipschitz*. Remarquons que la condition (2.5) n'est pas une conséquence de la continuité de $f(x, y)$. Par exemple, la fonction $\sqrt{|y|}$ est continue sans vérifier (2.5). D'autre part, chaque fonction de classe \mathcal{C}^1 vérifie (2.5). Ceci est une conséquence du théorème des accroissements finis (voir I.5).

Si $f(x, y)$ satisfait une condition de Lipschitz et si $\alpha L < 1$, l'opérateur T de (2.4) est une contraction sur B (notation comme dans le Lemme 2.2). Dans cette situation on peut appliquer le Théorème du point fixe de Banach (voir I.6) pour conclure que $Ty = y$ possède une solution unique dans B . Nous allons démontrer l'existence et l'unicité d'une solution sans cette condition supplémentaire sur α .

Théorème 2.3 (existence et unicité du problème de Cauchy) *Considérons l'ensemble $A = \{(x, y) \in \mathbb{R} \times \mathbb{R}^n ; |x - x_0| \leq a, \|y - y_0\| \leq b\}$ et supposons que $f : A \rightarrow \mathbb{R}^n$*

- soit continue,
- satisfasse une condition de Lipschitz.

Alors, le problème de Cauchy $y' = f(x, y)$, $y(x_0) = y_0$ possède une solution unique sur $I = [x_0 - \alpha, x_0 + \alpha]$, où $\alpha = \min(a, b/M)$ et $M = \max_{(x,y) \in A} \|f(x, y)\|$.

Démonstration. Existence. L'idée est de démontrer que l'itération de Picard-Lindelöf converge uniformément sur I vers une solution du problème de Cauchy. Dans une première partie nous allons démontrer que

$$\|y_{k+1}(x) - y_k(x)\| \leq ML^k \frac{|x - x_0|^{k+1}}{(k+1)!} \quad \text{pour } |x - x_0| \leq \alpha. \quad (2.6)$$

Pour $k = 0$ avec pour $y_0(x) = y_0$, cette estimation suit de $\|\int_{x_0}^x f(t, y(t)) dt\| \leq M|x - x_0|$. Supposons qu'elle soit vraie pour $k - 1$. Alors, on a pour $x_0 \leq x \leq x_0 + \alpha$ que

$$\begin{aligned} \|y_{k+1}(x) - y_k(x)\| &\leq \int_{x_0}^x \|f(t, y_k(t)) - f(t, y_{k-1}(t))\| dt \leq L \int_{x_0}^x \|y_k(t) - y_{k-1}(t)\| dt \\ &\leq ML^k \int_{x_0}^x \frac{|t - x_0|^k}{k!} dt = ML^k \frac{|x - x_0|^{k+1}}{(k+1)!} \end{aligned}$$

(la démonstration pour $x_0 - \alpha \leq x \leq x_0$ est analogue).

De l'estimation (2.6) nous déduisons que $\{y_k(x)\}$ est une suite de Cauchy pour la norme $\|y\|_\infty = \max_{x \in I} \|y(x)\|$. En effet,

$$\begin{aligned} \|y_{k+m}(x) - y_k(x)\| &\leq \|y_{k+m}(x) - y_{k+m-1}(x)\| + \dots + \|y_{k+1}(x) - y_k(x)\| \\ &\leq \frac{M}{L} \left(\frac{(L|x - x_0|)^{k+m}}{(k+m)!} + \dots + \frac{(L|x - x_0|)^{k+1}}{(k+1)!} \right) \leq \frac{M}{L} \sum_{j \geq k+1} \frac{(L\alpha)^j}{j!}, \end{aligned}$$

ce qui est borné par le reste d'une série convergente. Donc, cette expression est plus petite que ε si k est suffisamment grand. Comme l'espace des fonctions continues avec la norme $\|y\|_\infty$ est complet, la suite $\{y_k(x)\}$ converge vers une fonction continue $y : I \rightarrow \mathbb{R}^n$.

Pour démontrer que cette fonction $y(x)$ est une solution du problème de Cauchy, nous passons à la limite $k \rightarrow \infty$ dans (2.3). Comme $\{y_k(x)\}$ converge uniformément et $f(x, y)$ est uniformément continue sur le compact A , la suite $\{f(x, y_k(x))\}$ converge uniformément vers $f(x, y(x))$. On peut donc échanger la limite avec l'intégration dans (2.3) et on voit que $y(x)$ est solution de l'équation intégrale (2.2).

Unicité. Supposons que $y(x)$ et $z(x)$ soient deux solutions sur I . Ceci implique que $y(x) - z(x) = \int_{x_0}^x (f(t, y(t)) - f(t, z(t))) dt$. On en déduit que $\|y(x) - z(x)\| \leq 2M|x - x_0|$. Comme dans la première partie de la démonstration nous trouvons par récurrence que pour tout $k \geq 0$

$$\|y(x) - z(x)\| \leq 2ML^k \frac{|x - x_0|^{k+1}}{(k+1)!} \leq \frac{2M}{L} \frac{(L\alpha)^{k+1}}{(k+1)!}.$$

La limite $k \rightarrow \infty$ montre que $\|y(x) - z(x)\| = 0$ pour tout $x \in I$. □

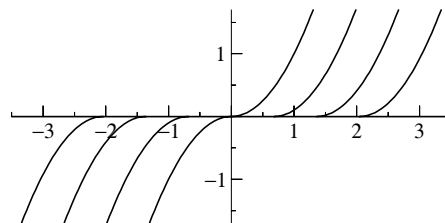
III.3 Théorème de Peano

Montrons d'abord par un exemple qu'on ne peut pas omettre l'hypothèse sur la condition de Lipschitz pour conclure l'existence et l'unicité d'une solution du problème de Cauchy.

Exemple 3.1 Considérons l'équation différentielle

$$y' = 2\sqrt{|y|}.$$

On vérifie que $y(x) \equiv 0$ ainsi que $y(x) = (x - c)^2$ pour $x \geq c$ et $y(x) = -(c - x)^2$ pour $x \leq c$ sont des solutions. Le problème de Cauchy avec pour valeur initiale $y(0) = 0$ possède donc une infinité de solutions (voir le dessin).



Observons que $f(x, y) = 2\sqrt{|y|}$ ne satisfait pas une condition de Lipschitz et que le Théorème 2.3 ne peut donc pas être appliqué.

Le but de ce paragraphe est de démontrer l'existence (mais pas l'unicité) d'une solution de (2.1) sans utiliser une condition de Lipschitz. Dans cette situation l'itération de Picard-Lindelöf n'est pas utile comme le démontre l'Exercice 10. Nous considérons donc une autre suite de fonctions approchant une solution du problème de Cauchy.

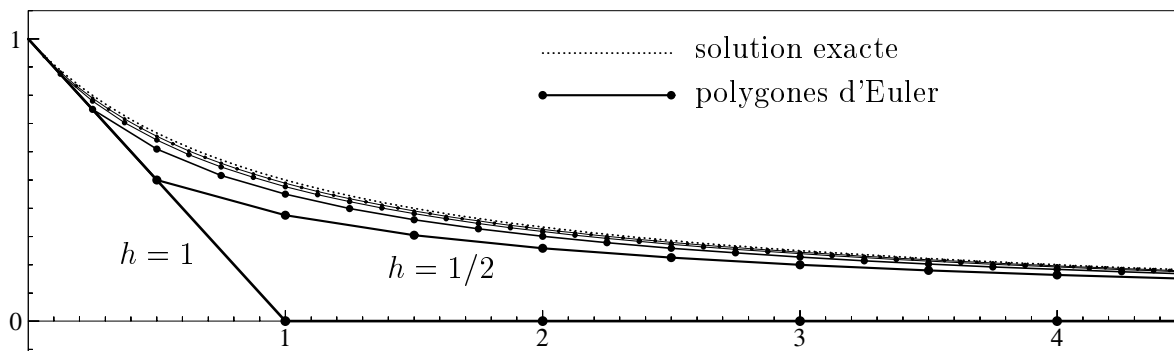


FIG. 3.1 – Polygones d'Euler pour $y' = -y^2$, $y(0) = 1$ avec solution exacte $y(x) = 1/(1+x)$

Polygones d'Euler (Euler 1768) L'idée est d'approximer la solution localement par sa tangente $y(x+h) \approx y(x) + hf(x, y(x))$. Pour un $h > 0$ donné, nous considérons alors la suite $\{x_n, y_n\}_{n \geq 0}$ donnée par

$$y_{n+1} = y_n + hf(x_n, y_n), \quad x_{n+1} = x_n + h,$$

et nous notons par $y_h(x)$ la fonction linéaire par morceaux qui passe par les points (x_n, y_n) . Ces polygones sont dessinés pour le problème $y' = -y^2$, $y(0) = 1$ et pour $h = 1, 1/2, \dots$ dans la Fig. 3.1.

Lemme 3.2 Soit $A = \{(x, y) \in \mathbb{R} \times \mathbb{R}^n; |x - x_0| \leq a, \|y - y_0\| \leq b\}$, $f : A \rightarrow \mathbb{R}^n$ une fonction continue, $M = \max_{(x,y) \in A} \|f(x, y)\|$ et $\alpha := \min(a, b/M)$. Pour $h = \alpha/N$ (avec $N \in \mathbb{N}$), le polygone d'Euler satisfait $(x, y_h(x)) \in A$ pour $x \in [x_0, x_0 + \alpha]$ et on a l'estimation

$$\|y_h(x) - y_h(\bar{x})\| \leq M |x - \bar{x}| \quad \text{pour } x, \bar{x} \in [x_0, x_0 + \alpha]. \quad (3.1)$$

Démonstration. Par récurrence on voit que $(x_n, y_n) \in A$ pour $n = 0, 1, \dots, N$. En effet, on a $\|y_{n+1} - y_n\| \leq hM$ et $\|y_{n+1} - y_0\| \leq (n+1)hM \leq \alpha M \leq b$ si $n+1 \leq N$. Ceci implique que $(x, y_h(x)) \in A$ pour tout $x \in [x_0, x_0 + \alpha]$. L'estimation (3.1) est une conséquence du fait que $y_h(x)$ est linéaire par morceaux et que la pente est partout bornée par M . \square

Définition 3.3 Une famille de fonctions $f_j : I \rightarrow \mathbb{R}^n$ (où I est un intervalle) s'appelle *équicontinue*, si pour tout $\varepsilon > 0$ il existe un $\delta > 0$ tel que

$$\forall j \quad \forall x, \bar{x} \quad (|x - \bar{x}| < \delta \implies \|f_j(x) - f_j(\bar{x})\| < \varepsilon).$$

Il est important que δ ne dépend que de ε et non de j .

Théorème 3.4 (Arzelà-Ascoli 1895) Soit $f_j : [a, b] \rightarrow \mathbb{R}^n$ une famille de fonctions satisfaisant

- $\{f_j(x)\}$ est équicontinue,
- pour tout $x \in [a, b]$ il existe un $M(x) \in \mathbb{R}$ tel que $\|f_j(x)\| \leq M(x)$ pour tout j .

Alors, la famille $\{f_j(x)\}$ possède une sous-suite $\{g_n(x)\}_{n \geq 1}$ qui converge uniformément vers une fonction continue $g : [a, b] \rightarrow \mathbb{R}^n$.

Démonstration. Construction de la sous-suite. L'ensemble $\mathcal{Q} \cap [a, b]$ est dénombrable et dense dans $[a, b]$. Il peut donc être écrit comme $\mathcal{Q} \cap [a, b] = \{x_1, x_2, x_3, \dots\}$. La famille $\{f_j(x_1)\}$ est bornée dans \mathbb{R}^n et, par le théorème de Bolzano-Weierstrass, possède une sous-suite convergente. Notons-la par $\{f_{1i}(x_1)\}_{i \geq 1}$. Donc

$$f_{11}(x), f_{12}(x), f_{13}(x), \dots \quad \text{converge pour } x_1.$$

Considérons maintenant la suite $\{f_{1i}(x_2)\}_{i \geq 1}$. De nouveau, par le théorème de Bolzano-Weierstrass, cette suite possède une sous-suite convergente et on trouve une suite telle que $f_{21}(x), f_{22}(x), f_{23}(x), \dots$ converge pour x_1 et pour x_2 . Après n étapes on trouve une sous-suite $\{f_{ni}(x)\}_{i \geq 1}$ de $\{f_j(x)\}$ telle que

$$f_{n1}(x), f_{n2}(x), f_{n3}(x), \dots \quad \text{converge pour } x_1, x_2, \dots, x_n.$$

L'idée est de considérer la "suite diagonale" $g_n(x) := f_{nn}(x)$. Cette suite converge pour tout $x_\ell \in \mathcal{Q} \cap [a, b]$, car $\{g_n(x_\ell)\}_{n \geq \ell}$ est une sous-suite de $\{f_{ln}(x_\ell)\}_{n \geq \ell}$, qui converge.

Convergence uniforme. Pour un $\varepsilon > 0$ donné, il existe un $\delta > 0$ tel que pour tout $n \geq 1$

$$|x - \bar{x}| < \delta \implies \|g_n(x) - g_n(\bar{x})\| < \varepsilon. \quad (3.2)$$

Choisissons un nombre fini de points $x_1, x_2, \dots, x_{q-1} \in \mathcal{Q} \cap [a, b]$ satisfaisant $a = x_0 < x_1 < \dots < x_{q-1} < x_q = b$ et $x_{i+1} - x_i < \delta$. Pour un $x \in [a, b]$, notons par ℓ l'indice tel que $x \in [x_\ell, x_{\ell+1}]$. Utilisons l'inégalité du triangle:

$$\|g_n(x) - g_m(x)\| \leq \|g_n(x) - g_n(x_\ell)\| + \|g_n(x_\ell) - g_m(x_\ell)\| + \|g_m(x_\ell) - g_m(x)\|.$$

Par (3.2) on a que $\|g_n(x) - g_n(x_\ell)\| + \|g_m(x_\ell) - g_m(x)\| < 2\varepsilon$. Comme $\{g_n(x_\ell)\}_{n \geq 1}$ est une suite convergente, il existe un entier N_ℓ tel que $\|g_n(x_\ell) - g_m(x_\ell)\| < \varepsilon$ pour tout $n, m \geq N_\ell$. En prenant $N := \max\{N_\ell; \ell = 0, 1, \dots, q\}$, on obtient alors $\|g_n(x) - g_m(x)\| < 3\varepsilon$ pour $n, m \geq N$. Comme N ne dépend pas de x , la convergence uniforme de $\{g_n(x)\}_{n \geq 1}$ sur $[a, b]$ est démontrée. La continuité de la fonction limite suit alors d'un résultat du cours "Analyse I" (voir [HW95, III.4]). \square

Théorème 3.5 (Peano 1890) *Considérons l'ensemble $A = \{(x, y) \in \mathbb{R} \times \mathbb{R}^n; |x - x_0| \leq a, \|y - y_0\| \leq b\}$ et supposons que $f : A \rightarrow \mathbb{R}^n$ soit continue, que $\|f(x, y)\| \leq M$ sur A et que $\alpha = \min(a, b/M)$. Alors, le problème de Cauchy $y' = f(x, y)$, $y(x_0) = y_0$ possède au moins une solution sur $[x_0 - \alpha, x_0 + \alpha]$.*

Démonstration. Nous considérons les polygones d'Euler $y_h(x)$ pour $h = \alpha/N$ avec $N \in \mathbb{N}$. Cette suite est bornée ($\|y_h(x) - y_0\| \leq M|x - x_0| \leq M\alpha$) et équicontinue (Lemme 3.2). Par le Théorème de Arzelà-Ascoli, la famille $\{y_h(x)\}$ possède une sous-suite qui converge uniformément vers une fonction continue $y : [x_0, x_0 + \alpha] \rightarrow \mathbb{R}^n$. Il reste à démontrer que $y(x)$ est une solution du problème de Cauchy. L'existence d'une solution sur $[x_0 - \alpha, x_0]$ se démontre de la même manière.

Pour un $x \in [x_0, x_0 + \alpha]$ donné, notons par $\ell = \ell(h)$ l'indice tel que $x \in [x_\ell, x_{\ell+1}]$ où $x_\ell = x_0 + \ell h$. La valeur de $y_h(x)$ peut donc être écrite comme

$$y_h(x) - y_0 = hf(x_0, y_0) + \dots + hf(x_{\ell-1}, y_{\ell-1}) + (x - x_\ell)f(x_\ell, y_\ell) \quad (3.3)$$

où (x_j, y_j) sont les approximations obtenues par la méthode d'Euler. Comme $f(t, y(t))$ est continue et donc intégrable au sens de Riemann, on a que

$$\int_{x_0}^x f(t, y(t)) dt = hf(x_0, y(x_0)) + \dots + hf(x_{\ell-1}, y(x_{\ell-1})) + (x - x_\ell)f(x_\ell, y(x_\ell)) + r(h) \quad (3.4)$$

avec $r(h) \rightarrow 0$ pour $h \rightarrow 0$. La continuité uniforme de f sur A et la convergence uniforme de la sous-suite de $\{y_h(x)\}$ vers $y(x)$ impliquent que $\|f(x, y_h(x)) - f(x, y(x))\| < \varepsilon$ pour h suffisamment petit (évidemment seulement pour les h tels que $y_h(x)$ appartienne à la sous-suite). En utilisant $y_h(x_j) = y_j$, la différence de (3.3) et (3.4) donne l'estimation

$$\left\| y_h(x) - y_0 - \int_{x_0}^x f(t, y(t)) dt \right\| \leq |x - x_0| \varepsilon + \|r(h)\| \leq \alpha \varepsilon + \|r(h)\|,$$

ce qui devient arbitrairement petit pour $h \rightarrow 0$. La fonction continue $y(x)$ satisfait donc l'équation intégrale (2.2), et elle est une solution du problème de Cauchy. \square

III.4 Prolongement des solutions et existence globale

Même si la fonction $f(x, y)$ est de classe \mathcal{C}^1 partout dans $\mathbb{R} \times \mathbb{R}^n$, on n'a pas la garantie que la solution du problème de Cauchy $y' = f(x, y), y(x_0) = y_0$ existe pour tout $x \in \mathbb{R}$. Par exemple, le problème $y' = 1 + y^2$ possède $y(x) = \tan(x - c)$ comme solution et cette solution ne peut pas être prolongée à un intervalle plus grand que $(c - \pi/2, c + \pi/2)$, car $y(x) \rightarrow \pm\infty$ si $x \rightarrow c \pm \pi/2$.

Nous allons démontrer que, pour un problème où $f(x, y)$ est de classe \mathcal{C}^1 sur $\mathbb{R} \times \mathbb{R}^n$, on a seulement deux possibilités: soit la solution existe sur $[x_0, \infty)$, soit elle tend vers l'infini en temps fini, c.-à-d., elle existe sur $[x_0, \omega_+)$ avec $\omega_+ < \infty$ et on a que $\|y(x)\| \rightarrow \infty$ si $x \rightarrow \omega_+$. On a des affirmations analogues pour $x < x_0$.

Définition 4.1 Une fonction $f : U \rightarrow \mathbb{R}^n$ (où U est un ouvert de $\mathbb{R} \times \mathbb{R}^n$) satisfait *localement une condition de Lipschitz*, si pour tout $(x_0, y_0) \in U$ il existe un voisinage $V \subset U$ tel que f satisfait une condition de Lipschitz sur V (voir (2.5)).

Si la fonction $f : U \rightarrow \mathbb{R}^n$ est de classe \mathcal{C}^1 , elle satisfait localement une condition de Lipschitz. Ceci est une conséquence immédiate du théorème des accroissements finis.

Lemme 4.2 Soit $U \subset \mathbb{R} \times \mathbb{R}^n$ un ouvert et supposons que $f : U \rightarrow \mathbb{R}^n$ soit continue et qu'elle satisfasse localement une condition de Lipschitz. Alors, pour tout $(x_0, y_0) \in U$ il existe un intervalle ouvert $I_{\max} = (\omega_-, \omega_+)$ avec $-\infty \leq \omega_- < x_0 < \omega_+ \leq \infty$ ayant les propriétés suivantes:

- le problème $y' = f(x, y), y(x_0) = y_0$ possède une solution unique sur I_{\max} ,
- si $z : I \rightarrow \mathbb{R}^n$ est une solution de $y' = f(x, y), y(x_0) = y_0$, alors $I \subset I_{\max}$ et $z = y|_I$.

Démonstration. a) Soient $y : I \rightarrow \mathbb{R}^n$ et $z : J \rightarrow \mathbb{R}^n$ deux solutions de $y' = f(x, y), y(x_0) = y_0$ sur des intervalles I et J contenant x_0 . Alors, on a $y(x) = z(x)$ sur $I \cap J$. Pour le démontrer, supposons (par l'absurde) qu'il existe un $\bar{x} \in I \cap J$ tel que $y(\bar{x}) \neq z(\bar{x})$ et considérons le premier point où les deux solutions se séparent. Le Théorème 2.3 montre qu'un tel point ne peut pas exister.

b) Définissons l'intervalle

$$I_{\max} := \bigcup \left\{ I \mid \begin{array}{l} I \text{ est un intervalle ouvert, } x_0 \in I \text{ et} \\ y' = f(x, y), y(x_0) = y_0 \text{ possède une solution sur } I \end{array} \right\}. \quad (4.1)$$

Cet intervalle est ouvert, car il est l'union d'intervalles ouverts. Nous définissons maintenant $y : I_{\max} \rightarrow \mathbb{R}^n$ de la manière suivante: chaque $x \in I_{\max}$ est contenu dans un intervalle I où le problème de Cauchy possède une solution. On peut donc définir $y(x)$ comme la valeur de cette solution. La partie (a) montre que la fonction $y(x)$ est bien définie sur I_{\max} et que $y(x)$ est une solution du problème de Cauchy. L'unicité de la solution est une conséquence du Théorème 2.3. \square

Pour l'exemple $y' = 1 + y^2$ du début de ce paragraphe, nous avons $I_{\max} = I_{\max}(x_0, y_0) = (c - \pi/2, c + \pi/2)$ où $c = x_0 - \arctan y_0$.

Théorème 4.3 Supposons que la fonction $f : U \rightarrow \mathbb{R}^n$ (U un ouvert de $\mathbb{R} \times \mathbb{R}^n$) soit continue et qu'elle satisfasse localement une condition de Lipschitz sur U . Alors, chaque solution de $y' = f(x, y)$ possède un prolongement jusqu'au bord de U . Plus précisément, soit $y : I_{\max} \rightarrow \mathbb{R}^n$ la solution passant par $(x_0, y_0) \in U$, alors pour tout compact $K \subset U$ il existe $x_1, x_2 \in I_{\max}$ avec $x_1 < x_0 < x_2$ tel que $(x_1, y(x_1)) \notin K$ et $(x_2, y(x_2)) \notin K$.

Démonstration. Soit $I_{\max} = (\omega_-, \omega_+)$. Si $\omega_+ = \infty$, il est évident qu'il existe un $x_2 > x_0$ satisfaisant $(x_2, y(x_2)) \notin K$ (car K est borné). Considérons alors le cas où $\omega_+ < \infty$, et supposons par l'absurde qu'il existe un compact $K \subset U$ tel que $(x, y(x)) \in K$ pour tout $x \in (x_0, \omega_+)$. Comme $f(x, y)$ est bornée sur K , nous avons que

$$\|y(x) - y(\bar{x})\| = \left\| \int_{\bar{x}}^x f(t, y(t)) dt \right\| \leq M \cdot |x - \bar{x}| < \varepsilon$$

si x et \bar{x} sont suffisamment proches de ω_+ . Par conséquent, la limite $\lim_{x \rightarrow \omega_+} y(x) = y_+$ existe. Le point (ω_+, y_+) est dans $K \subset U$, car K est fermé. On peut donc appliquer le Théorème 2.3 qui garantit l'existence d'une solution de $y' = f(x, y)$, $y(\omega_+) = y_+$ dans un voisinage de ω_+ . Ceci contredit la maximalité de l'intervalle I_{\max} . L'existence de x_1 vérifiant les propriétés demandées est démontrée de la même manière. \square

Une équation différentielle où f ne dépend que de y s'appelle une *équation différentielle autonome*. Si $y(x)$ est une solution de $y' = f(y)$, alors $z(x) = y(x - c)$ est aussi une solution (on vérifie que $z'(x) = y'(x - c) = f(y(x - c)) = f(z(x))$). On peut donc fixer la valeur de x_0 à 0.

Théorème 4.4 (systèmes dissipatifs) *Supposons que $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ satisfasse localement une condition de Lipschitz et que*

$$\langle f(y), y - v \rangle \leq \alpha - \beta \|y\|^2 \quad (4.2)$$

est vérifié avec un vecteur $v \in \mathbb{R}^n$ et avec $\alpha \geq 0$ et $\beta > 0$. Alors, quelque soit y_0 , le problème

$$y' = f(y), \quad y(0) = y_0 \quad (4.3)$$

possède une solution qui existe pour tout $x \geq 0$.

Démonstration. Considérons les boules $B_0 = \{y \in \mathbb{R}^n ; \|y\|^2 \leq \alpha/\beta\}$ et $B = \{y \in \mathbb{R}^n ; \|y - v\| \leq R\}$ avec un $R > \|v\| + \sqrt{\alpha/\beta}$ suffisamment grand pour que B_0 soit contenue dans l'intérieur de B . Pour $y \in \partial B$, on a $\langle f(y), y - v \rangle < 0$, ce qui signifie que le vecteur $f(y)$ pointe vers l'intérieur de B . Les solutions ne peuvent donc pas sortir de la boule B . Si l'on choisit R assez grand pour que $y_0 \in B$, l'existence globale de la solution sur $[0, \infty)$ est une conséquence du Théorème 4.3. \square

Exemple 4.5 Les équations de Lorenz sont données par

$$\begin{aligned} y_1' &= -\sigma y_1 + \sigma y_2 \\ y_2' &= -y_1 y_3 + r y_1 - y_2 \\ y_3' &= y_1 y_2 - b y_3 \end{aligned} \quad (4.4)$$

où $\sigma = 10$, $r = 28$ et $b = 8/3$. La Fig. 4.1 montre la première composante en fonction du temps pour $0 \leq x \leq 100$ et la Fig. 4.2 montre la solution dans l'espace de phase pour

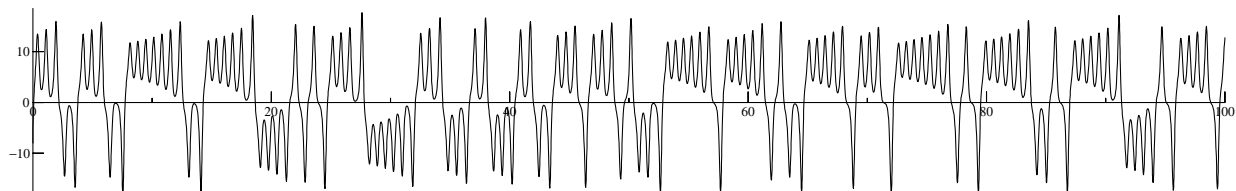


FIG. 4.1 – Première composante de (4.4) en fonction du temps

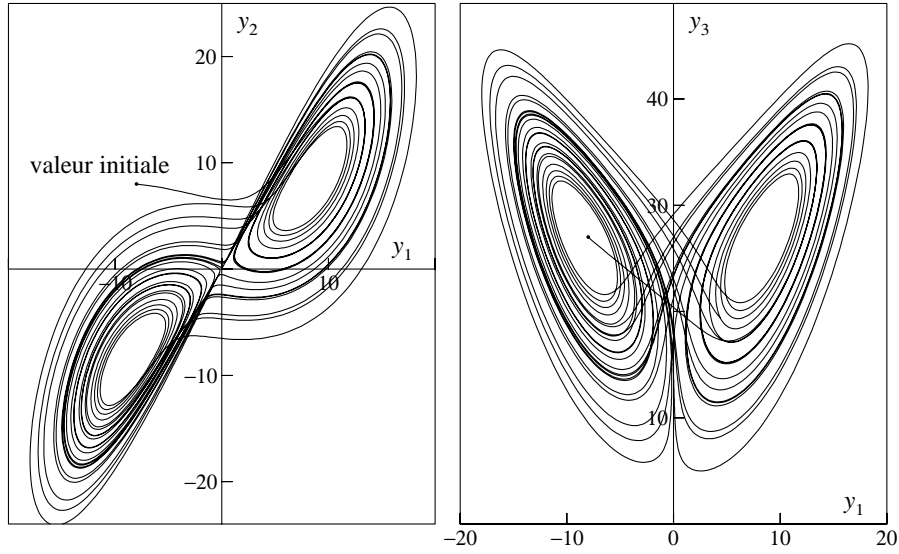


FIG. 4.2 – Deux vues de la solution de l'équation de Lorenz (4.4)

$0 \leq x \leq 30$ avec comme valeurs initiales $y_1(0) = -8$, $y_2(0) = 8$ et $y_3(0) = 27$. Elle est très “chaotique”.

Nous allons montrer que ce système est dissipatif. Pour cela, nous considérons un vecteur $v = (0, 0, \gamma)^T$ et nous obtenons

$$\langle f(y), y - v \rangle = -\sigma y_1^2 - y_2^2 - b y_3^2 + (\sigma + r - \gamma) y_1 y_2 + b \gamma y_3.$$

Avec le choix $\gamma = \sigma + r$ et en utilisant l'inégalité $2\gamma y_3 \leq y_3^2 + \gamma^2$ (ceci est une conséquence de $(y_3 - \gamma)^2 \geq 0$), on voit que (4.2) est vérifié avec pour $\alpha = b\gamma^2/2$ et pour $\beta = \min(\sigma, 1, b/2)$. La solution du système de Lorenz existe donc pour tout $x \in [0, \infty)$, quelle que soit sa valeur initiale.

III.5 Inégalités différentielles et solutions approchées

Une solution approchée du problème de Cauchy

$$y' = f(x, y), \quad y(x_0) = y_0 \tag{5.1}$$

est une fonction $v(x)$ pour laquelle le défaut $\delta(x) := v'(x) - f(x, v(x))$ est petit. On cherche à obtenir des estimations pour $\|v(x) - y(x)\|$. Souvent la fonction $v(x)$ n'est pas partout différentiable (pensons aux polygones d'Euler). Nous introduirons alors la dérivée de Dini et nous utiliserons un résultat sur les inégalités différentielles pour obtenir une estimation de $\|v(x) - y(x)\|$.

Définition 5.1 Soit $m : I \rightarrow \mathbb{R}$ une fonction continue sur l'intervalle I . La *dérivée de Dini* est définie par

$$D_+ m(x) = \liminf_{h \rightarrow 0^+} \frac{m(x+h) - m(x)}{h}.$$

Si m possède une dérivée à droite en x , alors $D_+ m(x) = m'(x+0)$. Dans cette situation, on a pour toute norme

$$D_+ \|m(x)\| \leq \|m'(x+0)\|. \tag{5.2}$$

Ceci découle de l'inégalité $\|m(x+h)\| - \|m(x)\| \leq \|m(x+h) - m(x)\|$. Un exemple sans dérivée à droite en 0 est la fonction $m(x) = x \sin(1/x)$. Pour cette fonction, on a $D_+m(0) = -1$.

Théorème 5.2 Soient $m : [x_0, x_0 + a) \rightarrow \mathbb{R}$ et $u : [x_0, x_0 + a) \rightarrow \mathbb{R}$ deux fonctions continues satisfaisant

$$\begin{aligned} D_+m(x) &\leq g(x, m(x)) && \text{pour } x \in [x_0, x_0 + a) \\ D_+u(x) &> g(x, u(x)) && \text{pour } x \in [x_0, x_0 + a) \\ m(x_0) &\leq u(x_0). \end{aligned}$$

Alors, on a $m(x) \leq u(x)$ pour tout $x \in [x_0, x_0 + a)$ (ici, g est une fonction arbitraire à deux variables réelles).

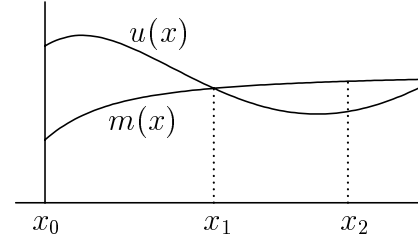
Démonstration. Supposons qu'il existe $x_2 \in (x_0, x_0 + a)$ avec $m(x_2) > u(x_2)$. Alors, la valeur

$$x_1 := \inf \{x ; x \geq x_0 \text{ et } m(t) > u(t) \text{ sur } [x, x_2]\}$$

est bien définie et, par continuité des fonctions m et u , on a que $m(x_1) = u(x_1)$ et $m(x_1 + h) > u(x_1 + h)$ pour $h > 0$ suffisamment petit. Par conséquent,

$$\frac{m(x_1 + h) - m(x_1)}{h} > \frac{u(x_1 + h) - u(x_1)}{h}$$

pour $h > 0$ et petit. On en déduit que $D_+m(x_1) \geq D_+u(x_1)$. Mais ceci est en contradiction avec $D_+m(x_1) \leq g(x_1, m(x_1)) = g(x_1, u(x_1)) < D_+u(x_1)$. \square



Le résultat suivant est fondamental pour toutes les estimations concernant des solutions approchées.

Théorème 5.3 (lemme fondamental) Soit $U \subset \mathbb{R} \times \mathbb{R}^n$ un ouvert, $f : U \rightarrow \mathbb{R}^n$ une fonction continue, $y : [x_0, x_0 + a) \rightarrow \mathbb{R}^n$ une solution de $y' = f(x, y)$ et $v : [x_0, x_0 + a) \rightarrow \mathbb{R}^n$ une fonction continue qui possède partout une dérivée à droite. Alors, les estimations

$$\begin{aligned} \|v'(x+0) - f(x, v(x))\| &\leq \delta && (v \text{ est solution } \delta\text{-approchée}) \\ \|f(x, y(x)) - f(x, v(x))\| &\leq L \|y(x) - v(x)\| \end{aligned}$$

impliquent que pour tout $x \in [x_0, x_0 + a)$

$$\|y(x) - v(x)\| \leq \|y(x_0) - v(x_0)\| e^{L(x-x_0)} + \frac{\delta}{L} (e^{L(x-x_0)} - 1).$$

Démonstration. Posons $m(x) := \|y(x) - v(x)\|$. En utilisant (5.2) et les hypothèses du théorème, nous avons que

$$D_+m(x) \leq \|y'(x) - v'(x+0)\| \leq L m(x) + \delta.$$

Définissons maintenant $u(x)$ par $u(x_0) = m(x_0)$ et

$$u'(x) = L u(x) + \delta + \varepsilon > L u(x) + \delta$$

avec un $\varepsilon > 0$. La solution de cette équation différentielle est donnée par

$$u(x) = u(x_0) e^{L(x-x_0)} + \frac{(\delta + \varepsilon)}{L} (e^{L(x-x_0)} - 1).$$

Le Théorème 5.2 avec pour $g(x, y) = Ly + \delta$ montre que

$$\|y(x) - v(x)\| \leq \|y(x_0) - v(x_0)\| e^{L(x-x_0)} + \frac{(\delta + \varepsilon)}{L} (e^{L(x-x_0)} - 1)$$

pour tout $\varepsilon > 0$. La limite $\varepsilon \rightarrow 0$ donne l'affirmation du théorème. \square

Dépendance continue des solutions des valeurs initiales Supposons que $f : U \rightarrow \mathbb{R}^n$ vérifie les hypothèses du Théorème 2.3 (existence et unicité du problème de Cauchy). Nous exprimons par $y(x, x_0, y_0)$ la dépendance de la solution des valeurs initiales. Cette fonction de trois variables est définie sur

$$D = \{(x, x_0, y_0) ; (x_0, y_0) \in U \text{ et } x \in I_{\max}(x_0, y_0)\}. \quad (5.3)$$

Nous allons démontrer que $y : D \rightarrow \mathbb{R}^n$ est une fonction continue. Remarquons que l'unicité de la solution implique que

$$y(x, x_0, y_0) = y(x, x_1, y(x_1, x_0, y_0)). \quad (5.4)$$

En effet, les deux fonctions de cette identité sont solutions de $y' = f(x, y)$ et elles passent par le même point $(x_1, y(x_1, x_0, y_0))$.

Lemme 5.4 *Supposons que $f : U \rightarrow \mathbb{R}^n$ (U un ouvert de $\mathbb{R} \times \mathbb{R}^n$) soit continue et qu'elle satisfasse localement une condition de Lipschitz. Alors, pour tout compact $K \subset U$ il existe un $L \geq 0$ tel que*

$$\|f(x, y) - f(x, z)\| \leq L \|y - z\| \quad \text{pour tout } (x, y), (x, z) \in K.$$

Démonstration. Supposons le contraire. Il existe alors des suites $(x_n, y_n), (x_n, z_n)$ dans K telles que

$$\|f(x_n, y_n) - f(x_n, z_n)\| > n \|y_n - z_n\|. \quad (5.5)$$

Soit (x, y) un point d'accumulation de la suite (x_n, y_n) (il existe par le théorème de Bolzano-Weierstrass). Par hypothèse, la fonction satisfait une condition de Lipschitz dans un voisinage V de (x, y) . Comme $f(x, y)$ est bornée sur K et comme $\|y_n - z_n\| < 2M/n$ où $M = \max_{(x,y) \in K} \|f(x, y)\|$, il y a un nombre infini d'indices n pour lesquelles (x_n, y_n) et (x_n, z_n) sont dans V . La condition de Lipschitz sur V est alors en contradiction avec (5.5). \square

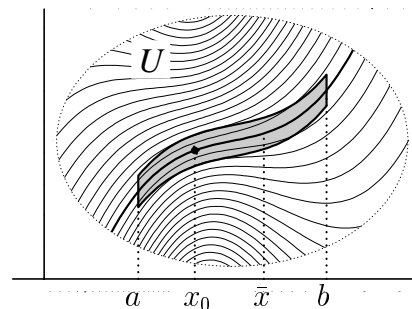
Théorème 5.5 *Supposons que $f : U \rightarrow \mathbb{R}^n$ (U un ouvert de $\mathbb{R} \times \mathbb{R}^n$) soit continue et qu'elle satisfasse localement une condition de Lipschitz. Alors, l'ensemble D de (5.3) est ouvert et la solution $y : D \rightarrow \mathbb{R}^n$ du problème de Cauchy est une fonction continue de trois variables (x, x_0, y_0) .*

Démonstration. Soit (\bar{x}, x_0, y_0) fixé et considérons un intervalle $[a, b] \subset I_{\max}(x_0, y_0)$ tel que $\bar{x}, x_0 \in (a, b)$. Nous choisissons maintenant un $\rho > 0$ suffisamment petit pour que le tuyau

$$K = \{(x, y) ; x \in [a, b], \|y - y(x, x_0, y_0)\| \leq \rho\}$$

soit entièrement inclus dans l'ouvert U . Par le Lemme 5.4 la fonction $f(x, y)$ satisfait une condition de Lipschitz (avec constante L) sur le compact K . L'ensemble

$$V = \{(x_1, y_1) ; x_1 \in (a, b), \|y_1 - y(x_1, x_0, y_0)\| < \rho e^{-L(b-a)}\}$$



est un voisinage de (x_0, y_0) qui satisfait $V \subset K \subset U$. Si $(x_1, y_1) \in V$, l'estimation du lemme fondamental garantit que $y(x, x_1, y_1)$ reste dans K pour tout $x \in [a, b]$, car

$$\begin{aligned} \|y(x, x_1, y_1) - y(x, x_0, y_0)\| &= \|y(x, x_1, y_1) - y(x, x_1, y(x_1, x_0, y_0))\| \\ &\leq e^{L|x-x_1|} \|y_1 - y(x_1, x_0, y_0)\| < \rho. \end{aligned} \quad (5.6)$$

Par conséquent, $[a, b] \subset I_{\max}(x_1, y_1)$ et $(a, b) \times V$ est un voisinage de (\bar{x}, x_0, y_0) inclus dans D . Donc, D est ouvert.

Pour démontrer la continuité de y , nous observons que pour $(x_1, y_1) \in V$ et $x \in (a, b)$

$$\|y(x, x_1, y_1) - y(\bar{x}, x_1, y_1)\| = \left\| \int_{\bar{x}}^x f(t, y(t, x_1, y_1)) dt \right\| \leq M|x - \bar{x}| \quad (5.7)$$

où M est une borne supérieure de $f(x, y)$ sur K . Comme dans (5.6) nous estimons

$$\begin{aligned} \|y(\bar{x}, x_1, y_1) - y(\bar{x}, x_0, y_0)\| &\leq e^{L|\bar{x}-x_1|} \|y_1 - y(x_1, x_0, y_0)\| \\ &\leq e^{L(b-a)} (\|y_1 - y_0\| + M|x_1 - x_0|). \end{aligned} \quad (5.8)$$

Les deux estimations (5.7) et (5.8) démontrent la continuité de $y : D \rightarrow \mathbb{R}^n$. \square

III.6 Systèmes d'équations différentielles linéaires

Dans ce paragraphe nous considérons des équations différentielles

$$y' = A(x)y + g(x) \quad (6.1)$$

où $A(x)$ est une matrice $n \times n$ et $g(x)$ est un vecteur. On dit que cette équation différentielle est *homogène* si $g(x) \equiv 0$, sinon elle est *inhomogène*.

Théorème 6.1 (existence globale et unicité) *Soit I un intervalle (arbitraire) et supposons que $A(x)$ et $g(x)$ soient des fonctions continues sur I . Alors, le problème $y' = A(x)y + g(x)$, $y(x_0) = y_0$ (avec $x_0 \in I$ et $y_0 \in \mathbb{R}^n$) possède une solution unique sur tout l'intervalle I .*

Démonstration. La fonction $f(x, y) = A(x)y + g(x)$ est continue sur $I \times \mathbb{R}^n$ et satisfait localement une condition de Lipschitz. La solution est donc unique là où elle existe. Pour démontrer l'existence globale, nous montrons que la solution $y(x)$ ne peut pas tendre vers ∞ pour $x \rightarrow a$ si a est un point à l'intérieur de I (Théorème 4.3). En effet, sur l'intervalle compact $[x_0, a]$, les fonctions $A(x)$ et $g(x)$ sont bornées ($\|A(x)\| \leq L$ et $\|g(x)\| \leq \delta$ sur $[x_0, a]$). Le lemme fondamental (Théorème 5.3) avec $v(x) \equiv 0$ donne alors l'estimation

$$\|y(x)\| \leq \|y_0\| e^{L(a-x_0)} + \frac{\delta}{L} (e^{L(a-x_0)} - 1)$$

pour tout $x \in [x_0, a]$. \square

Théorème 6.2 (principe de superposition) *Soit I un intervalle et soient $A(x)$, $g_1(x)$, $g_2(x)$ des fonctions continues sur I . Si*

$$\begin{aligned} y_1 : I &\rightarrow \mathbb{R}^n \text{ est solution de } y' = A(x)y + g_1(x), \\ y_2 : I &\rightarrow \mathbb{R}^n \text{ est solution de } y' = A(x)y + g_2(x), \end{aligned}$$

alors $y(x) := c_1 y_1(x) + c_2 y_2(x)$ est solution de

$$y' = A(x)y + g(x) \quad \text{avec} \quad g(x) := c_1 g_1(x) + c_2 g_2(x).$$

Démonstration. Ceci est un exercice très simple. \square

Alors que la démonstration de ce résultat est très simple, il a des conséquences très importantes :

- les solutions de $y' = A(x)y$ forment un espace vectoriel;
- les solutions de $y' = A(x)y$ dépendent linéairement de y_0 , c.-à-d.,

$$y(x, x_0, c_1 y_0 + c_2 z_0) = c_1 y(x, x_0, y_0) + c_2 y(x, x_0, z_0).$$

La solution de $y' = A(x)y$, $y(x_0) = y_0$ peut donc être écrite sous la forme

$$y(x, x_0, y_0) = R(x, x_0)y_0, \quad (6.2)$$

et la matrice $R(x, x_0)$ s'appelle la *résolvante* de l'équation différentielle $y' = A(x)y$. La i ème colonne de la matrice $R(x, x_0)$ est solution de $y' = A(x)y$ avec pour valeur initiale $y(x_0) = e_i = (0, \dots, 0, 1, 0, \dots, 0)^T$;

- si $\Phi(x)$ est une *matrice fondamentale* (aussi appelé *wronskienne*), c.-à-d., les colonnes de $\Phi(x)$ sont des solutions de $y' = A(x)y$ et $\Phi(x_0)$ est inversible, alors

$$R(x, x_0) = \Phi(x)\Phi(x_0)^{-1}. \quad (6.3)$$

En effet, $y(x) = \Phi(x)\Phi(x_0)^{-1}y_0$ est solution de $y' = A(x)y$, $y(x_0) = y_0$.

Théorème 6.3 (propriétés de la résolvante) Soit $A(x)$ continue sur un intervalle. Alors, la résolvante de $y' = A(x)y$ satisfait

- i) $R'(x, x_0) = A(x)R(x, x_0)$ (dérivée par rapport à x)
- ii) $R(x_0, x_0) = I$ (matrice identité)
- iii) $R(x, x_0) = R(x, x_1)R(x_1, x_0)$
- iv) $R(x, x_0)$ est inversible et $R(x, x_0)^{-1} = R(x_0, x)$.

Démonstration. Par (6.2) on a que $R'(x, x_0)y_0 = A(x)R(x, x_0)y_0$ et $R(x_0, x_0)y_0 = y_0$ pour tout $y_0 \in \mathbb{R}^n$. Ceci démontre les propriétés (i) et (ii). La propriété (iii) est une conséquence de (5.4) et (iv) suit de (iii) en posant $x = x_0$. \square

Exemple 6.4 L'équation différentielle $y'' + y = 0$ peut être écrite sous la forme (en posant $y_1 = y$ et $y_2 = y'$)

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}' = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

et possède comme résolvante

$$R(x, x_0) = \begin{pmatrix} \cos(x - x_0) & \sin(x - x_0) \\ -\sin(x - x_0) & \cos(x - x_0) \end{pmatrix}.$$

Il est intéressant que la propriété (iii) du théorème précédent, c.-à-d.,

$$\begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} \cos \beta & \sin \beta \\ -\sin \beta & \cos \beta \end{pmatrix} = \begin{pmatrix} \cos(\alpha + \beta) & \sin(\alpha + \beta) \\ -\sin(\alpha + \beta) & \cos(\alpha + \beta) \end{pmatrix}$$

avec $\alpha = x - x_1$, $\beta = x_1 - x_0$ et $\alpha + \beta = x - x_0$, est équivalente au théorème d'addition pour $\sin x$ et $\cos x$.

Théorème 6.5 (Liouville) Soit $A(x)$ continue sur un intervalle et soit $\Phi(x)$ une matrice fondamentale de $y' = A(x)y$. Alors,

$$\det \Phi(x) = \det \Phi(x_0) \cdot \exp\left(\int_{x_0}^x \text{trace } A(t) dt\right)$$

où $\text{trace } A(x) = a_{11}(x) + a_{22}(x) + \dots + a_{nn}(x)$.

Démonstration. Soit $\Phi(x) = (\Phi_{ij}(x))_{i,j=1}^n$. Comme $\det \Phi(x)$ est une forme multilinéaire par rapport aux lignes de $\Phi(x)$, on a que

$$\frac{d}{dx}(\det \Phi(x)) = \sum_{i=1}^n \det D_i(x) \quad \text{où} \quad D_i(x) = \begin{pmatrix} \Phi_{11}(x) & \dots & \Phi_{1n}(x) \\ \Phi'_{i1}(x) & \dots & \Phi'_{in}(x) \\ \Phi_{n1}(x) & \dots & \Phi_{nn}(x) \end{pmatrix}.$$

La matrice $D_i(x)$ est obtenue de $\Phi(x)$ en remplaçant la i ème ligne de $\Phi(x)$ par sa dérivée. En utilisant l'équation différentielle $\Phi'(x) = A(x)\Phi(x)$, c.-à-d., $\Phi'_{ij}(x) = \sum_{k=1}^n a_{ik}(x)\Phi_{kj}(x)$, et la multilinéarité du déterminant, on obtient que

$$\begin{aligned} \frac{d}{dx}(\det \Phi(x)) &= \sum_{i=1}^n \sum_{k=1}^n a_{ik}(x) \det \begin{pmatrix} \Phi_{11}(x) & \dots & \Phi_{1n}(x) \\ \Phi_{k1}(x) & \dots & \Phi_{kn}(x) \\ \Phi_{n1}(x) & \dots & \Phi_{nn}(x) \end{pmatrix} \leftarrow \text{ligne } i \\ &= \sum_{i=1}^n a_{ii}(x) \det \Phi(x). \end{aligned}$$

Il ne reste plus qu'à résoudre cette équation différentielle pour $\det \Phi(x)$. □

Le théorème de Liouville possède une interprétation intéressante. Si $V = (v_1, \dots, v_n)$ est une matrice dont les colonnes sont les vecteurs v_1, \dots, v_n , la valeur de $\det V$ représente le volume du parallélépipède engendré par v_1, \dots, v_n . Considérons maintenant le *flot* de l'équation différentielle $y' = A(x)y$, c.-à-d., l'application $\varphi_{x,x_0} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ définie par $\varphi_{x,x_0}(y_0) = R(x, x_0)y_0$. Si $\text{trace } A(x) \equiv 0$, on déduit du théorème de Liouville que le volume d'un sous-ensemble B de \mathbb{R}^n est invariant par l'application φ_{x,x_0} , c.-à-d., B et $\varphi_{x,x_0}(B)$ possèdent le même volume. Pour l'équation différentielle de l'Exemple 6.4, la préservation du volume est évidente, car l'application φ_{x,x_0} n'est rien d'autre qu'une rotation par l'angle $x - x_0$.

Equations linéaires inhomogènes Considérons maintenant l'équation différentielle (6.1) où $g(x)$ n'est pas nulle. Si l'on connaît la résolvante de l'équation homogène, on trouve la solution du problème inhomogène par la méthode des variations des constantes.

Théorème 6.6 (“variation des constantes”) Soient $A(x)$ et $g(x)$ continues sur un intervalle et soit $R(x, x_0)$ la résolvante de $y' = A(x)y$. Alors, la solution de $y' = A(x)y + g(x)$, $y(x_0) = y_0$ est donnée par

$$y(x) = R(x, x_0)y_0 + \int_{x_0}^x R(x, t)g(t) dt.$$

Démonstration. La solution générale de l'équation homogène est $R(x, x_0)c$ avec $c \in \mathbb{R}^n$. L'idée (d'où le nom “variation des constantes”) est de chercher une solution de $y' = A(x)y + g(x)$ sous la forme $y(x) = R(x, x_0)c(x)$. Il faut alors que

$$y'(x) = R'(x, x_0)c(x) + R(x, x_0)c'(x) = A(x)R(x, x_0)c(x) + g(x).$$

En utilisant la propriété (i) du Théorème 6.3, nous obtenons

$$c'(x) = R(x, x_0)^{-1}g(x) = R(x_0, x)g(x)$$

et par intégration $c(x) = c(x_0) + \int_{x_0}^x R(x_0, t)g(t) dt$. L'affirmation du théorème découle alors en insérant cette formule pour $c(x)$ dans $y(x) = R(x, x_0)c(x)$. \square

Ce résultat montre que, comme dans le cas particulier de dimension 1, la solution générale de $y' = A(x)y + g(x)$ est composée de la solution générale de l'équation homogène et d'une solution particulière de l'équation inhomogène. Il reste alors à discuter le calcul de la résolvante $R(x, x_0)$. Pour le cas $A(x) = A$ (matrice constante), ceci est le sujet du paragraphe suivant. Si $A(x)$ dépend de x , le calcul analytique de la résolvante est rarement possible.

III.7 Systèmes linéaires à coefficients constants

Considérons le problème de calculer la résolvante de

$$y' = Ay \tag{7.1}$$

où A est une matrice constante d'ordre n (les a_{ij} sont réels ou complexes). Motivés par la résolution de problèmes scalaires, essayons de trouver des solutions de la forme

$$y(x) = e^{\lambda x}v \quad \text{avec un vecteur } v \neq 0. \tag{7.2}$$

En insérant (7.2) dans (7.1) nous obtenons $y'(x) = \lambda e^{\lambda x}v = e^{\lambda x}Av$, ce qui est équivalent à $Av = \lambda v$. La fonction de (7.2) est alors une solution de (7.1) si et seulement si λ est une valeur propre de A et v est un vecteur propre correspondant.

Cas 1 (A est diagonalisable) Dans cette situation, il existe n vecteurs propres indépendants v_1, \dots, v_n avec valeurs propres $\lambda_1, \dots, \lambda_n$. Considérons la matrice

$$\Phi(x) = (e^{\lambda_1 x}v_1, \dots, e^{\lambda_n x}v_n). \tag{7.3}$$

Par le Théorème 6.5 de Liouville, $\Phi(x)$ est inversible pour tout x car $\Phi(0)$ l'est. La résolvante est alors donnée par $R(x, x_0) = \Phi(x)\Phi(x_0)^{-1}$ (voir l'équation (6.3)) ou aussi par $R(x, x_0) = \Phi(x - x_0)\Phi(0)^{-1}$.

Cas 2 (A n'est pas diagonalisable) L'idée est de transformer A sous une forme "plus simple", par exemple sous forme triangulaire ou sous forme de Jordan. On cherche alors une matrice inversible T telle que $T^{-1}AT = S$ possède une telle forme. Avec la transformation $y = Tz$ (et $y' = Tz'$) le système (7.1) devient $z' = Sz$. Pour le cas d'une matrice triangulaire on a

$$\begin{aligned} z'_1 &= s_{11}z_1 + s_{12}z_2 + \dots + s_{1n}z_n \\ z'_2 &= \phantom{s_{11}z_1} + s_{22}z_2 + \dots + s_{2n}z_n \\ & \phantom{+ s_{22}z_2} + \dots \\ z'_n &= \phantom{s_{11}z_1} \phantom{+ s_{22}z_2} + s_{nn}z_n \end{aligned}$$

On résoud d'abord l'équation différentielle pour z_n , ensuite celle pour z_{n-1} et à la fin celle pour z_1 . La solution de (7.1) est obtenue par $y(x) = Tz(x)$.

Exemple 7.1 Considérons un bloc de Jordan de dimension 4

$$J = \begin{pmatrix} \lambda & 1 & & \\ & \lambda & 1 & \\ & & \lambda & 1 \\ & & & \lambda \end{pmatrix}.$$

Pour la solution de $z' = Jz$ on commence par z_4 ; on trouve $z_4' = \lambda z_4$ et $z_4(x) = e^{\lambda x} c_4$. L'équation différentielle pour z_3 est $z_3' = \lambda z_3 + e^{\lambda x} c_4$. Sa solution est $z_3(x) = e^{\lambda x} c_3 + x e^{\lambda x} c_4$. De la même façon on calcule $z_2(x)$ et $z_1(x)$. Le résultat est

$$z(x) = e^{\lambda x} \begin{pmatrix} 1 & x & x^2/2! & x^3/3! \\ & 1 & x & x^2/2! \\ & & 1 & x \\ & & & 1 \end{pmatrix} z(0). \quad (7.4)$$

On trouve la solution (7.4) de $z' = Jz$ aussi à l'aide du théorème suivant.

Théorème 7.2 La résolvante de $y' = Ay$ est donnée par

$$R(x, x_0) = \exp(A(x - x_0)) := \sum_{i \geq 0} \frac{(x - x_0)^i}{i!} A^i. \quad (7.5)$$

Démonstration. Comme $\|A^i\| \leq \|A\|^i$, la série $\sum_{i \geq 0} \frac{\alpha^i}{i!} \|A\|^i = \exp(\alpha \|A\|)$ est une majorante de la série de (7.5) pour $|x - x_0| \leq \alpha$. Ceci montre la convergence uniforme sur chaque intervalle compact. On peut donc échanger la dérivée avec la sommation et on obtient

$$\frac{d}{dx} \exp(A(x - x_0)) = \frac{d}{dx} \sum_{i \geq 0} \frac{(x - x_0)^i}{i!} A^i = \sum_{i \geq 1} \frac{(x - x_0)^{i-1}}{(i-1)!} A^i = A \exp(A(x - x_0)).$$

Chaque colonne de $\exp(A(x - x_0))$ est une solution de $y' = Ay$ et on obtient (7.5) car $\exp(0) = I$. \square

Revenons à l'Exemple 7.1. Le bloc de Jordan peut être écrit sous la forme $J = \lambda I + N$ où

$$N = \begin{pmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & 0 & 1 \\ & & & 0 \end{pmatrix}, \quad N^2 = \begin{pmatrix} 0 & 0 & 1 & \\ & 0 & 0 & 1 \\ & & 0 & 0 \\ & & & 0 \end{pmatrix}, \quad N^3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ & 0 & 0 & 0 \\ & & 0 & 0 \\ & & & 0 \end{pmatrix}$$

et $N^j = 0$ pour $j \geq 4$. Par le Théorème 7.2, la résolvante est donnée par

$$R(x, 0) = \exp((\lambda I + N)x) = e^{\lambda x} \exp(Nx) = e^{\lambda x} \sum_{i \geq 0} \frac{x^i}{i!} N^i.$$

Attention Pour $y' = A(x)y$ où $A(x)$ dépend de x , on a en général que

$$R(x, x_0) \neq \exp\left(\int_{x_0}^x A(t) dt\right) \quad (7.6)$$

(voir Exercice 30). On a l'égalité dans (7.6) seulement si la dimension est 1 ou si les matrices $A(x)$ et $\int_{x_0}^x A(t) dt$ commutent.

III.8 Différentiabilité par rapport aux valeurs initiales

Considérons la solution du problème de Cauchy

$$y' = f(x, y), \quad y(x_0) = y_0 \quad (8.1)$$

comme fonction de trois arguments $y(x, x_0, y_0)$ et étudions la sensibilité de la solution par rapport à des perturbations dans les valeurs initiales x_0, y_0 . Nous savons déjà que $y(x, x_0, y_0)$ dépend continûment de (x_0, y_0) (voir le Théorème 5.5), mais que peut-on dire sur la différentiabilité? Quelle est l'approximation du premier ordre de $y(x, x_0 + \Delta x_0, y_0 + \Delta y_0)$?

A) Différentiabilité par rapport à y_0 . Supposons que $f : U \rightarrow \mathbb{R}^n$ (avec $U \subset \mathbb{R} \times \mathbb{R}^n$ ouvert) satisfasse les hypothèses sur l'existence et l'unicité de la solution. Alors, $y(x, x_0, y_0)$ est bien définie sur l'ensemble ouvert

$$D := \{(x, x_0, y_0) ; (x_0, y_0) \in U \text{ et } x \in I_{\max}(x_0, y_0)\}. \quad (8.2)$$

Pour obtenir une idée de la forme de la dérivée $\frac{\partial y}{\partial y_0}(x, x_0, y_0)$, écrivons le problème (8.1) comme

$$\frac{\partial y}{\partial x}(x, x_0, y_0) = f(x, y(x, x_0, y_0)), \quad y(x_0, x_0, y_0) = y_0 \quad (8.3)$$

et dérivons cette équation formellement par rapport à y_0 . En échangeant les dérivées par rapport à x et y_0 , nous obtenons

$$\frac{\partial}{\partial x} \left(\frac{\partial y}{\partial y_0}(x, x_0, y_0) \right) = \frac{\partial f}{\partial y}(x, y(x, x_0, y_0)) \cdot \frac{\partial y}{\partial y_0}(x, x_0, y_0), \quad \frac{\partial y}{\partial y_0}(x_0, x_0, y_0) = I.$$

D'après ce calcul formel, nous voyons que $\frac{\partial y}{\partial y_0}(x, x_0, y_0)$ est la solution de l'équation linéaire

$$\Psi' = \frac{\partial f}{\partial y}(x, y(x, x_0, y_0)) \Psi, \quad \Psi(x_0) = I. \quad (8.4)$$

Cette équation s'appelle *équation variationnelle* de (8.1). Elle est de la forme $y' = A(x, x_0, y_0)y$ où la matrice A dépend du paramètre (x_0, y_0) . La résolvante dépend donc aussi de ce paramètre. Notons-la par $R(x, s, x_0, y_0)$. Le calcul formel montre alors que

$$\frac{\partial y}{\partial y_0}(x, x_0, y_0) = R(x, x_0, x_0, y_0). \quad (8.5)$$

Dans la démonstration du théorème suivant, nous justifierons cette procédure.

Théorème 8.1 Soit $U \subset \mathbb{R} \times \mathbb{R}^n$ un ouvert, $f : U \rightarrow \mathbb{R}^n$ continue et supposons que $\frac{\partial f}{\partial y}(x, y)$ existe et soit continue sur U . La solution $y : D \rightarrow \mathbb{R}^n$ de (8.1) satisfait alors

- $y(x, x_0, y_0)$ est continûment différentiable par rapport à y_0 ;
- $\frac{\partial y}{\partial y_0}(x, x_0, y_0)$ est la résolvante de l'équation variationnelle (8.4).

Démonstration. Pour $(\bar{x}, x_0, y_0) \in D$ fixé, il faut démontrer que

$$y(\bar{x}, x_0, y_0 + \Delta y_0) - y(\bar{x}, x_0, y_0) - R(\bar{x}, x_0, x_0, y_0)\Delta y_0 = r(\Delta y_0)\|\Delta y_0\|$$

où $r(\Delta y_0) \rightarrow 0$ si $\Delta y_0 \rightarrow 0$. L'idée de la démonstration est de considérer

$$v(x) := y(x, x_0, y_0) + R(x, x_0, x_0, y_0)\Delta y_0$$

comme solution approchée de (8.1) et d'estimer la différence $y(x, x_0, y_0 + \Delta y_0) - v(x)$ à l'aide du lemme fondamental (Théorème 5.3). Avec l'abréviation $\delta(x) := R(x, x_0, x_0, y_0)\Delta y_0$ nous avons $\|\delta(x)\| \leq M\|\Delta y_0\|$ sur l'intervalle compact $[x_0, \bar{x}]$. Considérons le tuyau $K := \{(x, y); x \in [x_0, \bar{x}], \|y - y(x, x_0, y_0)\| \leq \rho\}$ avec $\rho > 0$ suffisamment petit pour que $K \subset U$. Pour $\|\Delta y_0\| \leq \rho/M$, la fonction $v(x)$ reste dans K et son défaut est donné par

$$f(x, v(x)) - v'(x) = f(x, y(x, x_0, y_0) + \delta(x)) - f(x, y(x, x_0, y_0)) - \frac{\partial f}{\partial y}(x, y(x, x_0, y_0))\delta(x).$$

Le théorème des accroissements finis appliqué à $F(\delta) := f(x, y + \delta) - f(x, y) - \frac{\partial f}{\partial y}(x, y)\delta$ (observons que $F(0) = 0$) implique que

$$\left\| f(x, y + \delta) - f(x, y) - \frac{\partial f}{\partial y}(x, y)\delta \right\| \leq \sup_{\lambda \in [0, 1]} \left\| \frac{\partial f}{\partial y}(x, y + \lambda\delta) - \frac{\partial f}{\partial y}(x, y) \right\| \cdot \|\delta\|.$$

Comme $\frac{\partial f}{\partial y}(x, y)$ est uniformément continue sur le compact K , on a l'estimation

$$\sup_{\lambda \in [0, 1]} \left\| \frac{\partial f}{\partial y}(x, y + \lambda\delta) - \frac{\partial f}{\partial y}(x, y) \right\| \leq r_1(\|\delta\|) \quad \text{où} \quad r_1(\|\delta\|) \rightarrow 0 \quad \text{si} \quad \delta \rightarrow 0.$$

La fonction $r_1(\cdot)$ peut être choisie monotone et indépendante de $(x, y) \in K$. Par conséquent,

$$\|f(x, v(x)) - v'(x)\| \leq r_1(M\|\Delta y_0\|)M\|\Delta y_0\|.$$

Sur le compact K , $f(x, y)$ satisfait une condition de Lipschitz. On peut donc appliquer le lemme fondamental et on obtient

$$\|y(x, x_0, y_0 + \Delta y_0) - v(x)\| \leq \frac{r_1(M\|\Delta y_0\|)M\|\Delta y_0\|}{L} (e^{L(x-x_0)} - 1),$$

ce qui démontre la différentiabilité de $y(x, x_0, y_0)$ par rapport à y_0 et la formule (8.5).

Il faut encore démontrer la continuité de $\frac{\partial y}{\partial y_0}(x, x_0, y_0)$. Elle est une conséquence du Lemme 8.2 (voir plus bas). On ne peut pas directement appliquer le Théorème 5.5, car l'argument (x_0, y_0) n'est pas une valeur initiale de l'équation variationnelle (8.4) mais un paramètre de celle-ci. \square

Lemme 8.2 *Soit $y' = A(x, c)y$ une équation différentielle linéaire qui dépend d'un paramètre c . Si $A(x, c)$ est continue sur $I \times V$ (I un intervalle ouvert et $V \subset \mathbb{R}^p$ ouvert), alors la résolvante $R(x, x_0, c)$ est une fonction continue sur $I \times I \times V$.*

Démonstration. Pour un (\bar{x}, x_0, c) fixé, il faut démontrer que la différence $R(x, x_1, c_1) - R(\bar{x}, x_0, c)$ est arbitrairement petite si (x, x_1, c_1) est suffisamment proche de (\bar{x}, x_0, c) . Une application du Théorème 5.5 montre que $R(x, x_1, c) - R(\bar{x}, x_0, c)$ est arbitrairement petit. Il suffit alors d'estimer la différence $R(x, x_1, c_1) - R(x, x_1, c)$.

Soit $[a, b] \subset I$ tel que x_0 et \bar{x} sont à l'intérieur de (a, b) , soit $K \subset V$ un voisinage compact de c , et soit L une borne de $\|A(x, c)\|$ sur le compact $[a, b] \times K$. Ceci implique $\|R(t, x_1, c_1)\| \leq e^{L(b-a)}$ pour $t, x_1 \in [a, b]$ et $c_1 \in K$ (voir la démonstration du Théorème 6.1).

La fonction $y(x) := R(x, x_1, c)y_0$ est une solution de $y' = A(x, c)y$, $y(x_1) = y_0$. La fonction $v(x) := R(x, x_1, c_1)y_0$ est une solution approchée, qui satisfait $v(x_1) = y_0$ et

$$\|v'(t) - A(t, c)v(t)\| = \|(A(t, c_1) - A(t, c))R(t, x_1, c_1)y_0\| \leq \varepsilon e^{L(b-a)}\|y_0\|$$

si c_1 est proche de c . Ceci découle de la continuité uniforme de $A(t, c)$ sur $[a, b] \times K$. Le lemme fondamental donne alors l'estimation

$$\left\| \left(R(x, x_1, c) - R(x, x_1, c_1) \right) y_0 \right\| \leq \frac{\varepsilon e^{L(b-a)}}{L} (e^{L|x-x_1|} - 1) \|y_0\|,$$

ce qui démontre l'affirmation. \square

B) Différentiabilité par rapport aux paramètres Considérons un problème avec des paramètres $c = (c_1, \dots, c_p)^T$

$$y' = f(x, y, c), \quad y(x_0) = y_0$$

où $f : U \rightarrow \mathbb{R}^n$ est continue et différentiable par rapport à (y, c) (U est un ouvert de $\mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^p$). La solution est une fonction $y(x, x_0, y_0, c)$. Pour étudier la différentiabilité par rapport à c , nous considérons le système augmenté

$$\begin{pmatrix} y \\ c \end{pmatrix}' = \begin{pmatrix} f(x, y, c) \\ 0 \end{pmatrix}, \quad \begin{pmatrix} y \\ c \end{pmatrix}(x_0) = \begin{pmatrix} y_0 \\ c \end{pmatrix}. \quad (8.6)$$

Avec $z = (y, c)^T$ et $F(x, z) = (f(x, y, c), 0)^T$, ce système devient $z' = F(x, z)$, $z(x_0) = (y_0, c)^T$ et c apparaît seulement dans la valeur initiale. On peut donc appliquer le Théorème 8.1. La fonction $z(x, x_0, z_0)$ et alors aussi $y(x, x_0, y_0, c)$ sont continûment différentiables par rapport à $z_0 = (y_0, c)^T$. En dérivant l'équation

$$\frac{\partial y}{\partial x}(x, x_0, y_0, c) = f(x, y(x, x_0, y_0, c), c)$$

par rapport à c , on obtient pour $\Psi(x) := \frac{\partial y}{\partial c}(x, x_0, y_0, c)$ l'équation différentielle

$$\Psi' = \frac{\partial f}{\partial y}(x, y(x, x_0, y_0, c), c)\Psi + \frac{\partial f}{\partial c}(x, y(x, x_0, y_0, c), c). \quad (8.7)$$

Exemple 8.3 La solution du problème

$$y' = f(x, y) + \varepsilon g(x, y), \quad y(x_0) = y_0$$

est, pour $|\varepsilon|$ petit, donnée par $y(x, \varepsilon) = y_0(x) + \varepsilon y_1(x) + \dots$ où

$$\begin{aligned} y_0'(x) &= f(x, y_0(x)), & y_0(x_0) &= y_0 \\ y_1'(x) &= \frac{\partial f}{\partial y}(x, y_0(x))y_1(x) + g(x, y_0(x)), & y_1(x_0) &= 0. \end{aligned}$$

C) Différentiabilité par rapport à x_0 . Supposons que $f(x, y)$ soit continûment différentiable par rapport à x et y . Nous considérons le système augmenté

$$\begin{pmatrix} y \\ x \end{pmatrix}' = \begin{pmatrix} f(x, y) \\ 1 \end{pmatrix}, \quad \begin{pmatrix} y \\ x \end{pmatrix}(x_0) = \begin{pmatrix} y_0 \\ x_0 \end{pmatrix}. \quad (8.8)$$

Avec $z = (y, x)^T$ et $F(z) = (f(x, y), 1)^T$ le système devient $z' = F(z)$, $z(x_0) = (y_0, x_0)^T$. Donc, x_0 peut être traité de la même manière que y_0 . La solution $y(x, x_0, y_0)$ est alors aussi différentiable par rapport à x_0 . En dérivant (8.3) par rapport à x_0 , on obtient pour $\Psi(x) := \frac{\partial y}{\partial x_0}(x, x_0, y_0)$ l'équation différentielle

$$\Psi' = \frac{\partial f}{\partial y}(x, y(x, x_0, y_0))\Psi, \quad \Psi(x_0) = -f(x_0, y_0). \quad (8.9)$$

Comme $\frac{\partial y}{\partial y_0}(x, x_0, y_0)$ est la résolvante de (8.9), on en déduit que

$$\frac{\partial y}{\partial x_0}(x, x_0, y_0) = -\frac{\partial y}{\partial y_0}(x, x_0, y_0)f(x_0, y_0). \quad (8.10)$$

Cette dernière formule peut aussi être obtenue en dérivant l'identité $y(x, x_0, y(x_0, x_1, y_1)) = y(x, x_1, y_1)$ par rapport à x_0 et en utilisant $y_0 := y(x_0, x, y_1)$. Comme résumé de ce paragraphe nous avons le résultat suivant.

Théorème 8.4 Soit $U \subset \mathbb{R} \times \mathbb{R}^n$ un ouvert et $f : U \rightarrow \mathbb{R}^n$ de classe \mathcal{C}^1 (de classe \mathcal{C}^k). Alors, la solution $y(x, x_0, y_0)$ de (8.1) est de classe \mathcal{C}^1 (de classe \mathcal{C}^k) sur D de (8.2).

Démonstration. Toutes les dérivées partielles de $y(x, x_0, y_0)$ sont continues. Alors, $y : D \rightarrow \mathbb{R}^n$ est de classe \mathcal{C}^1 . En appliquant cet argument itérativement à l'équation variationnelle, on voit par récurrence que y est de classe \mathcal{C}^k si f est de classe \mathcal{C}^k . \square

III.9 Stabilité

Dans le paragraphe précédent nous avons vu que, pour un x fixé, la solution $y(x, x_0, y_0)$ dépend continûment des valeurs initiales. Mais que peut-on dire pour $x \rightarrow \infty$?

Définition 9.1 Soit $U \subset \mathbb{R} \times \mathbb{R}^n$ ouvert et supposons que $f : U \rightarrow \mathbb{R}^n$ soit continue et satisfasse localement une condition de Lipschitz. Une solution $y(x, x_0, y_0)$ de $y' = f(x, y)$ (qui existe pour tout $x \geq x_0$) est dite *stable* (au sens de Liapunov) si

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad \forall \Delta y_0 (\|\Delta y_0\| < \delta) \quad \forall x (x \geq x_0) \quad \|y(x, x_0, y_0 + \Delta y_0) - y(x, x_0, y_0)\| < \varepsilon.$$

La solution est dite *asymptotiquement stable* si elle est stable et si en plus

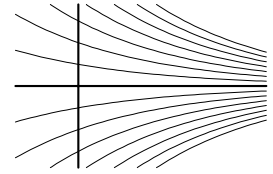
$$\exists \delta_0 > 0 \quad \forall \Delta y_0 (\|\Delta y_0\| < \delta_0) \quad \lim_{x \rightarrow \infty} \|y(x, x_0, y_0 + \Delta y_0) - y(x, x_0, y_0)\| = 0.$$

Une solution est *instable* si elle n'est pas stable.

Exemple 9.2 La solution $y(x) \equiv 0$ de l'équation différentielle

$$y' = \lambda y$$

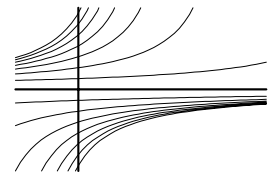
est asymptotiquement stable pour $\lambda < 0$ (voir le dessin ci-contre), stable pour $\lambda \leq 0$ et instable pour $\lambda > 0$.



Exemple 9.3 La solution $y(x, 0, y_0) = y_0/(1 - y_0 x)$ de l'équation différentielle

$$y' = y^2$$

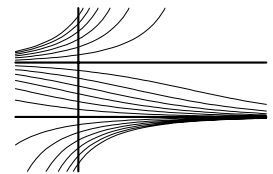
est asymptotiquement stable pour $y_0 < 0$, mais instable pour $y_0 = 0$.



Exemple 9.4 Pour l'équation différentielle

$$y' = y(y - 1)$$

la solution $y(x) \equiv 0$ est asymptotiquement stable, mais la solution $y(x) \equiv 1$ est instable.



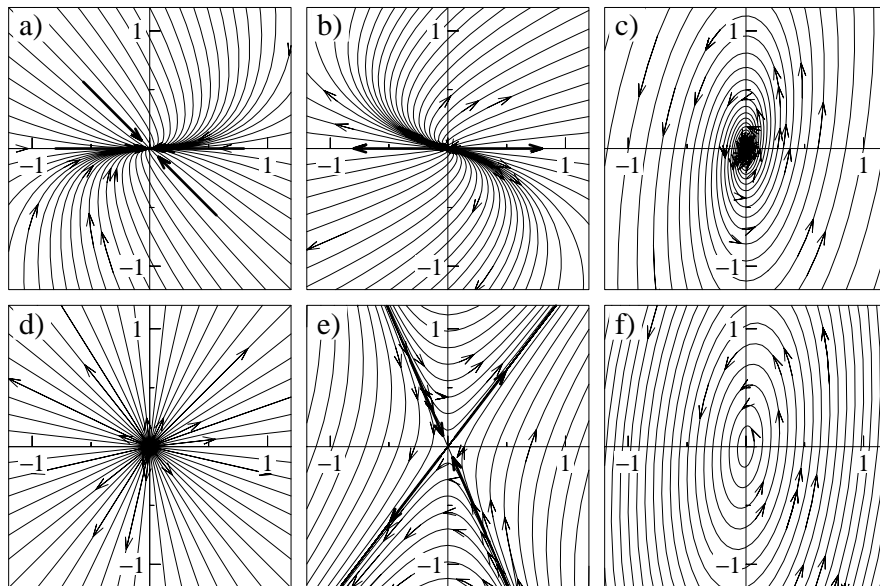
Pour étudier la stabilité d'une solution $y(x)$ de $y' = f(x, y)$, il est souvent avantageux d'appliquer la transformation $z = y - y(x)$. Ainsi, l'équation différentielle $y' = f(x, y)$ se transforme en

$$z' = g(x, z), \quad g(x, z) = f(x, z + y(x)) - f(x, y(x)) \quad (9.1)$$

avec $z(x) \equiv 0$ comme solution.

Lemme 9.5 Sous les hypothèses de la Définition 9.1, la solution $y(x)$ de $y' = f(x, y)$ est stable, si et seulement si $z(x) \equiv 0$ est une solution stable de (9.1). Dans cette affirmation, on peut remplacer 'stable' par 'asymptotiquement stable' ou par 'instable'.

Démonstration. L'affirmation est une conséquence immédiate du fait que $z(x, x_0, z_0) = y(x, x_0, y_0) - y(x)$ avec pour $z_0 = y_0 - y(x_0)$. \square



$$\begin{array}{lll}
 \text{a)} & \begin{pmatrix} -1 & 1 \\ 0 & -2 \end{pmatrix} & \text{b)} & \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} & \text{c)} & \begin{pmatrix} 1/3 & -1/3 \\ 2 & 0 \end{pmatrix} \\
 \text{d)} & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \text{e)} & \begin{pmatrix} 1/3 & 1/3 \\ 1 & 0 \end{pmatrix} & \text{f)} & \begin{pmatrix} 1/6 & -1/3 \\ 2 & -1/6 \end{pmatrix}
 \end{array}$$

FIG. 9.1 – Solutions de systèmes linéaires en dimension 2

Par conséquent, on peut toujours se ramener au cas où $f(x, 0) \equiv 0$ et où on est intéressé par la stabilité de la solution nulle $y(x) \equiv 0$. Un désavantage potentiel de la transformation (9.1) est qu'un système autonome $y' = f(y)$ peut devenir non-autonome.

Pour une équation linéaire $y' = A(x)y + g(x)$ la transformation (9.1) donne $z' = A(x)z$ et on obtient: une solution arbitraire de $y' = A(x)y + g(x)$ est stable (asymptotiquement stable ou instable) si la solution nulle de l'équation homogène $y' = A(x)y$ est stable (asymptotiquement stable ou instable). Pour une équation linéaire, toutes les solutions possèdent le même comportement de stabilité. Dans cette situation, on parle aussi de la stabilité de l'équation différentielle et non de la solution.

Théorème 9.6 a) L'équation différentielle $y' = Ay$ est asymptotiquement stable si et seulement si toutes les valeurs propres de A satisfont $\Re\lambda_i < 0$.

b) Elle est stable si et seulement si toutes les valeurs propres de A satisfont $\Re\lambda_i \leq 0$ et, pour les valeurs propres avec $\Re\lambda_i = 0$, les blocs de Jordan sont de dimension 1.

Démonstration. On transforme la matrice A sous forme de Jordan et on utilise la représentation explicite de la solution (Exemple 7.1). L'affirmation est une conséquence du fait que $|e^{\lambda x}|$ reste borné pour $x \geq x_0$ si $\Re\lambda \leq 0$, et que $x^k e^{\lambda x} \rightarrow 0$ pour $x \rightarrow \infty$ si $\Re\lambda < 0$. \square

Le comportement des solutions de $y' = Ay$ (pour des matrices de dimension 2) est illustré dans la Fig. 9.1. Pour le dessin (c) les valeurs propres sont complexes conjuguées avec $\Re\lambda_i > 0$, pour (e) les valeurs propres possèdent un signe opposé et pour (f) les deux valeurs propres sont sur l'axe imaginaire. Dans les dessins (a), (b) et (e) les directions des vecteurs propres sont indiquées par des fleches.

Le critère de stabilité du Théorème 9.6 ne reste pas vrai pour des équations différentielles non-autonomes $y' = A(x)y$.

Exemple 9.7 Considérons l'équation différentielle

$$y' = A(x)y, \quad A(x) = T(x)^{-1}BT(x), \quad (9.2)$$

où

$$B = \begin{pmatrix} -1 & 0 \\ 4 & -1 \end{pmatrix}, \quad T(x) = \begin{pmatrix} \cos(ax) & -\sin(ax) \\ \sin(ax) & \cos(ax) \end{pmatrix}.$$

Les valeurs propres de $A(x)$ sont réelles et négatives (égales à -1). Nous allons déterminer le paramètre a dans la matrice de rotation $T(x)$ afin que (9.2) devienne instable. Avec la transformation $z(x) := T(x)y(x)$ nous obtenons

$$z'(x) = T'(x)y(x) + T(x)A(x)y(x) = \left(\begin{pmatrix} 0 & -a \\ a & 0 \end{pmatrix} + \begin{pmatrix} -1 & 0 \\ 4 & -1 \end{pmatrix} \right) z(x),$$

une équation linéaire à coefficients constants. Pour $a = -2$, la matrice de l'équation différentielle pour $z(x)$ possède $\lambda^2 + 2\lambda - 3$ comme polynôme caractéristique. Donc, $\lambda_1 = 1$ est une valeur propre avec $(1, 1)^T$ comme vecteur propre. Une solution non-bornée de (9.2) est alors donnée par

$$y(x) = e^x \begin{pmatrix} \cos 2x + \sin 2x \\ -\sin 2x + \cos 2x \end{pmatrix}.$$

L'équation différentielle (9.2) est instable, tandis que toutes les valeurs propres de $A(x)$ possèdent une partie réelle négative.

Théorème 9.8 Soit $A(x)$ une matrice symétrique et continue sur $[a, \infty)$. Si les valeurs propres $\lambda_i(x)$ de $A(x)$ satisfont

$$\lambda_i(x) \leq \alpha \quad \text{pour } i = 1, \dots, n \text{ et pour } x \in [a, \infty),$$

alors on a l'estimation

$$\|y(x)\|_2 \leq e^{\alpha(x-x_0)} \|y(x_0)\|_2 \quad \text{pour } x \geq x_0$$

pour les solutions de $y' = A(x)y$. Le système $y' = A(x)y$ est donc stable si $\alpha \leq 0$ et asymptotiquement stable si $\alpha < 0$.

Démonstration. Montrons d'abord que les hypothèses sur $A(x)$ impliquent

$$v^T A(x)v \leq \alpha v^T v \quad \text{pour } v \in \mathbb{R}^n.$$

Comme $A(x)$ est symétrique, elle est diagonalisable à l'aide d'une matrice orthogonale $Q(x)$, c.-à-d., $Q(x)^T A(x)Q(x) = \Lambda(x) = \text{diag}(\lambda_1(x), \dots, \lambda_n(x))$. Avec $v = Q(x)w$ on a alors que $v^T A(x)v = w^T \Lambda(x)w \leq \alpha w^T w = \alpha v^T v$.

Pour une solution $y(x)$ de $y' = A(x)y$, considérons la fonction $m(x) := \|y(x)\|_2^2$. Sa dérivée satisfait

$$m'(x) = 2y(x)^T y'(x) = 2y(x)^T A(x)y(x) \leq 2\alpha y(x)^T y(x) = 2\alpha m(x).$$

En vue d'une application du Théorème 5.2, nous comparons $m(x)$ avec la solution de $u' = 2\alpha u$, $u(x_0) = m(x_0)$ (plus précisément avec la solution de $u' = 2\alpha u + \varepsilon$ où $\varepsilon > 0$ est arbitrairement petit) et nous obtenons $m(x) \leq m(x_0)e^{2\alpha(x-x_0)}$ pour $x \geq x_0$. En prenant la racine carrée, nous arrivons au résultat désiré. \square

Considérons un problème nonlinéaire $y' = f(x, y)$ satisfaisant $f(x, 0) = 0$ (voir la remarque après le Lemme 9.5). Une linéarisation autour de $y = 0$ donne

$$y' = A(x)y + g(x, y), \quad (9.3)$$

où $g(x, y)$ est une petite perturbation nonlinéaire.

Théorème 9.9 *Soit $A(x)$ continue sur (a, ∞) , $g : (a, \infty) \times U \rightarrow \mathbb{R}^n$ continue et localement Lipschitzienne où U est un voisinage ouvert de $0 \in \mathbb{R}^n$. Supposons que*

- la résolvante de $y' = A(x)y$ satisfasse $\|R(x, x_0)\| \leq Ke^{-\alpha(x-x_0)}$ pour $x \geq x_0$ avec un $\alpha > 0$,
- $\sup_{x>a} \frac{\|g(x, y)\|}{\|y\|} \rightarrow 0$ si $y \rightarrow 0$ (uniformément en x),

alors la solution nulle de (9.3) est asymptotiquement stable.

Démonstration. Les solutions de (9.3) satisfont

$$y(x) = R(x, x_0)y_0 + \int_{x_0}^x R(x, t)g(t, y(t)) dt \quad (9.4)$$

(ceci se voit directement par différentiation de (9.4)). On en déduit l'estimation

$$\|y(x)\| \leq Ke^{-\alpha(x-x_0)}\|y_0\| + \int_{x_0}^x Ke^{-\alpha(x-t)}\|g(t, y(t))\| dt.$$

Pour $\gamma := \alpha/(2K)$ il existe un $\varepsilon > 0$ tel que $\|g(x, y)\| \leq \gamma\|y\|$ pour $\|y\| \leq \varepsilon$ et pour tout $x > a$. Tant que la solution $y(x)$ reste dans $\{y; \|y\| \leq \varepsilon\}$, on a

$$\|y(x)\| \leq Ke^{-\alpha(x-x_0)}\|y_0\| + K\gamma \int_{x_0}^x e^{-\alpha(x-t)}\|y(t)\| dt. \quad (9.5)$$

En notant l'expression de droite par $z(x)$ et en dérivant, on obtient

$$z'(x) = -\alpha z(x) + K\gamma\|y(x)\| \leq -\frac{\alpha}{2} z(x), \quad z(x_0) = K\|y_0\|,$$

car $\|y(x)\| \leq z(x)$. En résolvant cette inégalité différentielle, on obtient

$$\|y(x)\| \leq z(x) \leq K\|y_0\|e^{-\alpha(x-x_0)/2}. \quad (9.6)$$

Posons $\delta := \varepsilon/K$. Pour y_0 satisfaisant $\|y_0\| \leq \delta$, l'estimation (9.6) montre que $y(x)$ reste dans $\{y; \|y\| \leq \varepsilon\}$ pour tout $x \geq x_0$ et que $y(x) \rightarrow 0$ pour $x \rightarrow \infty$. La solution nulle de (9.3) est donc asymptotiquement stable. \square

Théorème 9.10 *Soit A une matrice constante, et soit $g : (a, \infty) \times U \rightarrow \mathbb{R}^n$ continue et localement Lipschitzienne (U un voisinage ouvert de $0 \in \mathbb{R}^n$). Supposons que*

- au moins une valeur propre de A satisfait $\Re\lambda > 0$,
- $\sup_{x>a} \frac{\|g(x, y)\|}{\|y\|} \rightarrow 0$ si $y \rightarrow 0$ (uniformément en x),

alors la solution nulle de $y' = Ay + g(x, y)$ est instable.

Démonstration. a) Supposons pour le moment que A soit diagonalisable, $T^{-1}AT = \Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$, et que les valeurs propres de A soient ordonnées comme

$$\Re\lambda_1 \geq \dots \geq \Re\lambda_s > 0 \geq \Re\lambda_{s+1} \geq \dots \geq \Re\lambda_n.$$

Après le changement de coordonnées $y = Tz$, l'équation différentielle devient

$$z' = \Lambda z + \hat{g}(x, z), \quad \hat{g}(x, z) = T^{-1}g(x, Tz). \quad (9.7)$$

L'estimation

$$\frac{\|\hat{g}(x, z)\|}{\|z\|} = \frac{\|Tz\|}{\|z\|} \cdot \frac{\|T^{-1}g(x, Tz)\|}{\|Tz\|} \leq \|T\| \cdot \|T^{-1}\| \cdot \frac{\|g(x, Tz)\|}{\|Tz\|}$$

et l'hypothèse sur g impliquent que $\sup_{x>a} \|\hat{g}(x, z)\|/\|z\| \rightarrow 0$ si $z \rightarrow 0$. Pour $c := \Re\lambda_s > 0$, il existe alors un $\varepsilon > 0$ tel que

$$\|\hat{g}(x, z)\| \leq \frac{c}{2} \|z\| \quad \text{pour} \quad \|z\| \leq \varepsilon. \quad (9.8)$$

b) On cherche à démontrer que la solution nulle de $y' = Ay + g(x, y)$ est instable. Supposons, par l'absurde, qu'elle soit stable. Ceci implique que la solution $z(x) \equiv 0$ de (9.7) est aussi stable et que, par conséquence, il existe un $\delta > 0$ tel que

$$\|z_0\| \leq \delta \quad \implies \quad \|z(x)\| \leq \varepsilon \quad \text{pour} \quad x \geq x_0.$$

Nous écrivons maintenant le vecteur z comme $z = (u, v)^T$ où $u \in \mathbb{C}^s$ correspond aux valeurs propres avec partie réelle positive, et $v \in \mathbb{C}^{n-s}$. Nous choisissons une valeur initiale $z_0 = (u_0, v_0)^T$ satisfaisant $\|z_0\| \leq \delta$ et $\|u_0\| > \|v_0\|$. Pour la solution $z(x) = (u(x), v(x))^T$ nous considérons alors l'expression

$$d(x) = \frac{1}{2} (\|u(x)\|^2 - \|v(x)\|^2),$$

dont la dérivée satisfait

$$\begin{aligned} d'(x) &= \Re\langle u'(x), u(x) \rangle - \Re\langle v'(x), v(x) \rangle \\ &= \overline{u(x)}^T \Lambda_1 u(x) + \Re\langle \hat{g}_1(x, z(x)), u(x) \rangle - \overline{v(x)}^T \Lambda_2 v(x) - \Re\langle \hat{g}_2(x, z(x)), v(x) \rangle, \end{aligned}$$

où $\Lambda_1 = \text{diag}(\Re\lambda_1, \dots, \Re\lambda_s)$, $\Lambda_2 = \text{diag}(\Re\lambda_{s+1}, \dots, \Re\lambda_n)$ et $\hat{g}(x, z) = (\hat{g}_1(x, z), \hat{g}_2(x, z))^T$. L'inégalité de Cauchy-Schwarz et l'estimation (9.8) impliquent que

$$d'(x) \geq c \|u(x)\|^2 - \frac{c}{2} (\|u(x)\|^2 + \|v(x)\|^2) = c d(x).$$

En résolvant cette inégalité différentielle pour $d(x)$ et en utilisant $d(x_0) > 0$, nous obtenons

$$d(x) \geq d(x_0) e^{c(x-x_0)}.$$

Par conséquent, $d(x) \rightarrow \infty$ pour $x \rightarrow \infty$, ce qui contredit l'hypothèse que $z(x) \equiv 0$ est une solution stable de (9.7).

c) Dans la situation où A n'est pas diagonalisable, nous transformons A sous forme de Jordan, et nous faisons une transformation additionnelle avec $D = \text{diag}(1, \alpha, \alpha^2, \alpha^3, \dots)$ pour rendre les éléments hors-diagonaux aussi petits qu'on veut. Finalement, nous ajoutons ces termes à la fonction $g(x, y)$ et nous procédons comme sous (a) et (b). \square

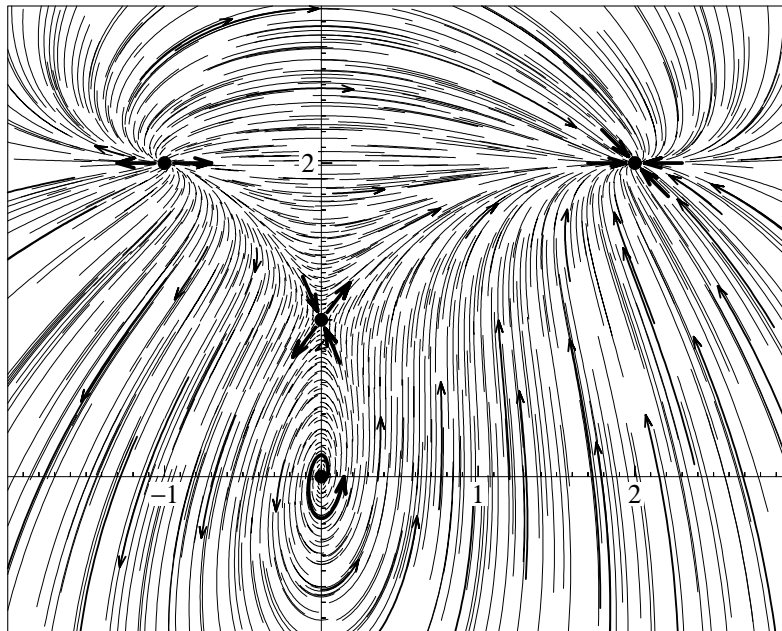


FIG. 9.2 – Solutions d'un système nonlinéaire avec 4 points critiques

Corollaire 9.11 Soit $U \subset \mathbb{R}^n$ un ouvert, soit $f : U \rightarrow \mathbb{R}^n$ continûment différentiable, et soit $y_0 \in U$ un point stationnaire, c.-à-d., $f(y_0) = 0$.

- (a) Si toutes les valeurs propres de $f'(y_0)$ satisfont $\Re \lambda < 0$, alors la solution stationnaire $y(x) \equiv y_0$ est asymptotiquement stable (on dit que y_0 est un attracteur).
- (b) Si au moins une valeur propre de $f'(y_0)$ satisfait $\Re \lambda > 0$, alors la solution stationnaire est instable.
- (c) Si les valeurs propres de $f'(y_0)$ satisfont $\Re \lambda \leq 0$, mais au moins une valeur propre satisfait $\Re \lambda = 0$, on ne peut rien dire. La solution stationnaire peut être stable ou bien instable.

Démonstration. La transformation $z = y - y_0$ donne $z' = f(z + y_0) = f'(y_0)z + g(z)$ où $\|g(z)\|/\|z\| \rightarrow 0$ si $z \rightarrow 0$. Les affirmations (a) et (b) sont alors une conséquence du Théorème 9.9 et du Théorème 9.10.

Pour les problèmes $y' = y^2$ et $y' = y^3$, on a $f'(0) = 0$ et la solution nulle est instable pour $y' = y^2$ (Exemple 9.3), mais asymptotiquement stable pour $y' = y^3$. Ceci se voit à partir de la forme analytique des solutions. \square

Exemple 9.12 Le système

$$\begin{aligned} y_1' &= (y_1 - y_2)(1 - y_1 - y_2)/3 \\ y_2' &= y_1(2 - y_2) \end{aligned} \tag{9.9}$$

possède 4 points critiques (c.-à-d., des points où $f(y) = 0$, notamment $(0, 0)$, $(0, 1)$, $(2, 2)$ et $(-1, 2)$). Seulement le point $(2, 2)$ est asymptotiquement stable (voir Fig.9.2).

III.10 Exercices

1. Déterminer le type des équations différentielles suivantes parmi (I) séparation des variables, (II) linéaire homogène, (III) linéaire inhomogène, (IV) Bernoulli et (V) substitution de variable, sans les résoudre. Si une équation est de type (V), donner la substitution. *Attention*, il peut y avoir plusieurs solutions pour une équation.

	I	II	III	IV	V
1. $u' \cos(u) = t$	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
2. $u' + u = u^3$	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
3. $u' = 2 + u/t$	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
4. $u' + u = e^{2t}$	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
5. $tu' = u + \sqrt{t^2 + u^2}$	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
6. $u' + u/t + t^3u^4 = 0$	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
7. $t^2u' - u^2 = 1$	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
8. $u' - 13u - 4ut = 2u^2 + 2t^2 + 5 + 13t$	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
9. $(1 - t^2)u' - 2tu = t^2$	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
10. $2tu' + u + 3t^2u^2 = 0$	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
11. $(t \cos(t))u' + (\cos(t) + t \sin(t))u = 1$	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
12. $-1/u^2 + 1/u - 2 = 0$	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
13. $u' = u - 2u^2$	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
14. $u + u/t - t \log t = 2(c \log(t^{1/2})u^2 + tuc \log(t))$	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
15. $u + u'(t - u) = 0$	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
16. $u' = u - 1/u$	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

2. Déterminer le type des équations différentielles suivantes et les résoudre:

$$\begin{aligned} (x+1)y' &= -xy, & x^2y' + xy + \sqrt{|y|} &= 0, \\ xy' &= \sqrt{x^2+y^2} + y, & (x+y-2) + (x-y+4)y' &= 0. \end{aligned}$$

Indication. Pour la dernière équation, chercher une fonction $U(x, y)$ telle que $\frac{\partial U}{\partial x}(x, y) = x + y - 2$ et $\frac{\partial U}{\partial y}(x, y) = x - y + 4$.

3. Trouver la solution de

$$y' - 2xy = \cos x - 2x \sin x$$

qui est bornée pour $x \rightarrow \infty$.

4. Considérons l'équation de Clairaut

$$y - xy' + f(y') = 0.$$

Montrer que l'enveloppe de la famille de solutions $y = Cx - f(C)$ est donnée par

$$x(t) = f'(t), \quad y(t) = tf'(t) - f(t).$$

Montrer également que cette enveloppe est solution de l'équation de Clairaut (dans un voisinage de $x(t)$ où $x'(t) \neq 0$).

Pour $f(t) = 5(t^3 - t)/2$, exprimer l'enveloppe sous la forme $y(x)$ (voir Fig. 1.1).

5. Un chien court à une vitesse w en direction de sa maîtresse, qui se promène à une vitesse v (sur l'axe y). Montrer que ceci conduit à l'équation différentielle

$$xy'' = \lambda \sqrt{1 + y'^2}, \quad \lambda = v/w < 1,$$

et la résoudre.

6. Appliquer l'itération de Picard-Lindelöf à

$$y' = \frac{1}{\sqrt{x}} + y^2, \quad y(0) = 0.$$

7. Appliquer l'itération de Picard-Lindelöf au problème

$$\begin{aligned} y_1' &= y_1 + 2y_2, & y_1(0) &= 0 \\ y_2' &= x^2 - y_1, & y_2(0) &= 0 \end{aligned}$$

et calculer les 5 premiers termes de la série de Taylor de la solution.

8. On suppose que $f(x, y)$ est continue et satisfait une condition de Lipschitz sur la bande infinie $D = \{(x, y); a \leq x \leq b\}$. Montrer que le problème $y' = f(x, y)$, $y(x_0) = y_0$ avec $x_0 \in [a, b]$ possède une solution unique sur tout l'intervalle $[a, b]$.

Indication. Suivre la démonstration du Théorème 2.3.

9. Sous les hypothèses de l'exercice précédent, on considère l'opérateur $T : \mathcal{C}([a, b]) \rightarrow \mathcal{C}([a, b])$ donné par

$$(Ty)(x) = y_0 + \int_{x_0}^x f(t, y(t)) dt .$$

- (a) Montrer que, pour la norme $\|y\|_\infty = \max\{|y(x)|; x \in [a, b]\}$, l'opérateur T n'est pas toujours une contraction.
 (b) Trouver $\gamma \in \mathbb{R}$ tel que T soit une contraction pour la norme

$$\|y\| = \max\{|y(x)|e^{-\gamma|x-x_0|}; x \in [a, b]\} .$$

10. Considérons le problème $y' = f(x, y)$, $y(0) = 0$, où $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ est donnée par

$$f(x, y) = \begin{cases} 0 & \text{si } x \leq 0, \quad y \in \mathbb{R} \\ 2x & \text{si } x > 0, \quad y \leq 0 \\ 2x - \frac{4y}{x} & \text{si } x > 0, \quad 0 < y < x^2 \\ -2x & \text{si } x > 0, \quad x^2 \leq y. \end{cases}$$

- (a) Montrer que f est continue. Que signifie ceci pour le problème?
 (b) Montrer que la fonction f ne satisfait une condition de Lipschitz dans aucun voisinage de l'origine.
 (c) Appliquer l'itération de Picard-Lindelöf avec $y_0(x) \equiv 0$. Les points d'accumulation de $\{y_k(x)\}$ sont-ils des solutions?
 (d) Montrer que le problème possède une unique solution. Laquelle?
11. Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ de classe \mathcal{C}^1 et soit l'équation différentielle

$$y'' = f(y, y') .$$

Si $f(0, 0) = 0$, montrer que toute solution non nulle de cette équation n'a que des zéros simples. Exemples: $y'' + y = 0$, $y'' + \sin y = 0$.

12. Calculer l'intervalle maximal d'existence, $I_{\max}(x_0, y_0)$, pour une solution unique de

$$y' = 1 - y^2 \quad \text{et} \quad y' = y^2 .$$

Pour le premier problème, trouver la réponse sans calculer explicitement la solution.

13. Soit $y' = f(y)$ une équation autonome, où $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ satisfait localement une condition de Lipschitz. Montrer que si pour une solution $y : I \rightarrow \mathbb{R}^n$, il existe $t_1, t_2 \in I$, $t_1 \neq t_2$ avec $y(t_1) = y(t_2)$, alors $I_{\max} = \mathbb{R}$ et la solution est soit constante soit périodique.
 14. Montrer l'affirmation suivante: Si le problème $y' = f(x, y)$, $y(x_0) = y_0$, où $f(x, y)$ est une fonction continue, possède une solution unique, alors les polygones d'Euler convergent vers cette solution.
 15. On considère le problème $y' = f(x, y)$, $y(x_0) = 0$, où

$$f(x, y) = \begin{cases} 4 \operatorname{sign}(y) \sqrt{|y|} & \text{si } |y| \geq x^2 \\ 4 \operatorname{sign}(y) \sqrt{|y|} + 4(x - \frac{|y|}{x}) \cos\left(\frac{\pi \log x}{\log 2}\right) & \text{si } |y| < x^2 \end{cases} ,$$

(voir la Fig. I.7.2 de [HNW93]). La fonction $f(x, y)$ est continue sur \mathbb{R}^2 .

- (a) Montrer que la suite des polygones d'Euler $\{y_h(x)\}$ ne converge pas pour $h \rightarrow 0$.
 (b) Calculer les points d'accumulation de la suite $\{y_h(x)\}$ où $h = 2^{-i}$, $i \geq 1$, et montrer qu'ils sont des solutions du problème de Cauchy.

Indication. Les solutions sont $\pm 4x^2$.

16. Résoudre l'équation

$$x' = -y + x(\epsilon - (x^2 + y^2))$$

$$y' = x + y(\epsilon - (x^2 + y^2))$$

en passant en coordonnées polaires. Discuter de l'allure des solutions en fonction de $\epsilon \in \mathbb{R}$.

Remarque. Si $\epsilon > 0$ on a une solution qui est un cercle de rayon $\sqrt{\epsilon}$, puis lorsque ϵ devient négatif, ce cercle disparaît, c'est une Bifurcation de Hopf (Fig. 10.1).

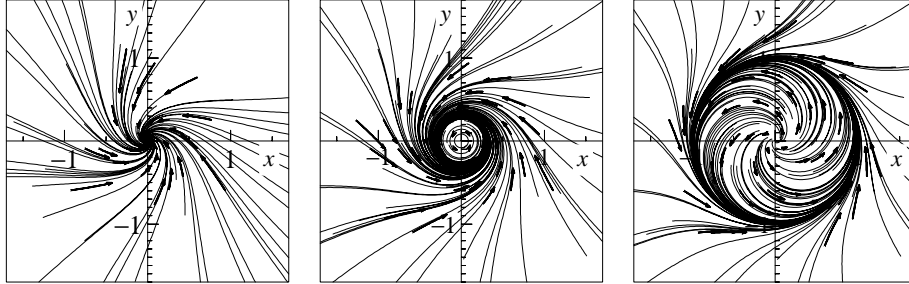


FIG. 10.1 – Comportement des solutions pour $\epsilon = -1$, $\epsilon = 0$ et $\epsilon = 1$.

17. Pour la solution de l'équation de Ricatti

$$y' = x^2 + y^2, \quad y(0) = 1,$$

démontrer que

$$\frac{1}{1-x} \leq y(x) \leq \tan\left(x + \frac{\pi}{4}\right).$$

Indication. Résoudre $u' = u^2 - \epsilon^2$ ($\epsilon \rightarrow 0$) et $v' = v^2 + 1$.

18. On considère le problème

$$y' = \lambda y, \quad y(0) = 1$$

sur l'intervalle $[0, 1]$ avec $\lambda > 0$. Démontrer que les polygones d'Euler $y_h(x)$ satisfont

$$\frac{\lambda}{1 + \lambda h} y_h(x) \leq D_+ y_h(x) \leq \lambda y_h(x).$$

En déduire que

$$\left(1 + \frac{\lambda}{n}\right)^n \leq e^\lambda \leq \left(1 + \frac{\lambda}{n}\right)^{n+\lambda}.$$

19. Démontrer le "lemme de Gronwall": S'il existe $L > 0$ tel que

$$m(x) \leq \rho + \epsilon(x - x_0) + L \int_{x_0}^x m(t) dt \quad \text{pour } x \geq x_0,$$

alors

$$m(x) \leq \rho e^{L(x-x_0)} + \frac{\epsilon}{L}(e^{L(x-x_0)} - 1).$$

Indication. Dériver une inégalité différentielle pour $u(x) := \rho + \epsilon(x - x_0) + L \int_{x_0}^x m(t) dt$.

20. Démontrer par un contre-exemple que l'affirmation (c.f. Théorème 5.2)

$$\left. \begin{array}{l} m(x_0) \leq u(x_0) \\ D_+ m(x) \leq g(x, m(x)) \\ D_+ u(x) \geq g(x, u(x)) \end{array} \right\} \implies m(x) \leq u(x)$$

n'est pas vraie sans hypothèses additionnelles (voir l'Exemple 3.1).

21. On considère le problème

$$y' = 1 - \cos(xy), \quad y(0) = 0.1. \tag{10.1}$$

(a) Montrer que (10.1) possède une solution unique sur $[0, \infty)$.

- (b) Calculer une solution approchée en remplaçant $\cos(xy)$ par $1 - (xy)^2/2$.
- (c) Estimer la différence entre la solution exacte et la solution approchée sur l'intervalle $[0, 1]$. La Fig. 10.2 illustre la solution exacte et approchée, on a une singularité pour la solution approchée en $x = 60^{1/3}$.

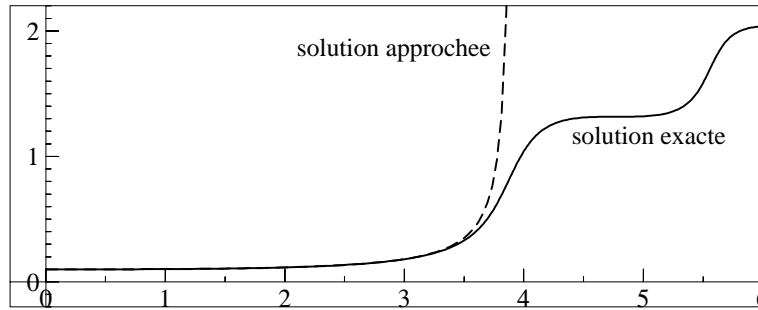


FIG. 10.2 – Allure de la solution exacte et de la solution approchée

22. Démontrer que si $A(x)$ est une matrice antisymétrique ($A^T = -A$), alors la résolvante de

$$y' = A(x)y$$

est une matrice orthogonale.

Indication. Montrer que le produit scalaire de deux solutions est constant.

23. Calculer la résolvante (représentation réelle) de l'équation différentielle

$$\begin{aligned} y_1' &= \cos(x)y_1 - \sin(x)y_2 \\ y_2' &= \sin(x)y_1 + \cos(x)y_2. \end{aligned}$$

Indication. Trouver une équation différentielle pour $z = y_1 + iy_2$.

24. Soit

$$y' = A(x)y \tag{10.2}$$

où $A(x)$ est une matrice d'ordre n . Montrer que si l'on connaît une solution non-triviale $\varphi(x)$ de (10.2), on peut réduire l'équation (10.2) à un problème analogue mais où la matrice est d'ordre $n - 1$.

Indication. Si l'on suppose que $\varphi_1(x) \neq 0$, chercher une solution de la forme $y = u\varphi(x) + z$, où u est une fonction scalaire et où z est de la forme $z = (0, z_2, \dots, z_n)^T$.

25. A l'aide de l'exercice précédent, calculer la résolvante $R(x, 1)$ de l'équation

$$y' = \begin{pmatrix} 1/x & -1 \\ 1/x^2 & 2/x \end{pmatrix} y$$

sachant qu'elle admet $y(x) = (x^2, -x)^T$ comme solution.

Remarque. Le résultat est $\begin{pmatrix} x^2(1 - \ln x) & -x^2 \ln x \\ x \ln x & x(1 + \ln x) \end{pmatrix}$.

26. A l'aide de l'exercice précédent, résoudre

$$y' = \begin{pmatrix} 1/x & -1 \\ 1/x^2 & 2/x \end{pmatrix} y + \begin{pmatrix} 1 \\ x^2 \end{pmatrix}, \quad y(1) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

27. Calculer la solution générale de

$$y' = \begin{pmatrix} 1 & -2 & 0 \\ 2 & 0 & -1 \\ 4 & -2 & -1 \end{pmatrix} y.$$

Indication. $\lambda_1 = 1$ est une valeur propre.

28. Transformer la matrice

$$A = \frac{1}{9} \begin{pmatrix} 14 & 4 & 2 \\ -2 & 20 & 1 \\ -4 & 4 & 20 \end{pmatrix}$$

sous forme normale de Jordan et calculer la résolvante de $y' = Ay$.

Indication. Toutes les valeurs propres de A sont égales à 2. La matrice T qui transforme A sous forme de Jordan est donnée par $T = (v_1, v_2, v_3)$ où $v_3 \neq 0$ est arbitraire, $v_2 = (A - 2I)v_3$ et v_1 est un vecteur propre de A différent de v_2 .

29. Supposons que la matrice $A(x)$ soit continue sur l'intervalle I . Montrer que si les matrices $A(x)$ et $\int_{x_0}^x A(t)dt$ commutent pour tout $x_0, x \in I$, alors la résolvante de $y' = A(x)y$ satisfait

$$R(x, x_0) = \exp \left(\int_{x_0}^x A(t)dt \right) .$$

30. On considère le système

$$y' = \begin{pmatrix} x & 1 \\ 0 & 0 \end{pmatrix} y . \quad (10.3)$$

(a) Montrer que $A(x)$ et $\int_{x_0}^x A(t)dt$ ne commutent pas.

(b) Calculer la résolvante de (10.3).

(c) Montrer que dans ce cas particulier $R(x, x_0) \neq \exp(\int_{x_0}^x A(t)dt)$.

31. On suppose que $A(x)$ soit périodique de période p , c.à.d. $A(x+p) = A(x)$.

(a) Montrer que si $\Phi(x)$ est une matrice fondamentale, il existe une matrice C telle que $\Phi(x+kp) = \Phi(x)C^k$ pour $k = 0, 1, 2, \dots$

(b) Montrer que les valeurs propres de C ne dépendent pas du choix de $\Phi(x)$. Les valeurs propres s'appellent *multiplicateurs de Floquet*.

(c) Montrer que pour chaque valeur propre λ de C , il existe une solution $y(x)$ de $y' = A(x)y$ qui vérifie $y(x+p) = \lambda y(x)$ pour tout x (solutions *quasi-périodiques*).

32. Soit le système

$$\begin{aligned} y' &= y(1-y) + z \\ z' &= z(1-z) . \end{aligned}$$

Pout $x_0 = 0, y_0 = 0, z_0 = 0$ calculer les quantités suivantes:

$$\frac{\partial y}{\partial y_0}(x, x_0, y_0, z_0), \quad \frac{\partial y}{\partial z_0}(x, x_0, y_0, z_0), \quad \frac{\partial z}{\partial y_0}(x, x_0, y_0, z_0), \quad \frac{\partial z}{\partial z_0}(x, x_0, y_0, z_0).$$

33. Soit l'équation différentielle

$$y' = \alpha y + \beta ,$$

avec $\alpha, \beta \in \mathbb{R}$. Elle admet la solution $y(x) = -\beta/\alpha$. Malgré la régularité de la fonction $f(x, y) = \alpha y + \beta$, cette solution est discontinue en $\alpha = 0$. Comment réconcilier ce fait avec le théorème sur la dérivabilité par rapport aux paramètres?

34. Considérons l'équation du pendule

$$y'' + \sin y = 0 , \quad y(0) = \epsilon , \quad y'(0) = 0 ,$$

où l'amplitude ϵ est supposée petite. Montrer que la solution peut être écrite sous la forme

$$y(x) = \epsilon y_1(x) + \epsilon^2 y_2(x) + \epsilon^3 y_3(x) + O(\epsilon^4) .$$

Calculer $y_1(x), y_2(x), y_3(x)$.

35. Calculer la résolvante $R(x, x_0)$ de l'équation

$$y' = \begin{pmatrix} 2/x & 1 \\ 3/x^2 & 3/x \end{pmatrix} y$$

pour $x, x_0 > 0$. Utiliser la transformation $t = \ln(x)$, $z_i(t) = x^{i-1}y_i(x)$.

Donner l'allure générale des systèmes différentiels de dimension ≥ 2 que l'on peut résoudre à l'aide de cette transformation.

36. Montrer que l'équation ($a_{ij} \in \mathbb{R}$)

$$y' = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} y$$

est asymptotiquement stable si et seulement si

$$a_{11} + a_{22} < 0 \quad \text{et} \quad a_{11}a_{22} - a_{12}a_{21} > 0 .$$

37. Montrer que pour un système autonome $y' = f(y)$ le flot

$$\Phi_t : \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad \Phi_t(y_0) = y(t, 0, y_0)$$

satisfait

$$\Phi_t \circ \Phi_s = \Phi_{t+s} .$$

38. Soit $y' = Ay$ où A est une matrice $n \times n$ avec des coefficients réels. Soit

$$\chi_A(\lambda) = \lambda^n + a_1\lambda^{n-1} + \cdots + a_{n-1}\lambda + a_n$$

le polynôme caractéristique de A . Montrer qu'une condition nécessaire pour la stabilité asymptotique de $y' = Ay$ est donnée par

$$a_1 > 0, a_2 > 0, \dots, a_n > 0 .$$

39. On considère le polynôme $p(\lambda) = \lambda^3 + a_1\lambda^2 + a_2\lambda + a_3$ à coefficients réels. Montrer que si

$$a_1 > 0, \quad a_3 > 0, \quad a_1a_2 > a_3$$

alors les racines λ_i du polynôme $p(\lambda)$ satisfont $\Re \lambda_i < 0$.

Indication. Etudier la condition sur les coefficients quand une racine de $p(\lambda)$ traverse l'axe imaginaire.

Remarque. Les conditions mentionnées sont également nécessaires, c'est un cas particulier du critère de Routh-Hurwitz.

40. Aider James Watt à résoudre son problème de stabilité pour son moteur à vapeur de locomotive. On a le système différentiel suivant:

$$\begin{aligned} \omega' &= k \cos(\varphi + \alpha) - F \\ \varphi'' &= \omega^2 \sin \varphi \cos \varphi - g \sin \varphi - b\varphi' \end{aligned}$$

avec $k, F, g, b > 0$, ω est la vitesse de rotation du moteur et l'accélération de φ est déterminée par la force centrifuge, poids et frottement [HNW93, p. 90–91].

Pour $\alpha = 0$, calculer le point d'équilibre et déterminer sous quelles conditions le point est asymptotiquement stable.

41. Calculer les solutions stationnaires de

$$\begin{aligned} y_1' &= -y_1y_2^2 - 2y_2 \\ y_2' &= y_1 - y_1^2y_2 \end{aligned}$$

et montrer que la solution nulle est stable.

Indication. Trouver $a > 0$ et $b > 0$ tels que la dérivée de $V(x) = ay_1^2(x) + by_2^2(x)$ soit non positive. (Une telle fonction s'appelle *fonction de Lyapunov*.)

42. Etudier la stabilité des quatre points critiques de l'équation différentielle (voir Fig. 9.2)

$$\begin{aligned} y_1' &= (y_1 - y_2)(1 - y_1 - y_2)/3 \\ y_2' &= y_1(2 - y_2) . \end{aligned}$$

43. (a) Montrer que

$$(y)'_1 y'_2 = (y)_2 - y_1 + (y_1^2 + y_2^2)^3 \sin\left(\frac{1}{y_1^2 + y_2^2}\right) (y)_1 y_2$$

possède dans chaque voisinage de l'origine, une infinité de solutions stables et une infinité de solutions instables.

Indication. Regarder l'équation différentielle en coordonnées polaires.

- (b) Est-ce que la solution nulle est stable, asymptotiquement stable?

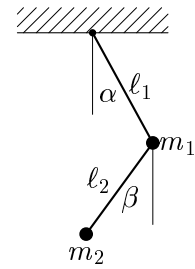
44. On considère l'équation $y'(x) = (A + B(x))y(x)$. Montrer que si les valeurs propres de A satisfont $\operatorname{Re}\lambda < 0$ et si $B(x) \rightarrow 0$ pour $x \rightarrow \infty$ alors l'équation différentielle est asymptotiquement stable.

Indication. Utiliser la démonstration du Théorème 9.9.

45. Pour le "pendule double" déterminer les équations du mouvement en utilisant le principe de Hamilton (voir Exercice II.19). Prendre les angles α et β comme coordonnées généralisées.

$$T = \frac{m_1}{2}(\dot{x}_1^2 + \dot{y}_1^2) + \frac{m_2}{2}(\dot{x}_2^2 + \dot{y}_2^2)$$

$$U = m_1 g y_1 + m_2 g y_2$$



Chapitre IV

Equations différentielles linéaires d'ordre 2

Contrairement au Chapitre III, où nous avons traité surtout le problème de Cauchy (équation différentielle avec valeurs initiales données), nous allons étudier des solutions d'équations différentielles assujéties à des conditions aux bords. Des exemples typiques sont ceux du paragraphe II.3, où on cherche des solutions d'une équation différentielle d'ordre 2 satisfaisant $y(a) = A$ et $y(b) = B$. Dans ce chapitre, nous restreignons notre étude à des équations différentielles linéaires.

IV.1 Problèmes aux limites – introduction

Pour illustrer la différence entre un problème de Cauchy et un problème aux limites, nous commençons par un exemple très simple.

Exemple 1.1 Considérons l'équation différentielle

$$y'' + y = 0,$$

dont la solution générale est donnée par $y(x) = c_1 \sin x + c_2 \cos x$. Avec des valeurs initiales données, $y(x_0) = y_0$, $y'(x_0) = y'_0$, le problème possède une solution unique quelles que soient les valeurs de x_0, y_0, y'_0 .

Si la solution cherchée est déterminée par des conditions aux bords d'un intervalle, on peut avoir les trois possibilités suivantes:

- a) $y(0) = 0$, $y(\pi) = 0$ une infinité des solutions, $y(x) = c_1 \sin x$;
- b) $y(0) = 0$, $y(\pi) = 1$ pas de solution;
- c) $y(0) = 0$, $y(1) = 2$ une solution unique, $y(x) = (2/\sin 1) \sin x$.

Dans une grande partie de ce chapitre, nous considérerons l'équation différentielle linéaire

$$a_2(x)y'' + a_1(x)y' + a_0(x)y = h(x), \tag{1.1}$$

où $a_0(x), a_1(x), a_2(x)$ et $h(x)$ sont continues sur l'intervalle $[a, b]$ et où $a_2(x) \neq 0$ sur $[a, b]$. Nous étudions les solutions $y : [a, b] \rightarrow \mathbb{R}$ assujéties aux *conditions aux limites* (conditions aux bords, conditions de frontière)

$$\begin{aligned} \alpha_1 y(a) + \alpha_2 y'(a) + \alpha_3 y(b) + \alpha_4 y'(b) &= \gamma_1 \\ \beta_1 y(a) + \beta_2 y'(a) + \beta_3 y(b) + \beta_4 y'(b) &= \gamma_2. \end{aligned} \tag{1.2}$$

Avec les abréviations

$$\begin{aligned}(Ly)(x) &:= a_2(x)y''(x) + a_1(x)y'(x) + a_0(x)y(x) \\ B_1y &:= \alpha_1y(a) + \alpha_2y'(a) + \alpha_3y(b) + \alpha_4y'(b) \\ B_2y &:= \beta_1y(a) + \beta_2y'(a) + \beta_3y(b) + \beta_4y'(b),\end{aligned}$$

où $L : \mathcal{C}^2([a, b]) \rightarrow \mathcal{C}^0([a, b])$ et $B_i : \mathcal{C}^1([a, b]) \rightarrow \mathbb{R}$, le problème s'écrit sous la forme

$$Ly = h, \quad B_1y = \gamma_1, \quad B_2y = \gamma_2. \quad (1.3)$$

On parle d'un problème *homogène*, si $h(x) \equiv 0$ et $\gamma_1 = \gamma_2 = 0$.

Théorème 1.2 (existence et unicité) *Soient $u(x), v(x)$ deux solutions indépendantes de $Ly = 0$. Alors, le problème (1.3) possède une solution unique si et seulement si*

$$\det \begin{pmatrix} B_1u & B_1v \\ B_2u & B_2v \end{pmatrix} \neq 0. \quad (1.4)$$

Par conséquent, le problème (1.3) possède une solution unique si et seulement si le problème homogène $Ly = 0, B_1y = 0, B_2y = 0$ possède une solution unique (alternative de Fredholm).

Démonstration. Soit $w(x)$ une solution particulière de $Ly = h$. La solution générale de $Ly = h$ est alors donnée par $y(x) = c_1u(x) + c_2v(x) + w(x)$. Cette solution vérifie les conditions aux limites (1.2) si et seulement si

$$\begin{pmatrix} B_1u & B_1v \\ B_2u & B_2v \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} \gamma_1 - B_1w \\ \gamma_2 - B_2w \end{pmatrix}.$$

Sous la condition (1.4), ce système linéaire pour c_1, c_2 possède une solution unique. \square

Exemple 1.3 Considérons le problème

$$y'' + y = 0, \quad y(a) = A, \quad y(b) = B.$$

On peut prendre $u(x) = \cos x$ et $v(x) = \sin x$. Par le Théorème 1.2, ce problème possède une solution unique si et seulement si

$$\det \begin{pmatrix} \cos a & \sin a \\ \cos b & \sin b \end{pmatrix} = \cos a \sin b - \cos b \sin a = \sin(a - b) \neq 0,$$

c.-à-d., si $b - a \neq k\pi$ avec un entier k .

IV.2 La fonction de Green

Le but de ce paragraphe est de dériver une formule analogue à celle de la “variation des constantes” pour le problème aux limites.

L'équation (1.1) peut être écrite sous la forme équivalente

$$\begin{pmatrix} y \\ y' \end{pmatrix}' = \begin{pmatrix} 0 & 1 \\ -a_0(x)/a_2(x) & -a_1(x)/a_2(x) \end{pmatrix} \begin{pmatrix} y \\ y' \end{pmatrix} + \begin{pmatrix} 0 \\ h(x)/a_2(x) \end{pmatrix}.$$

Notons la résolvante de ce système linéaire par

$$R(x, a) = \begin{pmatrix} R_1(x, a) & R_2(x, a) \\ R_1'(x, a) & R_2'(x, a) \end{pmatrix}.$$

Alors, la solution particulière de $Ly = h$ qui satisfait aux conditions initiales $y(a) = 0$, $y'(a) = 0$ est donnée par

$$y(x) = \int_a^x R_2(x, t) \frac{h(t)}{a_2(t)} dt = \int_a^b G_0(x, t) h(t) dt,$$

où

$$G_0(x, t) = \begin{cases} R_2(x, t)/a_2(t) & \text{si } a \leq t \leq x \leq b \\ 0 & \text{si } a \leq x \leq t \leq b. \end{cases} \quad (2.1)$$

Remarquons que la fonction $R_2(x, a)$ est la solution de $Ly = 0$ qui vérifie $R_2(a, a) = 0$ et $R_2'(a, a) = 1$.

Théorème 2.1 *Sous les hypothèses d'existence et d'unicité du Théorème 1.2, il existe une unique fonction continue $G(x, t)$ sur $[a, b] \times [a, b]$ telle que la solution de $Ly = h$, $B_1y = 0$, $B_2y = 0$ est donnée par*

$$y(x) = \int_a^b G(x, t) h(t) dt. \quad (2.2)$$

Cette fonction s'appelle fonction de Green.

Démonstration. Unicité. Soient $G_1(x, t)$ et $G_2(x, t)$ deux fonctions vérifiant (2.2). Comme la solution de $Ly = h$, $B_1y = 0$, $B_2y = 0$ est unique, on a

$$\int_a^b (G_1(x, t) - G_2(x, t)) h(t) dt = 0$$

pour toute fonction continue $h(t)$. Ceci implique $G_1(x, t) = G_2(x, t)$ sur $[a, b] \times [a, b]$.

Existence. Cherchons la fonction $G(x, t)$ sous la forme

$$G(x, t) = d_1(t)u(x) + d_2(t)v(x) + G_0(x, t), \quad (2.3)$$

où $u(x), v(x)$ sont deux solutions indépendantes de $Ly = 0$ et $d_1(t), d_2(t)$ sont des fonctions continues sur $[a, b]$. Avec cette fonction $G(x, t)$ la fonction $y(x)$ de (2.2) vérifie $Ly = h$.

Nous allons encore déterminer $d_1(t)$ et $d_2(t)$ de manière à ce que $B_1G(\cdot, t) = 0$ et $B_2G(\cdot, t) = 0$ pour tout $t \in (a, b)$. Ceci est équivalent à

$$\begin{pmatrix} B_1u & B_1v \\ B_2u & B_2v \end{pmatrix} \begin{pmatrix} d_1(t) \\ d_2(t) \end{pmatrix} = - \begin{pmatrix} B_1G_0(\cdot, t) \\ B_2G_0(\cdot, t) \end{pmatrix}.$$

Par hypothèse, ce système possède une solution unique pour $d_1(t), d_2(t)$. Evidemment, la propriété demandée pour $G(x, t)$ entraîne $B_1y = 0$ et $B_2y = 0$ pour la fonction de (2.2). \square

Exemple 2.2 Calculons la fonction de Green pour le problème

$$y'' = 0, \quad y(0) = 0, \quad y(1) = 0.$$

On a $R_2(x, t) = x - t$ et la fonction de Green est donc de la forme

$$G(x, t) = d_1(t) \cdot 1 + d_2(t) \cdot x + \begin{cases} x - t & \text{si } 0 \leq t \leq x \leq 1 \\ 0 & \text{si } 0 \leq x \leq t \leq 1. \end{cases} \quad (2.4)$$

Les conditions $G(0, t) = 0$ et $G(1, t) = 0$ impliquent $d_1(t) = 0$ et $d_2(t) + 1 - t = 0$. On en déduit

$$G(x, t) = \begin{cases} t(x - 1) & \text{si } 0 \leq t \leq x \leq 1 \\ (t - 1)x & \text{si } 0 \leq x \leq t \leq 1. \end{cases} \quad (2.5)$$

Par conséquent, la solution de $y'' = h(x)$, $y(0) = 0$, $y(1) = 0$ satisfait

$$y(x) = (x - 1) \int_0^x t h(t) dt + x \int_x^1 (t - 1) h(t) dt.$$

Remarques. La solution du problème (1.3) satisfait

$$y(x) = y_0(x) + \int_a^b G(x, t)h(t) dt,$$

où $y_0(x)$ est la solution de $Ly = 0$ satisfaisant les conditions aux bords $B_1y = \gamma_1$ et $B_2y = \gamma_2$.

On peut interpréter la fonction $x \mapsto G(x, t)$ comme solution de $Ly = \delta_t$ où δ_t est la "fonction δ de Dirac".

IV.3 Problèmes adjoints et auto-adjoints

Rappelons que, pour un opérateur linéaire A dans \mathbb{R}^n (donné par une matrice A), l'opérateur adjoint A^* est défini par la relation $\langle Au, v \rangle = \langle u, A^*v \rangle$. Il est donné par la transposée A^T de la matrice A .

Dans le contexte des équations différentielles, considérons le produit scalaire

$$\langle f, g \rangle = \int_a^b f(x)g(x) dx \quad (3.1)$$

sur $\mathcal{C}([a, b])$ (où $[a, b]$ est un intervalle compact) et un problème aux limites donné par L, B_1, B_2 . Le but de ce paragraphe est de déterminer L^*, B_1^*, B_2^* tels que

$$\langle Ly, z \rangle = \langle y, L^*z \rangle \quad (3.2)$$

pour toutes fonctions $y, z \in \mathcal{C}^2([a, b])$ satisfaisant $B_1y = B_2y = 0$ et $B_1^*z = B_2^*z = 0$.

Lemme 3.1 Soient $a_2 \in \mathcal{C}^2([a, b])$, $a_1 \in \mathcal{C}^1([a, b])$ et $a_0 \in \mathcal{C}^0([a, b])$ avec $a_2(x) \neq 0$ sur $[a, b]$. Avec l'opérateur

$$L^*z := (a_2(x)z)'' - (a_1(x)z)' + a_0(x)z \quad (3.3)$$

on a l'identité

$$\langle Ly, z \rangle - \langle y, L^*z \rangle = a_2(x) \left(y'(x)z(x) - y(x)z'(x) \right) \Big|_a^b + (a_1(x) - a_2'(x)) y(x)z(x) \Big|_a^b \quad (3.4)$$

pour toutes fonctions $y, z \in \mathcal{C}^2([a, b])$. L'opérateur L^* de (3.3) est appelé opérateur adjoint de L .

Démonstration. Dans l'expression

$$\langle Ly, z \rangle = \int_a^b (a_2(x)y''(x) + a_1(x)y'(x) + a_0(x)y(x))z(x) dx \quad (3.5)$$

il faut éliminer les dérivées de $y(x)$. Ceci se fait par intégration par parties comme suit:

$$\begin{aligned} \int_a^b a_2(x)y''(x)z(x) dx &= y'(x)a_2(x)z(x) \Big|_a^b - \int_a^b y'(x)(a_2z)'(x) dx \\ &= y'(x)a_2(x)z(x) \Big|_a^b - y(x)(a_2z)'(x) \Big|_a^b + \int_a^b y(x)(a_2z)''(x) dx \\ \int_a^b a_1(x)y'(x)z(x) dx &= y(x)a_1(x)z(x) \Big|_a^b - \int_a^b y(x)(a_1z)'(x) dx. \end{aligned}$$

En insérant ces deux intégrales dans (3.5), on obtient l'identité désirée. \square

Théorème 3.2 *Considérons le problème $Ly = 0$, $B_1y = 0$, $B_2y = 0$, où L satisfait les hypothèses du Lemme 3.1 et où les vecteurs $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)$ et $(\beta_1, \beta_2, \beta_3, \beta_4)$ définissant B_1 et B_2 sont linéairement indépendants. Alors, il existe*

$$\begin{aligned} B_1^*z &= \gamma_1z(a) + \gamma_2z'(a) + \gamma_3z(b) + \gamma_4z'(b) \\ B_2^*z &= \delta_1z(a) + \delta_2z'(a) + \delta_3z(b) + \delta_4z'(b), \end{aligned} \quad (3.6)$$

tels que $\langle Ly, z \rangle = \langle y, L^*z \rangle$ pour toutes les fonctions $y, z \in \mathcal{C}^2([a, b])$ satisfaisant $B_1y = B_2y = 0$ et $B_1^*z = B_2^*z = 0$. Les coefficients γ_i et δ_i ne sont pas uniques, mais l'espace des fonctions z qui satisfont $B_1^*z = B_2^*z = 0$ est unique.

Démonstration. Le membre droit de l'équation (3.4) peut être écrit sous la forme $\vec{y}^T C \vec{z}$ où $\vec{y} = (y(a), y'(a), y(b), y'(b))^T$, $\vec{z} = (z(a), z'(a), z(b), z'(b))^T$ et

$$C = \begin{pmatrix} -D(a) & 0 \\ 0 & D(b) \end{pmatrix} \quad \text{avec} \quad D(x) = \begin{pmatrix} a_1(x) - a_2'(x) & -a_2(x) \\ a_2(x) & 0 \end{pmatrix}.$$

Considérons une matrice inversible B de dimension 4, dont les deux premières colonnes sont les vecteurs $(\alpha_1, \alpha_2, \alpha_3, \alpha_4)^T$ et $(\beta_1, \beta_2, \beta_3, \beta_4)^T$. On a alors que $(B_1y, B_2y, \cdot, \cdot) = \vec{y}^T B$ et on obtient

$$\vec{y}^T C \vec{z} = (B_1y, B_2y, \cdot, \cdot) B^{-1} C \vec{z}.$$

Si l'on dénote par $(\gamma_1, \gamma_2, \gamma_3, \gamma_4)$ et $(\delta_1, \delta_2, \delta_3, \delta_4)$ les deux dernières lignes de la matrice $B^{-1}C$, on voit que $\vec{y}^T C \vec{z} = 0$ (et donc $\langle Ly, z \rangle = \langle y, L^*z \rangle$) pour des fonctions y, z satisfaisant $B_1y = B_2y = 0$ et $B_1^*z = B_2^*z = 0$.

Les coefficients γ_i et δ_i dépendent du choix de la matrice B . Si \hat{B} est une autre matrice inversible dont les deux premières colonnes sont comme pour B , on a

$$\hat{B} = B \begin{pmatrix} I & U \\ 0 & V \end{pmatrix} \quad \text{et} \quad \hat{B}^{-1} = \begin{pmatrix} I & -UV^{-1} \\ 0 & V^{-1} \end{pmatrix} B^{-1},$$

où I est la matrice identité de dimension 2. Les coefficients $\hat{\gamma}_i$ et $\hat{\delta}_i$, qui correspondent au choix \hat{B} , sont alors reliés avec γ_i et δ_i par

$$\begin{pmatrix} \hat{\gamma}_1 & \hat{\gamma}_2 & \hat{\gamma}_3 & \hat{\gamma}_4 \\ \hat{\delta}_1 & \hat{\delta}_2 & \hat{\delta}_3 & \hat{\delta}_4 \end{pmatrix} = V^{-1} \begin{pmatrix} \gamma_1 & \gamma_2 & \gamma_3 & \gamma_4 \\ \delta_1 & \delta_2 & \delta_3 & \delta_4 \end{pmatrix}.$$

Ils définissent alors des conditions aux bords qui sont équivalentes à (3.6). \square

Définition 3.3 Si le problème L, B_1, B_2 satisfait les hypothèses du Théorème 3.2, on appelle L^*, B_1^*, B_2^* , données par (3.3) et (3.6), le *problème aux limites adjoint*.

Le problème est dit *auto-adjoint*, si $L^* = L$ et si les conditions $B_1^*z = B_2^*z = 0$ sont équivalentes à $B_1z = B_2z = 0$.

En développant l'expression (3.3) pour L^*z on obtient

$$L^*z = a_2(x)z'' + (2a_2'(x) - a_1(x))z' + (a_2''(x) - a_1'(x) + a_0(x))z.$$

L'opérateur L est donc auto-adjoint, si et seulement si $a_2'(x) = a_1(x)$, c.-à-d., s'il est de la forme $Ly = (a_2(x)y')' + a_0(x)y$.

Exemple 3.4 Le problème

$$\begin{aligned} (a_2(x)y')' + a_0(x)y &= 0 & a_2 \in \mathcal{C}^1([a, b]) \text{ et } a_2(x) \neq 0 \text{ sur } [a, b], a_0 \in \mathcal{C}^0([a, b]), \\ \alpha y(a) + \beta y'(a) &= 0 & |\alpha| + |\beta| > 0, \\ \gamma y(b) + \delta y'(b) &= 0 & |\gamma| + |\delta| > 0 \end{aligned}$$

est un problème auto-adjoint. Pour voir ceci, il suffit de démontrer que

$$\langle Ly, z \rangle - \langle y, Lz \rangle = a_2(x) \left(y'(x)z(x) - y(x)z'(x) \right) \Big|_a^b$$

(voir le Lemme 3.1) s'annule pour toutes les fonctions y et z satisfaisant les conditions aux bords. On voit facilement que $y'(x)z(x) - y(x)z'(x)$ s'annule pour $x = a$ et pour $x = b$. Par exemple, si $\alpha \neq 0$, on a

$$y'(a)z(a) - y(a)z'(a) = y'(a) \left(-\frac{\beta}{\alpha} z'(a) \right) - \left(-\frac{\beta}{\alpha} y'(a) \right) z'(a) = 0.$$

Proposition 3.5 *Sous les hypothèses du Théorème 3.2, le problème $Ly = 0$, $B_1y = 0$, $B_2y = 0$ possède une solution unique, si et seulement si le problème adjoint $L^*z = 0$, $B_1^*z = 0$, $B_2^*z = 0$ possède une solution unique.*

Les fonctions de Green pour les deux problèmes satisfont alors

$$G^*(x, t) = G(t, x). \quad (3.7)$$

En particulier, la fonction de Green d'un problème auto-adjoint est symétrique.

Démonstration. Supposons que $Ly = 0$, $B_1y = 0$, $B_2y = 0$ possède une solution unique et notons par z une solution de $L^*z = 0$, $B_1^*z = 0$, $B_2^*z = 0$. Avec $y(x)$, solution de $Ly = z$, $B_1y = 0$, $B_2y = 0$, on obtient

$$\|z\|^2 = \langle z, z \rangle = \langle Ly, z \rangle = \langle y, L^*z \rangle = \langle y, 0 \rangle = 0.$$

Donc $z = 0$ et l'unicité de la solution du problème adjoint est démontrée.

Pour démontrer la relation (3.7), considérons deux fonctions continues $f(x)$ et $g(x)$ et définissons

$$y(x) = \int_a^b G(x, t) f(t) dt, \quad z(x) = \int_a^b G^*(x, t) g(t) dt.$$

Par définition de la fonction de Green, les fonctions y et z satisfont $Ly = f$, $B_1y = 0$, $B_2y = 0$ et $L^*z = g$, $B_1^*z = 0$, $B_2^*z = 0$. La relation $\langle Ly, z \rangle = \langle y, L^*z \rangle$ (c.-à-d., $\langle f, z \rangle = \langle y, g \rangle$) implique alors

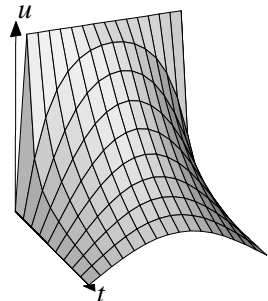
$$\int_a^b \int_a^b (G^*(x, t) - G(t, x)) f(x) g(t) dt dx = 0.$$

Comme cette relation est vérifiée pour toutes fonctions continues f et g , on en déduit la relation (3.7). \square

IV.4 Le problème de Sturm–Liouville

Une motivation du problème de Sturm–Liouville (voir plus bas pour sa définition) est l'étude de *l'équation de la chaleur*

$$\begin{aligned} \frac{\partial u}{\partial t} &= \frac{\partial^2 u}{\partial x^2} && \text{pour } x \in (0, 1), t > 0, \\ u(t, 0) &= 0, u(t, 1) = 0 && \text{(conditions aux bords)} \\ u(0, x) &= f(x) && \text{(condition initiale).} \end{aligned}$$



L'idée consiste à chercher des solutions de la forme $u(t, x) = z(t) \cdot y(x)$ ("séparation des variables"). L'équation aux dérivées partielles devient alors $z'(t)y(x) = z(t)y''(x)$ ce qui est équivalent à

$$\frac{z'(t)}{z(t)} = \frac{y''(x)}{y(x)} = \lambda$$

(λ est forcément constant, car $z'(t)/z(t)$ ne dépend que de t et $y''(x)/y(x)$ seulement de x). On est alors conduit au problème

$$y'' = \lambda y, \quad y(0) = 0, \quad y(1) = 0 \quad (4.1)$$

et à la question: *pour quel λ ce problème possède-t-il une solution non-nulle?* Dans ce cas particulier, il est facile de trouver la réponse. La solution générale de $y'' = \lambda y$ est $c_1 e^{\sqrt{\lambda}x} + c_2 e^{-\sqrt{\lambda}x}$ (si $\lambda \neq 0$). Donc, le problème (4.1) possède une solution non-nulle si et seulement si

$$0 = \det \begin{pmatrix} 1 & 1 \\ e^{\sqrt{\lambda}} & e^{-\sqrt{\lambda}} \end{pmatrix} = e^{-\sqrt{\lambda}} - e^{\sqrt{\lambda}} = e^{-\sqrt{\lambda}}(1 - e^{2\sqrt{\lambda}}).$$

Par conséquent, il faut que $2\sqrt{\lambda} = 2k\pi i$ avec un entier $k \neq 0$, et donc $\lambda = -k^2\pi^2$. La solution non-nulle de (4.1) est donnée par $y(x) = \sin(k\pi x)$. Comme $e^{\lambda t}$ est une solution de $z' = \lambda z$, on obtient $e^{-k^2\pi^2 t} \sin(k\pi x)$ comme solution de l'équation de la chaleur qui vérifie les conditions aux bords. Par linéarité, on voit que

$$u(t, x) = \sum_{k \geq 1} c_k e^{-k^2\pi^2 t} \sin(k\pi x) \quad (4.2)$$

est aussi une solution. Pour satisfaire la condition initiale $f(x) = u(0, x) = \sum_{k \geq 1} c_k \sin(k\pi x)$, on choisit c_k comme coefficient de Fourier de la fonction $f(x)$.

Pour un problème plus général que l'équation de la chaleur, par exemple,

$$\frac{\partial u}{\partial t} = \frac{1}{r(x)} \left(\frac{\partial}{\partial x} \left(p(x) \frac{\partial u}{\partial x} \right) + q(x)u \right) \quad (4.3)$$

avec les conditions aux bords $u(t, 0) = 0$ (Dirichlet) et $\frac{\partial u}{\partial x}(t, 1) = 0$ (Neumann) et la condition initiale $u(0, x) = f(x)$, la "séparation des variables" donne

$$\frac{z'(t)}{z(t)} = \frac{1}{r(x)y(x)} \left((p(x)y'(x))' + q(x)y(x) \right) = -\lambda.$$

Il faut alors étudier l'existence de solutions non-nulles du problème

$$(p(x)y')' + q(x)y + \lambda r(x)y(x) = 0, \quad y(0) = 0, \quad y'(1) = 0.$$

Formulation du problème de Sturm–Liouville. Considérons

$$Ly + \lambda r y = 0, \quad B_1 y = 0, \quad B_2 y = 0 \quad (4.4)$$

où

$$\begin{aligned} Ly &= (p(x)y')' + q(x)y & p &\in \mathcal{C}^1([a, b]) \text{ et } p(x) > 0 \text{ sur } [a, b], q \in \mathcal{C}^0([a, b]) \\ B_1 y &= \alpha y(a) + \beta p(a)y'(a) & |\alpha| + |\beta| &> 0 \\ B_2 y &= \gamma y(b) + \delta p(b)y'(b) & |\gamma| + |\delta| &> 0 \end{aligned}$$

et où $r(x)$ est continue et strictement positive sur $[a, b]$. Le problème consiste à chercher des nombres complexes λ et des fonctions non-nulles $y(x)$, tels que (4.4) est vérifié. On appelle λ une *valeur propre* et $y(x)$ une *fonction propre* du problème de Sturm–Liouville. Le but est d'étudier les valeurs propres et les fonctions propres et de montrer l'existence d'une infinité de valeurs propres comme c'est le cas pour le problème (4.1).

Théorème 4.1 *Les valeurs propres du problème de Sturm–Liouville sont*

- *réelles (parce que le problème est auto-adjoint),*
- *simples (c.-à-d., les fonction propres forment un espace de dimension 1).*

Démonstration. Soient $\lambda = \lambda_1 + i\lambda_2$ une valeur propre et $y = y_1 + iy_2$ une fonction propre correspondante. Les parties réelles et imaginaires de (4.4) donnent

$$\begin{aligned} Ly_1 + r(\lambda_1 y_1 - \lambda_2 y_2) &= 0 & B_1 y_1 &= 0 & B_2 y_1 &= 0 \\ Ly_2 + r(\lambda_2 y_1 + \lambda_1 y_2) &= 0 & B_1 y_2 &= 0 & B_2 y_2 &= 0. \end{aligned}$$

Comme le problème L, B_1, B_2 est auto-adjoint, on a

$$\begin{aligned} 0 &= \langle Ly_1, y_2 \rangle - \langle y_1, Ly_2 \rangle \\ &= \int_a^b r(x) (-\lambda_1 y_1(x) + \lambda_2 y_2(x)) y_2(x) dx + \int_a^b r(x) (\lambda_2 y_1(x) + \lambda_1 y_2(x)) y_1(x) dx \\ &= \lambda_2 \int_a^b r(x) (y_1^2(x) + y_2^2(x)) dx. \end{aligned}$$

Ceci démontre $\lambda_2 = 0$, car $r(x) > 0$ sur $[a, b]$ et $y_1 + iy_2 \neq 0$.

Supposons que, pour un λ fixé, le problème (4.4) possède deux solutions indépendantes $u(x)$ et $v(x)$. Alors, toutes les solutions de $Ly + \lambda ry = 0$ sont de la forme $y(x) = c_1 u(x) + c_2 v(x)$ et satisfont $B_1 y = B_2 y = 0$. Ceci est une contradiction, car on peut choisir $y(a), y'(a)$ tels que $\alpha y(a) + \beta p(a) y'(a) \neq 0$. \square

Une conséquence de ce théorème est que les fonctions propres peuvent être choisies réelles. En effet, si $y = y_1 + iy_2$ est une fonction propre pour un $\lambda \in \mathbb{R}$, y_1 et y_2 sont aussi des fonctions propres pour λ . Elles sont donc linéairement dépendantes.

Théorème 4.2 (orthogonalité des fonctions propres) *Soit $y(x)$ une fonction propre pour la valeur propre λ et soit $z(x)$ une fonction propre pour μ ($\mu \neq \lambda$). Alors,*

$$\int_a^b r(x) y(x) z(x) dx = 0.$$

Démonstration. Comme y et z satisfont les conditions aux bords, on a

$$0 = \langle Ly, z \rangle - \langle y, Lz \rangle = -\lambda \langle ry, z \rangle + \mu \langle y, rz \rangle = (\mu - \lambda) \int_a^b r(x) y(x) z(x) dx.$$

L'affirmation est alors une conséquence de $\mu \neq \lambda$. \square

Exemple 4.3 Pour le problème

$$y'' + \lambda y = 0, \quad y(0) = 0, \quad y(1) = 0$$

(voir (4.1)) les valeurs propres sont $\lambda_k = k^2 \pi^2$ et les fonctions propres sont $y_k(x) = \sin(k\pi x)$. Le Théorème 4.2 est une généralisation de la relation bien connue

$$\int_0^1 \sin(k\pi x) \sin(\ell\pi x) dx = 0 \quad \text{si } k \neq \ell.$$

IV.5 Théorème de comparaison de Sturm

Nous allons démontrer qu'un problème auto-adjoint possède une infinité de valeurs propres λ_k qui tendent vers $+\infty$.

Théorème 5.1 (Sturm 1836) *Considérons les deux équations différentielles*

$$\begin{aligned} (p(x)y')' + q(x)y &= 0 \\ (\hat{p}(x)\hat{y}')' + \hat{q}(x)\hat{y} &= 0 \end{aligned} \quad (5.1)$$

($p(x), \hat{p}(x)$ continûment différentiables et $q(x), \hat{q}(x)$ continues) et supposons que

$$0 < \hat{p}(x) \leq p(x), \quad \hat{q}(x) \geq q(x) \quad \text{sur } [a, b]. \quad (5.2)$$

Si les solutions $y(x), \hat{y}(x)$ sont linéairement indépendantes et si $y(x_1) = y(x_2) = 0$ avec $x_1 < x_2$, alors il existe $x_3 \in (x_1, x_2)$ tel que $\hat{y}(x_3) = 0$.

Démonstration. Les expressions $y(x)$ et $p(x)y'(x)$ ne s'annulent pas en même temps (sinon $y(x) \equiv 0$, ce qui n'est pas possible). On peut donc utiliser des coordonnées polaires

$$p(x)y'(x) = \rho(x) \cos \varphi(x), \quad y(x) = \rho(x) \sin \varphi(x)$$

où $\rho(x) > 0$ sur $[a, b]$. On calcule les dérivées

$$\begin{aligned} (p(x)y'(x))' &= \rho'(x) \cos \varphi(x) - \rho(x) \varphi'(x) \sin \varphi(x) \\ y'(x) &= \rho'(x) \sin \varphi(x) + \rho(x) \varphi'(x) \cos \varphi(x) \end{aligned}$$

et, en utilisant (5.1), on obtient les équations différentielles suivantes:

$$\varphi' = \frac{1}{p(x)} \cos^2 \varphi + q(x) \sin^2 \varphi, \quad \rho' = \left(\frac{1}{p(x)} - q(x) \right) \sin \varphi \cos \varphi \cdot \rho. \quad (5.3)$$

De la même manière on définit $\hat{\rho}(x), \hat{\varphi}(x)$ et on obtient des équations différentielles analogues.

Supposons maintenant que x_1, x_2 soient deux zéros consécutifs de $y(x)$. Alors $\varphi(x_1)$ et $\varphi(x_2)$ sont des multiples entiers de π . De l'équation différentielle (5.3), on voit que $\varphi'(x_1) > 0$ et $\varphi'(x_2) > 0$. Donc, on peut supposer que

$$\varphi(x_1) = 0, \quad \varphi(x_2) = \pi, \quad \hat{\varphi}(x_1) \in [0, \pi)$$

(on suppose $\hat{y}(x_1) \geq 0$, sinon on multiplie $\hat{y}(x)$ par -1). L'hypothèse (5.2) nous donne l'inégalité différentielle

$$\hat{\varphi}' = \frac{1}{\hat{p}(x)} \cos^2 \hat{\varphi} + \hat{q}(x) \sin^2 \hat{\varphi} \geq \frac{1}{p(x)} \cos^2 \hat{\varphi} + q(x) \sin^2 \hat{\varphi}. \quad (5.4)$$

Comme $\hat{\varphi}(x_1) \geq \varphi(x_1)$, on en déduit (utiliser le Théorème III.5.2; pour être précis, il faut soustraire un $\epsilon > 0$ et considérer la limite $\epsilon \rightarrow 0$) que

$$\hat{\varphi}(x) \geq \varphi(x) \quad \text{pour tout } x \in [x_1, x_2].$$

Dans le cas où il existe $x_0 \in (x_1, x_2)$ avec $\hat{\varphi}(x_0) > \varphi(x_0)$, on a $\hat{\varphi}(x) > \varphi(x)$ pour tout $x \geq x_0$ (car $\hat{\varphi}(x) \geq \tilde{\varphi}(x) > \varphi(x)$ où $\tilde{\varphi}(x)$ est la solution de (5.3) avec $\tilde{\varphi}(x_0) = \hat{\varphi}(x_0)$). Comme $\varphi(x_2) = \pi$, il existe alors $x_3 \in (x_1, x_2)$ avec $\hat{\varphi}(x_3) = \pi$, c.-à-d., $\hat{y}(x_3) = 0$.

Dans le cas contraire, on a $\widehat{\varphi}(x) = \varphi(x)$ sur $[x_1, x_2]$. Ceci implique $\widehat{\varphi}'(x) = \varphi'(x)$ et

$$\left(\frac{1}{\widehat{p}(x)} - \frac{1}{p(x)}\right) \cos^2 \varphi(x) + (\widehat{q}(x) - q(x)) \sin^2 \varphi(x) = 0.$$

Par l'hypothèse (5.2) les deux termes s'annulent séparément:

$$\frac{1}{\widehat{p}(x)} \cos \varphi(x) = \frac{1}{p(x)} \cos \varphi(x), \quad \widehat{q}(x) \sin \varphi(x) = q(x) \sin \varphi(x).$$

On voit donc que $\rho(x)$ et $\widehat{\rho}(x)$ sont solutions de la même équation différentielle du type $\rho' = a(x)\rho$. On en déduit que $\widehat{\rho}(x) = C\rho(x)$ et $\widehat{y}(x) = Cy(x)$, ce qui est une contradiction. \square

Exemple 5.2 Comme illustration pour la démonstration du théorème suivant, nous considérons le problème

$$((1 - 0.8 \sin^2 x)y')' - (x - \lambda)y = 0, \quad y(0) = 0, \quad y(\pi) = 0. \quad (5.5)$$

Fig. 5.1 (dessin de gauche) montre les solutions de l'équation différentielle (5.5) avec pour valeurs initiales $y(0) = 0, y'(0) = 1$ et pour les valeurs $\lambda = 2.1, 2.1 + \delta, 2.1 + 2\delta$, pour $\lambda = 4, 4 + 3\delta, 4 + 6\delta$, ainsi que pour $\lambda = 6, 6 + 6\delta, 6 + 12\delta$ ($\delta = 0.05$). On peut observer que les racines de la solution se déplacent vers la gauche (en direction de l'origine), si λ croît.

Le dessin de droite montre les cinq premières fonctions propres du problème (5.5), normalisées à $y'(0) = 1$. Les valeurs propres de ce problème sont 2.1224, 3.6078, 6.0016, 9.3773, 13.7298, 19.053, 25.347, 32.609, 40.841, 50.041, ...

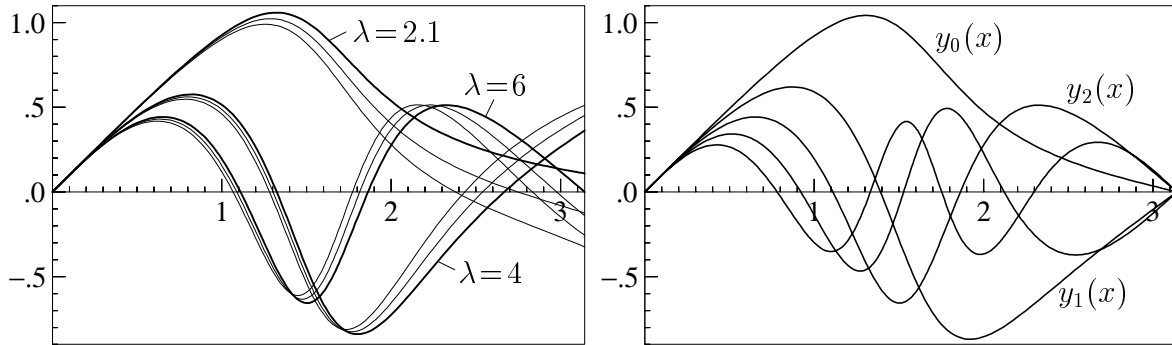


FIG. 5.1 – Solutions du problème de Sturm-Liouville (Exemple 5.2)

Théorème 5.3 Considérons le problème

$$(p(x)y')' + q(x)y + \lambda y = 0, \quad y(0) = 0, \quad y(1) = 0, \quad (5.6)$$

où $p(x)$ est continûment différentiable et strictement positive sur $[0, 1]$, et où $q(x)$ est continue. Alors, les valeurs propres et les fonctions propres satisfont:

- les valeurs propres forment une suite infinie $\lambda_0 < \lambda_1 < \lambda_2 < \dots$,
- la j ème fonction propre $y_j(x)$ possède exactement j zéros dans l'intervalle ouvert $(0, 1)$,
- entre deux zéros consécutifs de $y_j(x)$ il y a un zéro unique de $y_{j+1}(x)$,
- si $0 < K_1 \leq p(x) \leq K_2$ et $L_2 \leq q(x) \leq L_1$ sur $[0, 1]$, alors

$$-L_1 + K_1 j^2 \pi^2 \leq \lambda_j \leq -L_2 + K_2 j^2 \pi^2, \quad (5.7)$$

en particulier, $\lambda_j \rightarrow \infty$ pour $j \rightarrow \infty$.

Démonstration. Notons par $y(x, \lambda)$ la solution de

$$(p(x)y')' + (q(x) + \lambda)y = 0, \quad y(0) = 0, \quad y'(0) = 1 \quad (5.8)$$

(on peut supposer la normalisation $y'_j(0) = 1$ pour les fonctions propres). La fonction $y(x, \lambda)$ est de classe \mathcal{C}^1 (voir le paragraphe III.8). Notons par $\xi_j(\lambda)$ la j ème racine positive de $y(x, \lambda) = 0$. Cette racine est simple (à cause de l'unicité de la solution) et donc $\xi_j(\lambda)$ est continûment différentiable (par le théorème des fonctions implicites).

a) Montrons d'abord que $\xi_j(\lambda)$ est strictement décroissant

$$\xi_j(\lambda + \Delta\lambda) < \xi_j(\lambda) \quad \text{pour } \Delta\lambda > 0.$$

L'idée est d'appliquer le Théorème 5.1 avec $\hat{p}(x) = p(x)$ et avec $q(x)$ et $\hat{q}(x)$ remplacés par $q(x) + \lambda$ et $q(x) + \lambda + \Delta\lambda$ respectivement. Entre 0 et $\xi_1(\lambda)$ il y a donc un zéro de $y(x, \lambda + \Delta\lambda)$, c.-à-d., $\xi_1(\lambda + \Delta\lambda) < \xi_1(\lambda)$. Entre $\xi_1(\lambda)$ et $\xi_2(\lambda)$ il y a un autre zéro de $y(x, \lambda + \Delta\lambda)$, donc $\xi_2(\lambda + \Delta\lambda) < \xi_2(\lambda)$, etc.

b) Comparons l'équation différentielle (5.8) avec $K_1\hat{y}'' + (L_1 + \lambda)\hat{y} = 0$, dont $\hat{y}(x) = \sin(x\sqrt{(L_1 + \lambda)/K_1})$ est une solution. Si

$$L_1 + \lambda < K_1j^2\pi^2,$$

le nombre de zéros de $\hat{y}(x)$ dans $(0, 1]$ est $\leq j - 1$. Par le Théorème 5.1 de Sturm, le nombre de zéros de $y(x, \lambda)$ dans $(0, 1]$ est aussi $\leq j - 1$ (argumenter par l'absurde).

c) Comparons encore avec $K_2\hat{y}'' + (L_2 + \lambda)\hat{y} = 0$, et $\hat{y}(x) = \sin(x\sqrt{(L_2 + \lambda)/K_2})$. Si

$$L_2 + \lambda > K_2j^2\pi^2,$$

le nombre de zéros de $\hat{y}(x)$ dans $(0, 1)$ est $\geq j$. En échangeant les rôles de y et \hat{y} dans le Théorème 5.1, on voit que le nombre de zéros de $y(x, \lambda)$ dans $(0, 1)$ est aussi $\geq j$.

Les affirmations du théorème sont une conséquence directe de ces trois propriétés. L'existence d'une suite infinie de valeurs propres et l'estimation (5.7) suivent de (b) et (c). La monotonie de $\xi_j(\lambda)$ implique les propriétés affirmées concernant des fonctions propres. \square

Remarque. Sous les hypothèses du Théorème 5.3 les fonctions propres engendrent tout l'espace $L_2([a, b])$. Une série comme dans (4.2) nous permet donc de satisfaire la condition initiale pour (4.3).

IV.6 Equations différentielles avec des singularités

Les vibrations transversales de la membrane d'un tambour sont décrites par l'équation des ondes

$$\frac{\partial^2 u}{\partial t^2} = c^2 \Delta u \quad \text{pour } (x, y) \in \Omega \quad \text{et } t > 0, \quad (6.1)$$

où $\Omega = \{(x, y); x^2 + y^2 < 1\}$. La valeur $u(t, x, y)$ représente le déplacement transversal d'un point (x, y) de la membrane en fonction du temps. On suppose $u(t, x, y) = 0$ pour $(x, y) \in \partial\Omega$ (la membrane est fixée au bord) et des valeurs initiales pour $u(0, x, y)$ et $\frac{\partial u}{\partial t}(0, x, y)$. Comme dans le paragraphe IV.4 pour l'équation de la chaleur, la "séparation des variables" $u(t, x, y) = T(t) \cdot v(x, y)$ conduit à

$$\frac{T''}{c^2 T} = \frac{\Delta v}{v} = -\lambda^2.$$

L'équation différentielle pour $T(t)$ donne les oscillations en temps et $v(x, y)$ doit satisfaire

$$\begin{aligned} \Delta v + \lambda^2 v &= 0 & \text{sur } \Omega \\ v &= 0 & \text{sur } \partial\Omega. \end{aligned} \quad (6.2)$$

Pour résoudre ce problème sur le disque Ω , il est naturel d'introduire des coordonnées polaires $x = r \cos \varphi$, $y = r \sin \varphi$ et de considérer la fonction $w(r, \varphi) = v(x, y) = v(r \cos \varphi, r \sin \varphi)$. Un calcul direct montre que le problème (6.2) devient

$$\frac{\partial^2 w}{\partial r^2} + \frac{1}{r} \frac{\partial w}{\partial r} + \frac{1}{r^2} \frac{\partial^2 w}{\partial \varphi^2} + \lambda^2 w = 0 \quad (6.3)$$

$$w(1, \varphi) = 0, \quad w(r, 0) = w(r, 2\pi), \quad \frac{\partial w}{\partial \varphi}(r, 0) = \frac{\partial w}{\partial \varphi}(r, 2\pi), \quad (6.4)$$

où la limite $\lim_{r \rightarrow 0} w(r, \varphi)$ est constante et indépendante de φ . Une deuxième "séparation des variables" $w(r, \varphi) = a(r) \cdot z(\varphi)$ donne

$$\frac{r^2 a'' + r a' + \lambda^2 r^2 a}{a} = -\frac{z''}{z} = C.$$

La condition de périodicité $z(0) = z(2\pi)$, $z'(0) = z'(2\pi)$ implique que $C = n^2$ avec un entier n . Les solutions pour $z(\varphi)$ sont alors $z(\varphi) = c_1 \cos n\varphi + c_2 \sin n\varphi$. Pour la fonction $a(r)$ on obtient l'équation différentielle

$$r^2 a'' + r a' + (\lambda^2 r^2 - n^2) a = 0 \quad (6.5)$$

avec les conditions aux bords

$$a(1) = 0 \quad \text{et} \quad a(0) = \begin{cases} 0 & \text{si } n > 0 \\ \text{une valeur finie} & \text{si } n = 0. \end{cases}$$

Pour $\lambda \neq 0$ la transformation $x = \lambda r$ et $y(x) = a(r)$ élimine le paramètre λ de l'équation différentielle et (6.5) devient l'équation de Bessel

$$x^2 y'' + x y' + (x^2 - n^2) y = 0. \quad (6.6)$$

Nous allons démontrer plus tard dans ce paragraphe que cette équation différentielle possède une solution $J_n(x)$ qui est bornée pour $x \rightarrow 0$ (on a $J_n(0) = 0$ pour $n > 0$) et qui possède une infinité de zéros positifs $0 < j_{n1} < j_{n2} < j_{n3} < \dots$ (voir la Fig. 6.1 et l'Exercice 17). La fonction $a(r) = J_n(j_{nk} r)$ satisfait alors (6.5) avec $\lambda = j_{nk}$ et aussi $a(1) = 0$. Par conséquent,

$$J_n(j_{nk} r) \cos n\varphi \quad \text{et} \quad J_n(j_{nk} r) \sin n\varphi$$

sont des solutions de (6.2) (Fig. 6.2). Par superposition, on obtient la solution générale.

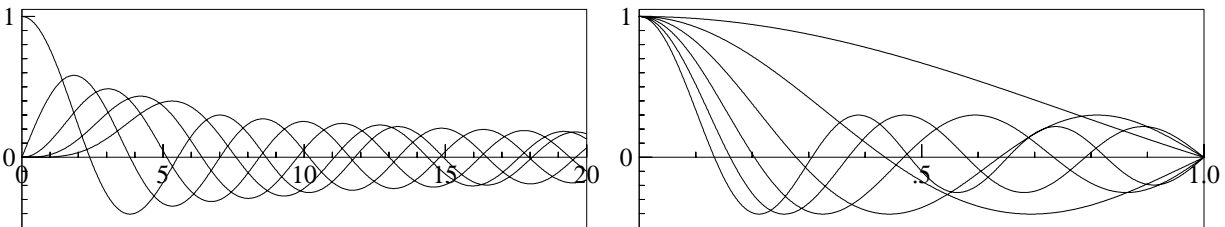


FIG. 6.1 – Fonctions de Bessel $J_n(x)$ (gauche), $J_0(j_{0k} r)$ (droite)

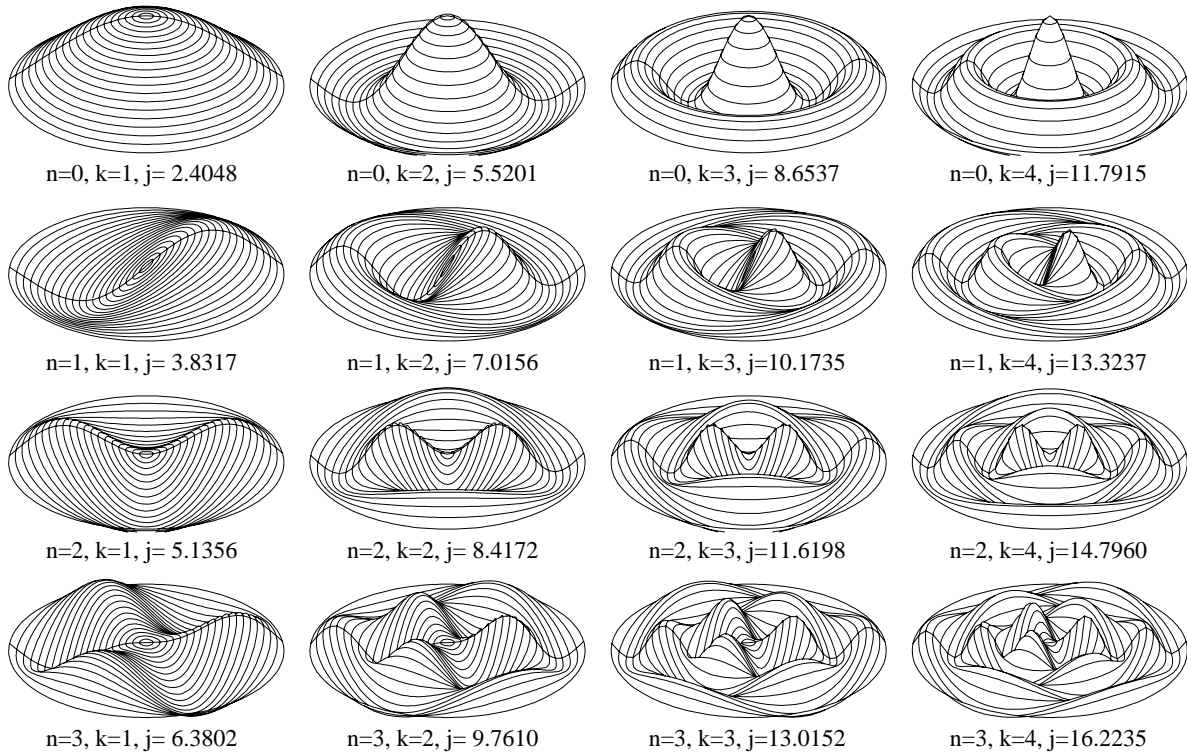


FIG. 6.2 – Fonctions propres $J_n(j_{nk}r) \cos n\varphi$ du Laplacien sur le disque

Une particularité de l'équation de Bessel (6.6) est que le coefficient $a_2(x) = x^2$ de y'' s'annule en un point où on s'intéresse à la solution.

Définition 6.1 Si l'équation différentielle est sous la forme

$$(x - x_0)^2 y'' + (x - x_0)p(x)y' + q(x)y = 0, \quad (6.7)$$

où $p(x)$ et $q(x)$ sont analytiques dans un voisinage de x_0 , alors le point x_0 s'appelle *point singulier régulier* de l'équation différentielle.

Exemple 6.2 Le point $x = 0$ est un point singulier régulier de l'équation de Bessel (6.6). Pour l'équation hypergéométrique

$$x(1 - x)y'' + (c - (a + b + 1)x)y' - aby = 0 \quad (6.8)$$

les points $x = 0$ et $x = 1$ sont singuliers réguliers (multiplier l'équation par $x/(1 - x)$ pour voir que (6.8) est sous la forme (6.7) proche de l'origine).

Pour résoudre (6.7), écrivons les fonctions $p(x)$, $q(x)$ sous la forme d'une série

$$p(x) = \sum_{j \geq 0} p_j (x - x_0)^j, \quad q(x) = \sum_{j \geq 0} q_j (x - x_0)^j. \quad (6.9)$$

Si $p(x) = p_0$ et $q(x) = q_0$ sont constantes (équation de Cauchy, voir [HW95, Sect. II.8]), on trouve des solutions de la forme $y(x) = (x - x_0)^\alpha$ où $\alpha(\alpha - 1) + p_0\alpha + q_0 = 0$. Ceci suggère de chercher des solutions de (6.7) sous la forme

$$y(x) = (x - x_0)^\alpha \sum_{j \geq 0} c_j (x - x_0)^j = \sum_{j \geq 0} c_j (x - x_0)^{j+\alpha} \quad \text{avec } c_0 \neq 0. \quad (6.10)$$

On calcule les dérivées

$$y'(x) = \sum_{j \geq 0} c_j(j + \alpha)(x - x_0)^{j+\alpha-1}, \quad y''(x) = \sum_{j \geq 0} c_j(j + \alpha)(j + \alpha - 1)(x - x_0)^{j+\alpha-2}$$

et on les insert dans l'équation différentielle (6.7). En comparant les coefficients de $(x - x_0)^{k+\alpha}$, on voit que $y(x)$ de (6.10) est une solution de (6.7) si et seulement si pour tout $k \geq 0$

$$c_k(k + \alpha)(k + \alpha - 1) + \sum_{j=0}^k c_j(j + \alpha)p_{k-j} + \sum_{j=0}^k c_j q_{k-j} = 0.$$

Cette relation peut être écrite sous la forme

$$c_k f(k + \alpha) + \sum_{j=0}^{k-1} c_j \left((j + \alpha)p_{k-j} + q_{k-j} \right) = 0 \quad (6.11)$$

avec

$$f(\alpha) = \alpha(\alpha - 1) + \alpha p_0 + q_0. \quad (6.12)$$

Pour $k = 0$ on obtient $c_0 f(\alpha) = 0$ et donc $f(\alpha) = 0$ (car $c_0 \neq 0$). Comme $f(\alpha)$ est un polynôme de degré 2, on a en général deux choix possibles pour la valeur de α .

Lemme 6.3 *Soit c_0 arbitrairement donné et supposons que*

$$f(\alpha) = 0 \quad \text{et} \quad f(\alpha + k) \neq 0 \quad \text{pour} \quad k = 1, 2, \dots$$

Alors, la série $\sum_{k \geq 0} c_k (x - x_0)^k$ avec les coefficients c_k donnés par (6.11), possède un rayon de convergence $\geq \min(\rho_p, \rho_q)$, où ρ_p et ρ_q sont les rayons de convergence des séries (6.9).

Démonstration. Le fait que $f(\alpha + k)$ est un polynôme de degré 2 en k et l'hypothèse $f(\alpha + k) \neq 0$ pour $k = 1, 2, \dots$ impliquent que

$$|f(\alpha + k)| \geq \mu k^2 \quad \text{pour} \quad k = 1, 2, \dots$$

avec un $\mu > 0$. Fixons maintenant un $\rho < \min(\rho_p, \rho_q)$. On a alors

$$|p_j| \rho^j \leq M_p \quad \text{et} \quad |q_j| \rho^j \leq M_q \quad \text{pour} \quad j = 1, 2, \dots$$

et la relation (6.11) implique que pour $k \geq 1$

$$|c_k| \leq \frac{1}{\mu k^2} \sum_{j=0}^{k-1} |c_j| (|j + \alpha| M_p + M_q) \rho^{j-k} \leq \frac{M}{k} \sum_{j=0}^{k-1} |c_j| \rho^{j-k}$$

avec une constante $M \geq 1$. Par récurrence sur k , on en déduit que

$$|c_k| \leq |c_0| M k^{M-1} \rho^{-k} \quad \text{pour} \quad k = 1, 2, \dots$$

(utiliser l'inégalité $1 + M \sum_{j=1}^{k-1} j^{M-1} \leq k^M$). Ceci entraîne que le rayon de convergence de la série $\sum_{k \geq 0} c_k (x - x_0)^k$ est $\geq \rho$. \square

Si α_1 et α_2 (avec $\Re\alpha_1 \geq \Re\alpha_2$) sont les deux racines de l'équation $f(\alpha) = 0$, la fonction

$$y_1(x) = (x - x_0)^{\alpha_1} \sum_{j \geq 0} c_j(\alpha_1)(x - x_0)^j$$

est une solution de (6.7), car $f(\alpha_1 + k) \neq 0$ pour tout $k \geq 1$. Nous avons noté les coefficients c_j de (6.11) par $c_j(\alpha)$ pour indiquer la dépendance du paramètre α . Comment trouve-t-on une deuxième solution de (6.7) qui est linéairement indépendante de $y_1(x)$?

Cas 1. Si $\alpha_1 - \alpha_2 \notin \mathbb{Z}$, on est sûr que $f(\alpha_2 + k) \neq 0$ pour $k \geq 1$. Donc,

$$y_2(x) = (x - x_0)^{\alpha_2} \sum_{j \geq 0} c_j(\alpha_2)(x - x_0)^j$$

est une deuxième solution de (6.7). Elle est indépendante de $y_1(x)$, car les termes dominants possèdent des puissances différentes.

Cas 2. Si $\alpha_1 - \alpha_2 = m \in \mathbb{Z}$, on a que $f(\alpha_2 + m) = 0$ et la relation (6.11) ne permet pas de calculer c_m . Inspirés par les équations de Cauchy, où x^α et $x^\alpha \ln x$ sont des solutions si α est une racine double du polynôme caractéristique, nous cherchons une deuxième solution sous la forme

$$y_2(x) = C y_1(x) \ln(x - x_0) + (x - x_0)^{\alpha_2} \sum_{j \geq 0} d_j (x - x_0)^j. \quad (6.13)$$

Si l'on insert $y_2(x)$ dans l'équation différentielle (6.7), le terme avec le logarithme se simplifie et on peut déterminer d_j et C par comparaison des mêmes puissances de $(x - x_0)^{\alpha_2 + j}$.

Exemple 6.4 (Equation de Bessel) Appliquons cette procédure à l'équation de Bessel

$$x^2 y'' + x y' + (x^2 - n^2) y = 0, \quad (6.14)$$

où $n \in \mathbb{R}$, $n \geq 0$. En insérant (6.10) avec $x_0 = 0$ dans (6.14) on obtient

$$\sum_{j \geq 0} c_j(j + \alpha)(j + \alpha - 1)x^{j+\alpha} + \sum_{j \geq 0} c_j(j + \alpha)x^{j+\alpha} + (x^2 - n^2) \sum_{j \geq 0} c_j x^{j+\alpha} = 0.$$

Une comparaison des coefficients donne

$$c_0(\alpha(\alpha - 1) + \alpha - n^2) = 0 \quad (6.15)$$

$$c_1((1 + \alpha)\alpha + (1 + \alpha) - n^2) = 0 \quad (6.16)$$

$$c_j((j + \alpha)(j + \alpha - 1) + (j + \alpha) - n^2) + c_{j-2} = 0. \quad (6.17)$$

Comme $c_0 \neq 0$, l'équation (6.15) implique $\alpha^2 - n^2 = 0$, c.-à-d., $\alpha_1 = n$ et $\alpha_2 = -n$. En utilisant $\alpha^2 = n^2$, l'équation (6.16) devient $c_1(2\alpha + 1) = 0$ et on obtient $c_1 = 0$ à condition que $\alpha \neq -1/2$ (supposons pour le moment que ceci est le cas). L'équation (6.17) donne $c_j \cdot j(j + 2\alpha) + c_{j-2} = 0$. Le fait que $c_1 = 0$ implique que $c_j = 0$ pour tout j impair. Pour $j = 2k$ on a $c_{2k} \cdot 4k(k + \alpha) + c_{2k-2} = 0$ et on voit par récurrence que

$$c_{2k} = \frac{(-1)^k}{4^k k! \prod_{i=1}^k (i + \alpha)} c_0$$

si $\alpha \neq -1, -2, \dots$. Avec le choix $c_0 = (2^\alpha \Gamma(\alpha + 1))^{-1}$, où $\Gamma(\alpha + 1) = \int_0^\infty e^{-x} x^\alpha dx$ est la fonction gamma d'Euler (voir [HW95, Sect. III.8]), la fonction $y(x)$ de (6.10) devient

$$J_\alpha(x) = \left(\frac{x}{2}\right)^\alpha \sum_{k \geq 0} \frac{(-1)^k}{k! \Gamma(\alpha + k + 1)} \left(\frac{x}{2}\right)^{2k}. \quad (6.18)$$

Cette fonction s'appelle *fonction de Bessel* d'indice α . Avec l'interprétation $\frac{1}{\Gamma(0)} = \frac{1}{\Gamma(-1)} = \frac{1}{\Gamma(-2)} = \dots = 0$, elle est définie pour tout $\alpha \in \mathbb{R}$. Le critère du quotient

$$\left| \frac{c_{2k}}{c_{2(k+1)}} \right| = 4(k+1)(k+1+\alpha) \rightarrow \infty$$

montre que la série dans (6.18) converge pour tout $x \in \mathbb{R}$.

Par construction, $J_n(x)$ et $J_{-n}(x)$ sont des solutions de (6.14). Pour $n \notin \mathbb{Z}$, $n \geq 0$ on a que $J_n(x) \rightarrow 0$ et $J_{-n}(x) \rightarrow \infty$ si $x \rightarrow 0$. Les deux fonctions sont donc linéairement indépendantes et la solution générale de (6.14) est donnée par $C_1 J_n(x) + C_2 J_{-n}(x)$.

Pour $n \in \mathbb{Z}$ on a

$$J_{-n}(x) = (-1)^n J_n(x).$$

Cette formule est obtenue par un calcul direct:

$$\begin{aligned} J_{-n}(x) &= \left(\frac{x}{2}\right)^{-n} \sum_{k \geq n} \frac{(-1)^k}{k!(-n+k)!} \left(\frac{x}{2}\right)^{2k} = \left(\frac{x}{2}\right)^{-n} \sum_{j \geq 0} \frac{(-1)^{j+n}}{(j+n)!j!} \left(\frac{x}{2}\right)^{2j+2n} \\ &= (-1)^n \left(\frac{x}{2}\right)^n \sum_{j \geq 0} \frac{(-1)^{j+n}}{j!(n+j)!} \left(\frac{x}{2}\right)^{2j} = (-1)^n J_n(x). \end{aligned}$$

Une deuxième solution indépendante peut être trouvée avec l'ansatz

$$Y_n(x) = C J_n(x) \ln x + x^{-n} \sum_{j \geq 0} d_j x^j$$

(voir l'équation (6.13)).

IV.7 Exercices

1. Est-ce que le problème

$$y'' = h(x)$$

avec pour conditions aux limites périodiques $y(0) = y(1)$, $y'(0) = y'(1)$ possède une solution unique?

2. On considère le problème

$$(a_2(x)y')' + a_0(x)y = 0 \tag{7.1}$$

$$\alpha y(a) + \beta y'(a) = 0 \tag{7.2}$$

$$\gamma y(b) + \delta y'(b) = 0 \tag{7.3}$$

où $a_0(x)$ est continue, $a_2(x)$ est continûment différentiable et $a_2(x) \neq 0$ pour $x \in [a, b]$.

- (a) En utilisant le théorème de Liouville, montrer que pour deux solutions $u(x)$ et $v(x)$ de (7.1) on a

$$a_2(x) \left(u(x)v'(x) - v(x)u'(x) \right) = C = \text{const.} \tag{7.4}$$

- (b) Supposons que la solution de (7.1)-(7.3) soit unique et notons par $u(x)$ une solution de (7.1) satisfaisant (7.2), et par $v(x)$ une solution de (7.1) satisfaisant (7.3). Montrer que la fonction de Green est donnée par

$$G(x, t) = \begin{cases} \frac{1}{C} u(x)v(t) & \text{si } x \leq t \\ \frac{1}{C} u(t)v(x) & \text{si } x \geq t \end{cases}$$

où C est la constante de (7.4).

3. En utilisant le résultat de l'exercice précédent, calculer la fonction de Green de

$$\begin{aligned} xy'' + y' &= 0, & y(1) &= 0, & y(2) &= 0, \\ y'' + y &= 0, & y(a) &= 0, & y(b) &= 0. \end{aligned}$$

4. Calculer la fonction de Green de

$$y'' = 0, \quad y(0) + 2y(1) = 0, \quad 2y'(0) + y'(1) = 0.$$

Le résultat est $G(x, t) = |x - t|/2 + (x + t)/6 - 4/9$.

5. Soit le problème

$$a_2(x)y'' + a_1(x)y' + a_0(x)y = 0, \quad y(0) = 0, \quad y'(0) = 0$$

sur $[0, 1]$. Calculer le problème adjoint.

6. Soit le problème générale

$$Ly = a_2(x)y'' + a_1(x)y' + a_0(x)y = 0, \quad B_1y = 0, \quad B_2y = 0,$$

on suppose avoir existence et unicité de la solution. Démontrer que pour $\varphi \in C^\infty([a, b])$ satisfaisant $\varphi(x) \equiv 0$ en dehors de $(a + \epsilon, b - \epsilon)$ avec un $\epsilon > 0$, on a que

$$\int_a^b G(x, t)(L^*\varphi)(x) dx = \varphi(t).$$

Ceci est équivalent à $LG_t = \delta_t$ dans le sens des distributions où $G_t(x) = G(x, t)$.

7. On considère l'équation différentielle linéaire d'ordre 4

$$y^{(4)} = h(x), \quad x \in [0, 1], \tag{7.5}$$

et les conditions aux bords

$$B_1y = 0, \quad B_2y = 0, \quad B_3y = 0, \quad B_4y = 0,$$

où B_iy est une combinaison linéaire de $y(0), y'(0), y''(0), y'''(0), y(1), y'(1), y''(1), y'''(1)$. Etudier l'existence et l'unicité des solutions. Sous l'hypothèse d'existence et d'unicité, démontrer qu'il existe une unique fonction continue $G(x, t)$ sur $[0, 1] \times [0, 1]$ (la fonction de Green) telle que la solution soit donnée par

$$y(x) = \int_0^1 G(x, t)h(t)dt.$$

8. L'équation différentielle de la ligne élastique d'une poutre est donnée par (7.5), où $h(x)$ représente la charge. Calculer la fonction de Green pour les situations suivantes:

- (a) la poutre est fixée aux deux extrémités:

$$y(0) = y'(0) = 0, \quad y(1) = y'(1) = 0.$$

- (b) la poutre est librement posée aux deux extrémités:

$$y(0) = y''(0) = 0, \quad y(1) = y''(1) = 0.$$

- (c) la poutre est fixée à gauche, mais complètement libre à droite:

$$y(0) = y'(0) = 0, \quad y''(1) = y'''(1) = 0.$$

Résultat. $G(x, t)$ est symétrique et pour $t \leq x$ on a

(a) $G(x, t) = t^2(1 - x^2)(3x - t - 2xt)/6,$

(b) $G(x, t) = t(1 - x)(2x - x^2 - t^2)/6,$

(c) $G(x, t) = t^2(3x - t)/6.$

9. Calculer les valeurs propres et les fonctions propres du problème

$$(xy')' + \frac{\lambda}{x}y = 0, \quad y(1) = 0, \quad y(2) = 0.$$

10. Calculer les valeurs propres et les fonctions propres du problème

$$y'' + \lambda y = 0, \quad y(0) = y'(0), \quad y(1) = 0$$

et montrer que les valeurs propres satisfont $\sqrt{\lambda_k} = \pi/2 + k\pi + \beta_k$ avec $\beta_k \rightarrow 0$.

11. Transformation sous forme auto-adjointe. Considérons l'opérateur

$$Ly = a_2(x)y'' + a_1(x)y' + a_0(x)y,$$

où les $a_i(x)$ sont suffisamment différentiables sur l'intervalle considéré. Comment faut-il choisir $\varphi(x)$ pour que la transformation $y = \varphi(x)z$ nous conduise à un opérateur auto-adjoint?

12. Considérons

$$y'' + \frac{\cos^2 x}{1 + \cos^2 x}y - \lambda \frac{\cos x}{1 + \cos^2 x}y = 0, \quad y'(0) = 0, \quad y'(\pi) = 0.$$

Montrer que $y(x) = i + \cos x$ est une fonction propre pour la valeur propre $\lambda = i$.

Comment expliquer cette valeur propre complexe?

13. Résoudre par séparation des variables le problème

$$\begin{aligned} \frac{\partial u}{\partial t} &= \frac{\partial^2 u}{\partial x^2} - u, & 0 < x < 1, \quad t > 0 \\ \frac{\partial u}{\partial x}(t, 0) &= \frac{\partial u}{\partial x}(t, 1) = 0, & t > 0 \\ u(0, x) &= f(x), & 0 < x < 1 \end{aligned} \tag{7.6}$$

(la température d'une barre dont les extrémités sont isolées mais qui échange de la chaleur en chaque point avec le milieu extérieur).

14. L'opérateur différentiel

$$Ly = (p(x)y')' + q(x)y, \quad x \in (a, \infty)$$

($p \in \mathcal{C}^1(a, \infty)$ et $p(x) > 0$ sur (a, ∞) , $q \in \mathcal{C}^0(a, \infty)$), s'appelle *oscillatoire*, si chaque solution non-nulle de $Ly = 0$ possède une infinité de zéros dans (a, ∞) . En utilisant le théorème de comparaison de Sturm, montrer que L est oscillatoire si et seulement si il existe *une* solution non-nulle ayant une infinité de zéros dans (a, ∞) .

15. Déterminer les valeurs de α pour lesquelles l'opérateur

$$y'' + \frac{\alpha}{x^2}y, \quad \alpha \in \mathbb{R}$$

est oscillatoire.

16. Démontrer que l'opérateur différentiel

$$Ly = y'' + q(x)y$$

(a) est oscillatoire, si $\liminf_{x \rightarrow \infty} x^2 q(x) > 1/4$;

(b) n'est pas oscillatoire, si $\limsup_{x \rightarrow \infty} x^2 q(x) < 1/4$.

Indication. Comparer l'opérateur L avec celui de l'Exercice 15.

17. A l'aide de $y = \varphi(x)z$ transformer l'équation de Bessel

$$x^2 y'' + xy' + (x^2 - n^2)y = 0$$

sous la forme $z'' + q(x)z = 0$. Montrer que l'équation de Bessel est oscillatoire pour tout $n \in \mathbb{R}$, et calculer $J_n(x)$ pour $n = \pm 1/2$.

Résultat. $J_{1/2}(x) = \sqrt{2/\pi x} \sin x$, $J_{-1/2}(x) = \sqrt{2/\pi x} \cos x$.

18. Démontrer la formule de récurrence

$$J_{n+1}(x) - \frac{2n}{x}J_n(x) + J_{n-1}(x) = 0.$$

En utilisant l'Exercice 17, calculer $J_{3/2}(x)$.

19. On considère l'équation hypergéométrique

$$x(1-x)y'' + (c - (a+b+1)x)y' - aby = 0.$$

Trouver une solution dans un voisinage de $x = 0$.

Le résultat est

$$F(x; a, b, c) = \sum_{k \geq 0} \frac{(a)_k (b)_k}{(c)_k} \frac{x^k}{k!}$$

où $(a)_k = a(a+1) \cdots (a+k-1)$.

20. On considère l'équation différentielle

$$a_2(x)y'' + a_1(x)y' + a_0(x)y = 0 .$$

Pour étudier le comportement des solutions à l'infini utiliser la transformation $w = 1/x$, $z(w) = y(x)$. Démontrer que ∞ est un point singulier régulier de l'équation hypergéométrique. Est-ce aussi vrai pour l'équation de Bessel?

21. Justifier l'ansatz (6.13) pour une deuxième solution indépendante si $\alpha_1 - \alpha_2 = m \in \mathbb{Z}$ et $m \geq 0$. Montrer l'existence de constantes C et d_j tels que (6.13) est une solution de l'équation différentielle.

Chapitre V

Sous-variétés de \mathbb{R}^n

Dans ce chapitre nous donnons une application intéressante du théorème d'inversion locale (paragraphe I.7) et nous discutons la définition de sous-variétés de \mathbb{R}^n . Nous introduisons l'espace tangent et des équations différentielles sur des sous-variétés de \mathbb{R}^n . Pour plus des détails, nous référons aux livres [AMR88] et [Ar74].

V.1 Théorème du rang

Un changement de coordonnées peut être très utile pour l'étude des fonctions. Par exemple, si la fonction $f(x_1, x_2)$ ne dépend que de $x_1^2 + x_2^2$, il est naturel d'introduire des coordonnées polaires. Le but de ce paragraphe est de chercher, pour une fonction $y = f(x)$ donnée, des coordonnées ξ et η à la place de x et y , pour lesquelles la fonction f devient plus simple (par exemple linéaire).

Nous nous contentons d'étudier ce problème localement, c.-à-d., près d'un point a . Comme un changement de coordonnées est équivalent à un difféomorphisme, le problème se pose de la manière suivante: soient $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $a \in \mathbb{R}^n$ et $b = f(a) \in \mathbb{R}^m$ donnés, chercher un difféomorphisme local φ près de a ($\xi = \varphi(x)$) et un difféomorphisme local ψ près de b ($\eta = \psi(y)$), tels que $\psi \circ f \circ \varphi^{-1}$ est "plus simple" dans un voisinage de $0 \in \mathbb{R}^m$.

Théorème 1.1 (Théorème du rang) Soient $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ de classe \mathcal{C}^k (avec $k \geq 1$), $a \in \mathbb{R}^n$ et $b = f(a)$. Les propositions suivantes sont équivalentes:

- il existe des difféomorphismes locaux φ près de a et ψ près de b (de classe \mathcal{C}^k) tels que

$$(\psi \circ f \circ \varphi^{-1})(\xi_1, \dots, \xi_n) = (\xi_1, \dots, \xi_r, 0, \dots, 0)^T \quad \text{près de } \xi = 0; \quad (1.1)$$

- le rang de $f'(x)$ est constant (égal à r) dans un voisinage de a .

Démonstration. a) Supposons l'existence des difféomorphismes φ et ψ tels que (1.1) est satisfait. En dérivant (1.1) on obtient

$$\psi'(y) f'(x) (\varphi^{-1})'(\xi) = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}, \quad (1.2)$$

où $x = \varphi^{-1}(\xi)$, $y = f(x)$, I est la matrice identité de dimension r et les trois "0" sont tels que la matrice (1.2) possède la même dimension que $f'(x)$. Les matrices $(\varphi^{-1})'(\xi)$ et $\psi'(y)$ sont inversibles pour $\xi = 0$ et $y = b$, et donc aussi dans un voisinage de 0 et b , respectivement. Ceci implique que le rang de $f'(x)$ est constant et égal à r .

b) Considérons une fonction f pour laquelle le rang de $f'(a)$ est r . Après une éventuelle permutation des x_1, \dots, x_n et/ou des y_1, \dots, y_m , on peut supposer que la sous-matrice

$$A = \left(\frac{\partial f_i}{\partial x_j}(a) \right)_{i,j=1}^r \quad \text{est inversible.} \quad (1.3)$$

Définissons alors l'application $\xi = \varphi(x)$ par

$$\varphi_i(x) = \xi_i = \begin{cases} f_i(x_1, \dots, x_n) - b_i & \text{pour } i = 1, \dots, r \\ x_i - a_i & \text{pour } i = r + 1, \dots, n. \end{cases} \quad (1.4)$$

Comme la matrice Jacobienne

$$\varphi'(a) = \begin{pmatrix} A & * \\ 0 & I \end{pmatrix}$$

est inversible, (1.4) définit un difféomorphisme local près de a . Considérons la fonction $g(\xi) = (f \circ \varphi^{-1})(\xi)$, dont les premières r composantes sont données par $g_i(\xi) = \xi_i + b_i$. Comme le changement de coordonnées $x \leftrightarrow \xi$ ne change pas le rang, la matrice

$$g'(\xi) = \begin{pmatrix} I & 0 \\ C(\xi) & D(\xi) \end{pmatrix}$$

est de rang r pour tout ξ dans un voisinage de $\xi = 0$. Donc $D(\xi) \equiv 0$ et la fonction $g(\xi)$ ne dépend que de ξ_1, \dots, ξ_r . Finalement, définissons l'application $\eta = \psi(y)$ par

$$\psi_i(y) = \eta_i = \begin{cases} y_i - b_i & \text{pour } i = 1, \dots, r \\ y_i - g_i(y_1 - b_1, \dots, y_r - b_r) & \text{pour } i = r + 1, \dots, m. \end{cases} \quad (1.5)$$

La matrice Jacobienne

$$\psi'(b) = \begin{pmatrix} I & 0 \\ * & I \end{pmatrix}$$

est inversible et (1.5) définit donc un difféomorphisme local près de b . Par définition de φ et ψ , on voit que $\psi \circ f \circ \varphi^{-1}$ est donnée par (1.1). \square

Corollaire 1.2 (Théorème de la submersion) *Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ une submersion en $a \in \mathbb{R}^n$ de classe \mathcal{C}^k (avec $k \geq 1$), c.-à-d., $f'(a)$ est surjective (rang de $f'(a)$ est m). Alors, il existe un difféomorphisme local φ près de 0 (de classe \mathcal{C}^k) tel que*

$$(f \circ \varphi)(\xi_1, \dots, \xi_n) = f(a) + (\xi_1, \dots, \xi_m)^T.$$

Démonstration. Dans la démonstration du Théorème 1.1, il ne faut pas faire de permutation des y_1, \dots, y_m et la fonction ψ est une simple translation. La fonction φ de ce corollaire correspond à φ^{-1} dans le Théorème 1.1. \square

Corollaire 1.3 (Théorème de l'immersion) *Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ une immersion en $a \in \mathbb{R}^n$ de classe \mathcal{C}^k (avec $k \geq 1$), c.-à-d., $f'(a)$ est injective (rang de $f'(a)$ est n). Alors, il existe un difféomorphisme local ψ près de $b = f(a)$ (de classe \mathcal{C}^k) tel que*

$$(\psi \circ f)(x_1, \dots, x_n) = (x_1, \dots, x_n, 0, \dots, 0)^T. \quad (1.6)$$

Démonstration. On suppose (1.3) avec $r = n$ et on définit l'application $\eta = \psi(y)$ implicitement par

$$y_i = \begin{cases} f_i(\eta_1, \dots, \eta_m) & \text{pour } i = 1, \dots, n \\ f_i(\eta_1, \dots, \eta_m) - \eta_i & \text{pour } i = n + 1, \dots, m. \end{cases}$$

Comme la matrice Jacobienne de cette application est inversible, ψ est un difféomorphisme local près de b et il satisfait (1.6). \square

V.2 Définition des sous-variétés de \mathbb{R}^n

Il y a plusieurs manières de décrire une sous-variété lisse (courbe, surface, ...) de \mathbb{R}^n .

Théorème 2.1 Soit $\mathcal{M} \neq \emptyset$ un sous-ensemble de \mathbb{R}^n , $a \in \mathcal{M}$, et soit $0 \leq r \leq n$. Les propositions suivantes sont équivalentes:

- (i) il existe un voisinage $U \subset \mathbb{R}^n$ de a et une application différentiable $g : U \rightarrow \mathbb{R}^{n-r}$ avec $g(a) = 0$ et $g'(a)$ surjective (une submersion) tels que

$$\mathcal{M} \cap U = g^{-1}(0);$$

- (ii) il existe un voisinage $U \subset \mathbb{R}^n$ de a , un voisinage $V \subset \mathbb{R}^n$ de 0 et un difféomorphisme $\varphi : V \rightarrow U$ avec $\varphi(0) = a$ tels que

$$\mathcal{M} \cap U = \varphi((\mathbb{R}^r \times \{0\}) \cap V);$$

- (iii) il existe un voisinage $U \subset \mathbb{R}^n$ de a , un voisinage $W \subset \mathbb{R}^r$ de 0 et une application différentiable $\eta : W \rightarrow U$ avec $\eta(0) = a$ et $\eta'(0)$ injective (une immersion) tels que

$$\mathcal{M} \cap U = \eta(W)$$

et $\eta : W \rightarrow \mathcal{M} \cap U$ est un homéomorphisme.

Définition 2.2 Un ensemble $\mathcal{M} \neq \emptyset$ qui satisfait aux conditions du Théorème 2.1 pour tout $a \in \mathcal{M}$ s'appelle une *sous-variété* (lisse) dans \mathbb{R}^n de dimension r et de codimension $n - r$. L'application η s'appelle une *paramétrisation locale* de \mathcal{M} près de a .

Exemple 2.3 Considérons la surface de la sphère dans \mathbb{R}^3 près du pôle nord $a = (0, 0, 1)^T$. Elle peut être décrite par $n = 3$, $r = 2$ et

$$g(x, y, z) = x^2 + y^2 + z^2 - 1,$$

mais aussi par $\varphi(x, y, \zeta) = (x, y, \sqrt{1 - x^2 - y^2} - \zeta)^T$ et par $\eta(x, y) = \varphi(x, y, 0)$. Une autre possibilité est de considérer des coordonnées sphériques:

$$\varphi(\alpha, \beta, \rho) = ((1 + \rho) \sin \alpha \cos \beta, (1 + \rho) \sin \alpha \sin \beta, (1 + \rho) \cos \alpha)^T$$

et $\eta(\alpha, \beta) = \varphi(\alpha, \beta, 0)$.

Démonstration du Théorème 2.1. Considérons l'injection canonique j et la projection canonique π :

$$\begin{array}{ccccc} \mathbb{R}^r & \xrightarrow{j} & \mathbb{R}^r \times \mathbb{R}^{n-r} & \xrightarrow{\pi} & \mathbb{R}^{n-r} \\ x & \mapsto & (x, 0) & & \\ & & (y_1, y_2) & \mapsto & y_2 \end{array}$$

Sans entrer dans les détails, nous donnons l'idée de la démonstration.

(i) \Rightarrow (ii): Le théorème de la submersion (Corollaire 1.2) implique l'existence d'un difféomorphisme local φ tel que

$$(g \circ \varphi)(\xi_1, \dots, \xi_n) = (\xi_{r+1}, \dots, \xi_n)^T$$

(après une permutation des ξ_i). La condition $g(x) = 0$ est alors équivalente à $\xi_{r+1} = \dots = \xi_n = 0$ si $x = \varphi(\xi)$.

(ii) \Rightarrow (i): Il suffit de prendre $g = \pi \circ \varphi^{-1}$.

(iii) \Rightarrow (ii): Le théorème de l'immersion (Corollaire 1.3) implique l'existence d'un difféomorphisme local ψ tel que

$$(\psi \circ \eta)(\xi_1, \dots, \xi_r) = (\xi_1, \dots, \xi_r, 0, \dots, 0)^T.$$

Donc, $x \in \eta(W)$ est équivalent à $\psi(x) \in \mathbb{R}^r \times \{0\}$. L'hypothèse supplémentaire " $\eta : W \rightarrow \mathcal{M} \cap U$ est un homéomorphisme" exclut que, pour certains éléments de $\eta(W)$, la préimage soit loin de 0.

(ii) \Rightarrow (iii): On peut prendre $\eta = \varphi \circ j$. □

Remarque On ne peut pas omettre l'hypothèse " $\eta : W \rightarrow \mathcal{M} \cap U$ est un homéomorphisme" dans la caractérisation (iii) du Théorème 2.1. Par exemple, la fonction $\eta(t) = ((1 + 0.1t^2) \cos t, (1 + 0.1t^2) \sin t)$ est une immersion pour chaque t , mais l'image $\eta(\mathbb{R})$ n'est pas une sous-variété de \mathbb{R}^2 à cause des croisements de la courbe (Fig. 2.1, gauche). Même l'injectivité de $\eta(t)$ ne serait pas suffisante comme le montre le dessin de droite de la Fig. 2.1.



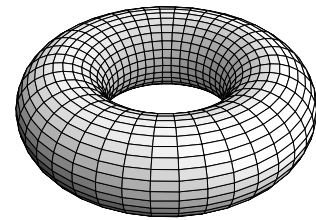
FIG. 2.1 – Sous-ensembles de \mathbb{R}^2 qui ne sont pas des sous-variétés

Les sous-variétés de dimension $r = 0$ sont des points discrets dans \mathbb{R}^n . Les sous-variétés de dimension $r = n$ sont des ouverts de \mathbb{R}^n . Tout sous-espace linéaire ou affine de \mathbb{R}^n est une sous-variété. Par contre, l'ensemble $\{(x, y) ; xy = 0\}$ n'est pas une sous-variété de \mathbb{R}^2 . Près de l'origine, cet ensemble n'est pas difféomorphe à une droite.

Exemple 2.4 (tore de révolution) Considérons le cercle $(x, z) = (d + \rho \cos \alpha, \rho \sin \alpha)$ (avec $0 < \rho < d$) et tournons-le autour de l'axe z . Ceci donne la paramétrisation

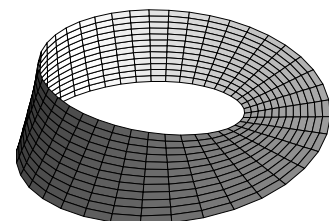
$$\eta(\alpha, \beta) = \begin{pmatrix} (d + \rho \cos \alpha) \cos \beta \\ (d + \rho \cos \alpha) \sin \beta \\ \rho \sin \alpha \end{pmatrix}$$

d'un tore. On peut vérifier que $\eta'(\alpha, \beta)$ est injective (c.-à-d., de rang 2) et que η est localement un homéomorphisme.



Exemple 2.5 (ruban de Möbius) Considérons une tige de longueur 2 (paramétrisée par $-1 < t < 1$) et tournons-la autour de son centre et, en même temps, deux fois plus vite autour d'un axe à distance d . Ceci donne la paramétrisation

$$\eta(t, \alpha) = \begin{pmatrix} (d + t \cos \alpha) \cos 2\alpha \\ (d + t \cos \alpha) \sin 2\alpha \\ t \sin \alpha \end{pmatrix}.$$



Exemple 2.6 (groupe orthogonal) L'ensemble

$$\mathcal{O}(n) = \{X ; X^T X = I\}$$

est une sous-variété de dimension $n(n-1)/2$ de $\mathcal{M}_n(\mathbb{R}) = \mathbb{R}^{n \times n}$ (espace de toutes les matrices de dimension n). Pour voir ceci, considérons l'application

$$g : \mathcal{M}_n(\mathbb{R}) \rightarrow \text{Sym}_n(\mathbb{R}) \approx \mathbb{R}^{n(n+1)/2}$$

définie par $g(X) = X^T X - I$ (le symbole $\text{Sym}_n(\mathbb{R})$ dénote l'espace des matrices symétriques de dimension n). On a bien $g^{-1}(0) = \mathcal{O}(n)$. Il faut alors voir que g est une submersion en tout point $A \in \mathcal{O}(n)$. La dérivée de $g(X)$ en A est $g'(A)H = A^T H + H^T A$. Pour une matrice symétrique B , le choix $H = AB/2$ montre que $g'(A)H = B$. Donc, $g'(A) : \mathcal{M}_n(\mathbb{R}) \rightarrow \text{Sym}_n(\mathbb{R})$ est surjective, et $\mathcal{O}(n)$ est une sous-variété de codimension $n(n+1)/2$.

V.3 Espace tangent

La tangente en un point a d'une courbe dans \mathbb{R}^2 ou \mathbb{R}^n est une droite, c.-à-d., un espace affine. En plaçant l'origine dans a , cette tangente devient un espace linéaire. Le plan tangent en un point a d'une surface est un espace de dimension 2. Un vecteur dans ce plan peut être interprété comme $\gamma'(0)$, où $\gamma(t)$ est une courbe différentiable dans la surface qui satisfait $\gamma(0) = a$. Ceci motive la définition suivante.

Définition 3.1 Soit $\mathcal{M} \subset \mathbb{R}^n$ une sous-variété et $a \in \mathcal{M}$. L'espace tangent à \mathcal{M} en a est

$$T_a \mathcal{M} = \left\{ h \in \mathbb{R}^n ; \begin{array}{l} \text{il existe } \gamma : (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}^n \text{ de classe } \mathcal{C}^1 \text{ telle que} \\ \gamma(t) \in \mathcal{M} \text{ pour } t \in (-\varepsilon, \varepsilon), \gamma(0) = a \text{ et } \gamma'(0) = h \end{array} \right\}.$$

Théorème 3.2 Soit $\mathcal{M} \subset \mathbb{R}^n$ une sous-variété et $a \in \mathcal{M}$.

a) si \mathcal{M} est donnée par une submersion $g : U \rightarrow \mathbb{R}^{n-r}$ (c.-à-d., $\mathcal{M} \cap U = g^{-1}(0)$), alors

$$T_a \mathcal{M} = \text{Ker } g'(a).$$

b) si \mathcal{M} est donnée par une paramétrisation $\eta : W \rightarrow \mathbb{R}^n$ (c.-à-d., $\mathcal{M} \cap U = \eta(W)$), alors

$$T_a \mathcal{M} = \text{Im } \eta'(0).$$

Démonstration. a) Pour une courbe $\gamma(t)$ avec $g(\gamma(t)) = 0$ on a que $g'(a)\gamma'(0) = 0$ et donc $T_a \mathcal{M} \subset \text{Ker } g'(a)$. En utilisant le théorème de la submersion (Corollaire 1.2), on peut voir que pour chaque $h \in \text{Ker } g'(a)$ il existe une courbe $\gamma(t)$ dans \mathcal{M} avec $\gamma(0) = a$ et $\gamma'(0) = h$.

b) Soit $\sigma(t)$ une courbe différentiable dans \mathbb{R}^r satisfaisant $\sigma(0) = 0$. Alors $\gamma(t) = \eta(\sigma(t))$ est une courbe dans \mathcal{M} qui satisfait $\gamma'(0) = \eta'(0)\sigma'(0)$. Donc $\text{Im } \eta'(0) \subset T_a \mathcal{M}$. Cette fois, le théorème de l'immersion (Corollaire 1.3) implique qu'on obtient toutes les courbes $\gamma(t)$ de cette manière. \square

V.4 Equations différentielles sur des sous-variétés

Soit $y(t)$ une courbe différentiable à valeurs dans une sous-variété $\mathcal{M} \subset \mathbb{R}^n$ de dimension r . Par définition de l'espace tangent, la dérivée $y'(t)$ satisfait $y'(t) \in T_{y(t)}\mathcal{M}$ pour tout t .

Définition 4.1 On dit que

$$y' = f(y) \quad (4.1)$$

est une *équation différentielle sur la sous-variété \mathcal{M}* , si $f(y) \in T_y\mathcal{M}$ pour tout $y \in \mathcal{M}$. Une solution est une fonction $y : I \rightarrow \mathcal{M} \subset \mathbb{R}^n$ qui vérifie $y'(t) = f(y(t))$ pour tout $t \in I$.

Pour étudier l'existence et l'unicité d'une solution, on utilise une paramétrisation locale $y = \eta(z)$ de la sous-variété \mathcal{M} près d'une valeur initiale $y_0 \in \mathcal{M}$, et on cherche une équation différentielle pour z dans \mathbb{R}^r . Pour ceci, on définit $z(t)$ par $y(t) = \eta(z(t))$ dans un voisinage de $z = 0$ et on obtient de (4.1) que

$$\eta'(z(t))z'(t) = f(\eta(z(t))), \quad z(0) = 0.$$

Une multiplication par $\eta'(z)^T$ et ensuite par $(\eta'(z)^T\eta'(z))^{-1}$ (cette matrice est inversible, car $\eta'(z)$ est injective) donne l'équation différentielle pour z

$$z' = \tilde{f}(z), \quad \tilde{f}(z) = (\eta'(z)^T\eta'(z))^{-1}\eta'(z)^T f(\eta(z)). \quad (4.2)$$

Si f est de classe \mathcal{C}^1 et si η est de classe \mathcal{C}^2 , la fonction \tilde{f} est de classe \mathcal{C}^1 et on peut appliquer la théorie locale du Chapitre III (existence, unicité, ...). Dès qu'on a trouvé une solution $z(t)$ de (4.2), le fait que $f(\eta(z)) \in T_{\eta(z)}\mathcal{M} = \text{Im } \eta'(z)$ implique que $y(t) = \eta(z(t))$ est une solution de (4.1).

Exemple 4.2 Considérons le problème

$$y_1' = y_1 + y_2 - y_1z, \quad y_2' = 2y_2 - y_2z, \quad 0 = y_1^2 + y_2^2 - 1, \quad (4.3)$$

où la variable de contrôle z doit être déterminée afin qu'on obtienne une équation différentielle sur la sous-variété $\mathcal{M} = \{(y_1, y_2) ; y_1^2 + y_2^2 = 1\}$. En dérivant la relation algébrique dans (4.3), on obtient

$$0 = 2y_1y_1' + 2y_2y_2' = 2y_1^2 + 2y_1y_2 + 2y_2^2 - 2(y_1^2 + y_2^2)z.$$

Ceci est la condition pour z qui garantit que $y' \in T_y\mathcal{M}$ pour tout $y \in \mathcal{M}$.

Une application très importante sont les *systèmes mécaniques soumis à des liaisons*. Supposons que l'énergie cinétique soit $T(q, \dot{q}) = \frac{1}{2}\dot{q}^T M \dot{q}$, l'énergie potentielle $U(q)$ et que le mouvement soit contraint par la relation $g(q) = 0$ où $g : \mathbb{R}^d \rightarrow \mathbb{R}^m$ ($m < d$). Les solutions du problème variationnel $\int L dt \rightarrow \min$ avec comme Lagrangien

$$L(q, \dot{q}) = T(q, \dot{q}) - U(q) - g(q)^T \lambda$$

décrivent le mouvement du système (sans démonstration). Les équations d'Euler-Lagrange $\frac{\partial L}{\partial q} - \frac{d}{dt}(\frac{\partial L}{\partial \dot{q}}) = 0$ (Chapitre II) donnent le système

$$\begin{aligned} \dot{q} &= v \\ M\dot{v} &= -U'(q) - G(q)^T \lambda \\ 0 &= g(q) \end{aligned} \quad (4.4)$$

où $G(q) = g'(q)$. Toutes les solutions de ce système satisfont $g(q) = 0$, mais aussi $G(q)v = 0$ (la dérivée totale de $g(q(t)) = 0$). Donc, elles sont des courbes dans la sous-variété

$$\mathcal{M} = \{(q, v) ; g(q) = 0, G(q)v = 0\} \quad (4.5)$$

de \mathbb{R}^{2d} . Il faut encore déterminer λ pour que $(\dot{q}, \dot{v}) \in T_{(q,v)}\mathcal{M}$. Pour ceci, il faut que la dérivée totale de $G((q(t))v(t)) = 0$ soit nulle:

$$0 = g''(q)(v, v) + G(q)M^{-1}(-U'(q) - G(q)^T \lambda). \quad (4.6)$$

En supposant $G(q)M^{-1}G(q)^T$ inversible, on peut calculer λ de cette relation et on obtient une équation différentielle pour (q, v) sur la sous-variété \mathcal{M} .

Exemple 4.3 (pendule simple) Dans ce cas l'énergie cinétique est $T = \frac{m}{2}(\dot{x}^2 + \dot{y}^2)$, l'énergie potentielle est $U = mgy$, et on a la liaison $g(x, y) = \frac{1}{2}(x^2 + y^2 - \ell^2) = 0$. Le système (4.4) devient alors

$$\begin{aligned} \dot{x} &= u, & m\dot{u} &= -x\lambda \\ \dot{y} &= v, & m\dot{v} &= -mg - y\lambda \end{aligned}$$

et on peut calculer λ de la relation

$$u^2 + v^2 - gy - m^{-1}(x^2 + y^2)\lambda = 0.$$

Exemple 4.4 (pendule double) Comme dans l'Exercice 45 du Chapitre III nous avons

$$T = \frac{m_1}{2}(\dot{x}_1^2 + \dot{y}_1^2) + \frac{m_2}{2}(\dot{x}_2^2 + \dot{y}_2^2), \quad U = m_1gy_1 + m_2gy_2$$

et les liaisons

$$x_1^2 + y_1^2 - \ell_1^2 = 0, \quad (x_2 - x_1)^2 + (y_2 - y_1)^2 - \ell_2^2 = 0. \quad (4.7)$$

Les équations du mouvement sont alors données par (en divisant les équations (4.7) par 2)

$$\begin{aligned} \dot{x}_1 &= u_1, & m_1\dot{u}_1 &= -x_1\lambda_1 + (x_2 - x_1)\lambda_2 \\ \dot{x}_2 &= u_2, & m_2\dot{u}_2 &= -(x_2 - x_1)\lambda_2 \\ \dot{y}_1 &= v_1, & m_1\dot{v}_1 &= -m_1g - y_1\lambda_1 + (y_2 - y_1)\lambda_2 \\ \dot{y}_2 &= v_2, & m_2\dot{v}_2 &= -m_2g - (y_2 - y_1)\lambda_2. \end{aligned}$$

Comme dans l'exemple précédent, on obtient λ_1, λ_2 de l'équation (4.6).

Observons que cette forme des équations du mouvement est beaucoup plus simple que celle obtenue dans l'Exercice 45 du Chapitre III pour les angles α et β . La différence est encore plus prononcée si le système mécanique est plus compliqué.

V.5 Exercices

1. Calculer le rang de la matrice jacobienne pour

$$f(x) = \begin{pmatrix} x_1 + x_2 + x_3 + x_4 \\ x_1^2 + x_2^2 + x_3^2 + x_4^2 \\ x_1^3 + x_2^3 + x_3^3 + x_4^3 \end{pmatrix}.$$

2. Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ de classe \mathcal{C}^1 . Montrer que si $\text{rang } f'(a) = r$, alors $\text{rang } f'(x) \geq r$ dans un voisinage de a . Donner un exemple pour lequel $\text{rang } f'(a) = r$ et $\text{rang } f'(x) > r$ pour $x \in U \setminus \{a\}$, où U est un voisinage de a .
3. Soit A une matrice (avec n colonnes et m lignes) de rang r . Montrer qu'il existe des matrices inversibles S et T telles que

$$SAT = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}, \quad (5.1)$$

où I est la matrice identité de dimension r . Calculer une décomposition de la forme (5.1) pour la matrice

$$A = \begin{pmatrix} 1 & -1 & 2 \\ -2 & 2 & -4 \end{pmatrix}.$$

4. Pour les ensembles définis ci-dessous, décider si ce sont des sous-variétés ou non (faire un dessin si possible):

$$\begin{array}{ll} \{(t, t^2) \in \mathbb{R}^2 ; t \in \mathbb{R}\} & \{(t, t^2) \in \mathbb{R}^2 ; t \geq 0\} \\ \{(t^2, t^3) \in \mathbb{R}^2 ; t \in \mathbb{R}\} & \{(t^2, t^3) \in \mathbb{R}^2 ; t \neq 0\} \\ \{(x, y) \in \mathbb{R}^2 ; x > 0, y > 0\} & \{(x, y, z) \in \mathbb{R}^3 ; x = y = z = 0\} \\ \{(x, y, z) \in \mathbb{R}^3 ; x^2 + y^2 - z^2 = 1\} & \{(x, y, z) \in \mathbb{R}^3 ; x^2 + y^2 - z^2 = 0\} \end{array}$$

5. Donner une application différentiable $g : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ telle que l'ensemble

$$\mathcal{M} = \{x \in \mathbb{R}^3 ; g(x) = 0\}$$

est une sous-variété de dimension 1 dans \mathbb{R}^3 , mais $g'(x)$ n'est surjective pour aucun point $x \in \mathcal{M}$.

6. Vérifier que l'ensemble $\{(x, y) ; xy = 0\}$ n'est pas une sous-variété de \mathbb{R}^2 , mais que $\{(x, y) ; xy = 0\} \setminus \{(0, 0)\}$ est une sous-variété.

7. Soient $X \subset \mathbb{R}^n$ et $Y \subset \mathbb{R}^m$ deux sous-variétés. Montrer que le produit

$$X \times Y = \{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m ; x \in X, y \in Y\}$$

est une sous-variété de $\mathbb{R}^n \times \mathbb{R}^m$, on l'appelle la variété produit.

8. Démontrer que l'ensemble

$$\{(\cos t + 2) \cos \lambda t, (\cos t + 2) \sin \lambda t, \sin t\} \in \mathbb{R}^3 ; t \in \mathbb{R}\} \quad (5.2)$$

est une sous-variété de \mathbb{R}^3 pour $\lambda = 2/13$ (voir le dessin). Pour $\lambda = \sqrt{2}$ l'ensemble (5.2) n'est pas une sous-variété et il est partout dense dans le tore

$$\{(\cos u + 2) \cos v, (\cos u + 2) \sin v, \sin u\}.$$

Indication. En utilisant des propriétés des fractions continues [HW95, Sect. I.6] montrer que l'ensemble $\{\ell + k\sqrt{2} ; \ell, k \in \mathbb{Z}\}$ est dense dans \mathbb{R} .

