

TOPOLOGIE ET ANALYSE

Notes de cours

J. HENRY
août 2006

Table des matières

| | | |
|----------|--|-----------|
| 1 | Espaces Métriques | 2 |
| 1.1 | Espaces métriques | 2 |
| 1.2 | Topologie d'un espace métrique | 4 |
| 1.3 | Sous ensembles d'un espace métrique | 6 |
| 1.4 | Suites dans un espace métrique | 8 |
| 1.5 | Limites | 10 |
| 1.6 | Continuité. Continuité uniforme | 11 |
| 1.7 | Homéomorphisme | 12 |
| 2 | Espaces vectoriels normés | 13 |
| 2.1 | Espaces vectoriels normés | 13 |
| 2.2 | Applications linéaires continues | 14 |
| 2.3 | Applications multilinéaires | 15 |
| 2.4 | Exemples | 16 |
| 3 | Compacité | 17 |
| 3.1 | Définitions | 17 |
| 3.2 | Propriétés | 17 |
| 3.3 | Fonctions continues sur un compact | 19 |
| 3.4 | Application aux espaces vectoriels normés de dimension finie | 19 |
| 3.5 | Application: théorème de d'Alembert-Gauss | 20 |
| 3.6 | Complément : caractérisation de Borel Lebesgue des compacts | 21 |
| 4 | Connexité | 23 |
| 4.1 | Définition et propriétés élémentaires | 23 |
| 4.2 | Connexes de \mathbb{R} | 24 |
| 4.3 | Connexité par arcs | 24 |
| 4.4 | Applications | 25 |
| 4.5 | Complément: composantes connexes | 26 |
| 5 | Espaces Complètes | 29 |
| 5.1 | Définitions et propriétés élémentaires | 29 |
| 5.2 | Propriétés des espaces complets | 29 |

| | | |
|-----------|--|-----------|
| 5.3 | Théorème du point fixe | 30 |
| 5.4 | Espaces de Banach | 31 |
| 6 | Suites et séries de fonctions à valeurs dans un espace vectoriel normé | 33 |
| 6.1 | Convergence simple et convergence uniforme d'une suite de fonctions | 33 |
| 6.2 | Conservation de la continuité | 34 |
| 6.3 | Echange de limites | 34 |
| 6.4 | Séries de fonctions à valeurs dans un Banach | 35 |
| 7 | Exemples d'espaces de Banach | 37 |
| 7.1 | L'espace $\mathcal{B}(X, E)$ | 37 |
| 7.2 | L'espace $\mathcal{CB}(X, E)$ | 37 |
| 7.3 | L'espace $L_c(E, F)$ | 38 |
| 7.4 | L'espace $l^1(\mathbb{K})$ | 38 |
| 8 | Approximation des fonctions à valeurs dans un espace vectoriel de dimension finie | 41 |
| 8.1 | Introduction et notations | 41 |
| 8.2 | Théorèmes d'approximation | 41 |
| 8.3 | Complément : polynômes de Bernstein | 43 |
| 8.4 | Complément : une autre démonstration du théorème de Weierstrass | 44 |
| 9 | Dérivation | 47 |
| 9.1 | Définition et propriétés élémentaires | 47 |
| 9.2 | Cas des fonctions numériques (rappels) | 47 |
| 9.3 | Opérations | 48 |
| 9.4 | Inégalité des accroissements finis et conséquences | 50 |
| 10 | Intégration des fonctions continues par morceaux sur un intervalle compact | 53 |
| 10.1 | Intégration des fonctions en escalier | 53 |
| 10.2 | Intégration des fonctions continues par morceaux | 54 |
| 10.3 | Cas des fonctions numériques | 54 |
| 10.4 | Sommes de Riemann | 56 |
| 10.5 | Intégration et primitives | 57 |
| 11 | Espaces préhilbertiens | 61 |
| 11.1 | Produit scalaire | 61 |
| 11.2 | Orthogonalité | 63 |
| 11.3 | Projection orthogonale sur un sev de dimension finie | 65 |
| 11.4 | Suites orthonormales | 65 |
| 11.5 | Polynômes orthogonaux | 67 |
| 11.6 | L'espace $l^2_{\mathbb{N}}(\mathbb{C})$ | 69 |
| 12 | Séries de Fourier | 71 |

| | | |
|-----------|--|------------|
| 12.1 | Rappels du cours d'Intégration | 71 |
| 12.2 | Polynômes et séries trigonométriques | 73 |
| 12.3 | Coefficients et série de Fourier | 74 |
| 12.4 | Densité des polynômes trigonométriques | 75 |
| 12.5 | Convergence en moyenne quadratique | 78 |
| 12.6 | Etude de la convergence uniforme | 80 |
| 12.7 | Etude de la convergence simple | 81 |
| 12.8 | Cas des fonctions T -périodiques | 83 |
| 13 | Calcul différentiel | 85 |
| 13.1 | Préliminaires : notations | 85 |
| 13.2 | Différentiabilité | 85 |
| 13.3 | Fonctions de classe C^1 | 95 |
| 13.4 | Exemples et applications | 96 |
| 13.5 | Inégalités des accroissements finis | 98 |
| 13.6 | Dérivées d'ordre supérieur | 99 |
| 13.7 | Formule de Taylor Young à l'ordre 2 | 102 |
| 13.8 | Problèmes d'extremums | 102 |
| | Appendice : preuve de quelques théorèmes | 107 |
| A | Preuve du théorème de caractérisation des fonctions de classe C^1 | 107 |
| B | Preuve du théorème de Schwarz | 108 |
| C | Preuve de la formule de Taylor | 109 |
| 14 | Difféomorphismes, fonctions implicites | 111 |
| 14.1 | Difféomorphismes | 111 |
| 14.2 | Le théorème d'inversion locale | 113 |
| 14.3 | Théorème des fonctions implicites | 114 |
| 15 | Equations et systèmes différentiels linéaires | 119 |
| 15.1 | Equations différentielles linéaires scalaires du premier ordre | 119 |
| 15.2 | Systèmes linéaires d'ordre 1 | 121 |
| 15.3 | Equations différentielles linéaires d'ordre n | 124 |
| 15.4 | Equations différentielles linéaires scalaires du second ordre | 126 |
| 15.5 | Appendice : preuve du théorème de Cauchy | 130 |
| 16 | Systèmes différentiels linéaires à coefficients constants | 133 |
| 16.1 | Exponentielle d'endomorphisme | 133 |
| 16.2 | Systèmes différentiels linéaires du premier ordre à coefficients constants | 137 |
| 16.3 | Equations différentielles linéaires à coefficients constants | 138 |
| 16.4 | Systèmes linéaires d'ordre 2 à coefficients réels | 140 |

| | |
|--|------------|
| 17 Equations différentielles | 143 |
| 17.1 Systèmes d'ordre 1 | 143 |
| 17.2 Equation différentielle scalaire du premier ordre | 144 |
| 17.3 Equations scalaires du second ordre | 147 |
| 17.4 Champs de vecteurs | 148 |

1

Espaces Métriques

Ces notes constituent un résumé de notions usuelles. Les définitions et énoncés sont donnés en détail. Les démonstrations de nombreuses affirmations, entre autre de celles précédées du signe ♠ sont laissées au lecteur. Le signe ■ indique la fin d'une démonstration.

1.1 Espaces métriques

1.1.1 Distance

DEFINITION 1.1.1

Soit X un ensemble non vide. On appelle distance sur X une application $d : X \times X \rightarrow \mathbb{R}$ vérifiant les trois propriétés suivantes:

- (1) $\forall x, y \in X, d(x, y) \geq 0$ et $d(x, y) = 0 \Leftrightarrow x = y$
- (2) $\forall x, y \in X, d(x, y) = d(y, x)$
- (3) $\forall x, y, z \in X, d(x, z) \leq d(x, y) + d(y, z)$ (inégalité triangulaire).

On appelle espace métrique un couple (X, d) où X est un ensemble non vide et d une distance sur X .

De (3), on déduit

$$(4) \forall x, y, z \in X, |d(x, z) - d(y, z)| \leq d(x, y)$$

EXEMPLE 1.1.1

1. \mathbb{R} ou \mathbb{C} muni de la distance $d(z, z') = |z - z'|$. Sauf mention explicite du contraire, on considérera toujours \mathbb{R} ou \mathbb{C} muni de cette distance.
2. La droite numérique achevée $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ peut être munie d'une distance $d(x, y) = |\arctan x - \arctan y|$, en convenant de poser $\arctan(-\infty) = -\pi/2$ et $\arctan(+\infty) = \pi/2$
3. Soit X un ensemble quelconque ayant au moins deux éléments ; l'application $e_X : X \times X \rightarrow \mathbb{R}$ définie par $e_X(x, x) = 0$ pour tout x et $e_X(x, y) = 1$ si $x \neq y$ est une distance sur X que l'on appellera distance discrète.

1.1.2 Boules

DEFINITION 1.1.2

Soit (X, d) un espace métrique, r un réel positif et $a \in X$.

On appelle boule ouverte de centre a et de rayon r l'ensemble $B(a, r) = \{x \in X \mid d(a, x) < r\}$.

On appelle boule fermée de centre a et de rayon r l'ensemble $B'(a, r) = \{x \in X \mid d(a, x) \leq r\}$

On appelle sphère de centre a et de rayon r l'ensemble $S(a, r) = \{x \in X \mid d(a, x) = r\}$.

On notera que si $r = 0$, $B(a, 0) = \emptyset$ et $B'(a, 0) = S(a, 0) = \{a\}$.

1.1.3 Sous espaces métriques

Soit (X, d) un espace métrique, Y un sous ensemble non vide de X . La restriction d_Y de d à $Y \times Y$ définit une distance sur Y . Le couple (Y, d_Y) est appelé le sous espace métrique Y de X . En pratique, on notera encore d au lieu de d_Y la distance ainsi définie sur Y .

Si $a \in Y$, la boule ouverte (resp. fermée) de centre a et de rayon r dans l'espace métrique Y est l'intersection avec Y de la boule ouverte (resp. fermée) de centre a et de rayon r dans X : $B_Y(a, r) = B(a, r) \cap Y$.

1.1.4 Parties bornées

♠ Soit (X, d) un espace métrique non vide et A une partie non vide de X . Les propriétés suivantes sont équivalentes:

1. $\exists a \in X, \exists r > 0, A \subset B(a, r)$
2. $\exists a \in X, \exists r > 0, A \subset B'(a, r)$
3. $\forall a \in X, \exists r > 0, A \subset B(a, r)$
4. $\forall a \in X, \exists r > 0, A \subset B'(a, r)$
5. La restriction de d à $A \times A$ est majorée.

On dit qu'une partie non vide A est bornée si elle vérifie l'une (et donc les cinq) propriétés ci dessus. Si A est une partie bornée non vide, le diamètre de A est $\text{diam}(A) = \sup_{x, y \in A} d(x, y)$. Par convention, la partie vide est bornée.

1.1.5 Distance d'un point à une partie

Soit (X, d) un espace métrique, A une partie non vide de X et $x \in X$. On appelle distance de x à A , le nombre $d(x, A) = \inf_{y \in A} d(x, y)$.

Si A et B sont deux sous ensembles non vides de X , la distance de A à B est le nombre $d(A, B) = \inf_{(x, y) \in A \times B} d(x, y)$. On a donc $d(x, B) = d(\{x\}, B)$.

Attention! Ceci n'est pas une distance sur l'ensemble des parties non vides de X .

1.1.6 Fonctions lipschitziennes

DEFINITION 1.1.3

Soient (X, d) et (Y, d') deux espaces métriques et f une application de X dans Y . Si k est un réel positif, on dit que f est k -lipschitzienne si $\forall x, x' \in X, d'(f(x), f(x')) \leq kd(x, x')$.

On dit que f est lipschitzienne si il existe un $k \geq 0$ telle que f soit k -lipschitzienne.

♠ La composée de deux applications lipschitziennes est lipschitzienne.

♠ Si A est une partie non vide d'un espace métrique (X, d) , l'application $x \rightarrow d(x, A)$ est 1-lipschitzienne.

1.1.7 Distances équivalentes

DEFINITION 1.1.4

Soit X un ensemble non vide et d et δ deux distances sur X . On dit que d et δ sont équivalentes (on précise quelquefois fortement équivalentes) si il existe des réels positifs a et b tels que $d \leq a\delta$ et $\delta \leq bd$.

Il revient au même de dire qu'il existe des réels strictement positifs k et k' tels que $k\delta \leq d \leq k'\delta$.

♠ Deux distances d et δ sur X sont équivalentes ssi les applications $Id_X : (X, d) \rightarrow (X, \delta)$ et $Id_X : (X, \delta) \rightarrow (X, d)$ sont lipschitziennes.

♠ La relation d est équivalente à δ est une relation d'équivalence sur l'ensemble des distances sur X .

♠ Soit (X', d') un espace métrique et d et δ deux distances équivalentes sur X . Alors $f : (X, d) \rightarrow (X', d')$ est lipschitziennne si et seulement si $f : (X, \delta) \rightarrow (X', d')$ est lipschitziennne.

1.1.8 Espaces produits

Soient p un entier, $p \geq 2$, (X_i, d_i) , $1 \leq i \leq p$, p espaces métriques et $X = X_1 \times X_2 \times \dots \times X_p$.

Pour $x = (x_1, x_2, \dots, x_p)$ et $y = (y_1, y_2, \dots, y_p)$ dans X , on définit

$$D_1(x, y) = \sum_{i=1}^p d_i(x_i, y_i); \quad D_2(x, y) = \sqrt{\sum_{i=1}^p d_i(x_i, y_i)^2} \quad \text{et} \quad D_\infty(x, y) = \sup_{1 \leq i \leq p} d_i(x_i, y_i).$$

D_1, D_2 et D_∞ sont des distances sur X qui sont équivalentes. On a en effet $D_\infty \leq D_1 \leq pD_\infty$ et $D_\infty \leq D_2 \leq \sqrt{p}D_\infty$.

Soit, pour chaque i , $F_i \subset E_i$ et $F = F_1 \times \dots \times F_p$. Notons δ_i la distance induite sur F_i par d_i . Sur l'espace produit F , on dispose des distances Δ_1, Δ_2 , et Δ_∞ fabriqués à partir des δ_i comme les D_1, D_2 , et D_∞ à partir des d_i . Les distances Δ_1, Δ_2 , et Δ_∞ sont égales aux distances induites par D_1, D_2 , et D_∞ sur $F \subset E$.

♠ Les projections $x = (x_1, \dots, x_p) \rightarrow x_i$ sont alors lipschitziennes pour chacune des distances ci dessus.

♠ Soit (X, d) un espace métrique. Lorsqu'on munit $X \times X$ de l'une quelconque des distances définies ci dessus, l'application $d : X \times X \rightarrow \mathbb{R}$ est lipschitziennne.

1.2 Topologie d'un espace métrique

Soit (X, d) un espace métrique.

1.2.1 Ouverts

DEFINITION 1.2.1

On dit qu'un sous ensemble Ω de X est ouvert si $\forall a \in \Omega, \exists r > 0, B(a, r) \subset \Omega$.

♠ Toute boule ouverte est un ouvert.

♠ Si $\Omega \subset X$, les trois propriétés suivantes sont équivalentes:

1. Ω est ouvert.
2. $\forall a \in \Omega, \exists n \in \mathbb{N}^*, B(a, \frac{1}{n}) \subset \Omega$
3. Ω est réunion de boules ouvertes.

THEOREME 1.2.1

Soit $\mathcal{O} = \mathcal{O}(X, d) \subset \mathcal{P}(X)$ la famille des ouverts de l'espace métrique (X, d) .

1. \mathcal{O} est stable par réunion quelconque.(i.e. toute réunion d'ouverts est un ouvert).
2. \mathcal{O} est stable par intersection **finie**.(i.e. toute intersection **finie** d'ouverts est un ouvert).
3. \emptyset et X appartiennent à \mathcal{O} . (L'ensemble vide et l'espace tout entier sont ouverts).

Remarque: On appelle topologie sur un ensemble X toute famille de parties de X vérifiant les propriétés 1), 2) et 3). Ainsi, la donnée d'une distance sur X définit une topologie sur X .

Attention. Il est faux qu'une intersection quelconque d'ouverts soit un ouvert.

THEOREME 1.2.2

Deux distances équivalentes d et δ sur l'ensemble X définissent la même topologie. Autrement dit, si d et δ sont équivalentes, on a $\mathcal{O}(X, d) = \mathcal{O}(X, \delta)$.

Preuve

Il suffit de montrer que $\delta \leq kd \Rightarrow \mathcal{O}(X, \delta) \subset \mathcal{O}(X, d)$ puis d'échanger les rôles de d et δ . Soit $\Omega \in \mathcal{O}(X, \delta)$. On veut montrer que Ω est un ouvert pour la topologie définie par d . C'est le cas si $\Omega = \emptyset$. Sinon, soit $a \in \Omega$. Puisque $\Omega \in \mathcal{O}(X, \delta)$, il existe $\rho > 0$ tel que $B_\delta(a, \rho) \subset \Omega$ où $B_\delta(a, \rho)$ est la boule ouverte de centre a de rayon ρ pour la distance δ . L'inégalité $\delta \leq kd$ implique l'inclusion $B_d(a, \rho/k) \subset B_\delta(a, \rho)$. On en déduit que $B_d(a, \rho/k) \subset \Omega$. ■

Attention. La réciproque est fautive. Deux distances sur le même ensemble X peuvent définir la même topologie sans être équivalentes.

Exercices:

1. Pour $x, y \in \mathbb{R}$, on pose $\delta(x, y) = \inf(|x - y|, 1)$. Montrer que δ est une distance sur \mathbb{R} , qu'elle définit la même topologie que la distance usuelle, mais qu'elle n'est pas équivalente à cette distance.
2. La topologie de \mathbb{R} considéré comme sous espace métrique de l'espace métrique $\overline{\mathbb{R}}$ (exemple 1.1.1) est la topologie usuelle.
3. Si X est un ensemble quelconque munie de la distance discrète e_X (définie exemple 1.1.1) toutes les parties de X sont ouvertes. On a donc $\mathcal{O}(X) = \mathcal{P}(X)$; la topologie ainsi définie s'appelle la topologie discrète sur X .

1.2.2 Fermés

DEFINITION 1.2.2

Une partie F de X est fermée si son complémentaire $F^c = X \setminus F$ est ouvert.

Du théorème 1.2.1 on déduit immédiatement :

THEOREME 1.2.3

Une intersection quelconque de fermés est un fermé. Une réunion **finie** de fermés est un fermé. \emptyset et X sont fermés.

♠ Toute boule fermée, toute sphère, tout singleton est une partie fermée.

1.2.3 Voisinages

DEFINITION 1.2.3

Soit A une partie quelconque de X . On appelle voisinage de A toute partie W de X qui contient un ouvert contenant A .

Si W et W' sont deux voisinages de A , $W \cap W'$ en est un. Si $a \in X$, on appelle voisinage de a tout voisinage de $A = \{a\}$. On notera $\mathcal{V}(a)$ l'ensemble des voisinages de a .

THEOREME 1.2.4

Soit (X, d) un espace métrique et Ω une partie de X . Ω est ouvert si et seulement si il est voisinage de chacun de ses points.

preuve

Soit Ω un ouvert; si $x \in \Omega$, Ω contient un ouvert (Ω lui-même) contenant x , donc $\Omega \in \mathcal{V}(x)$. Réciproquement, supposons que Ω soit voisinage de chacun de ses points et montrons que Ω est ouvert. Soit $a \in \Omega$; il existe un ouvert U tel que $a \in U$ et $U \subset \Omega$. Comme U est ouvert, il existe un $r > 0$ tel que $B(a, r) \subset U \subset \Omega$. Finalement, pour tout $a \in \Omega$, il existe un $r > 0$ tel que $B(a, r) \subset \Omega$ ce qui prouve que Ω est ouvert. ■

THEOREME 1.2.5

♠ Soient (X, d) un espace métrique, $a \in X$ et $\mathcal{V}(a)$ l'ensemble des voisinages de a . On a

1. $\forall V \in \mathcal{V}(a), a \in V$

2. $\forall V \in \mathcal{V}(a), \forall W \subset X, V \subset W \Rightarrow W \in \mathcal{V}(a)$
3. $\mathcal{V}(a)$ est stable par intersection finie.
4. Soit $V \in \mathcal{V}(a)$. Il existe $U \in \mathcal{V}(a), U \subset V$ tel que $\forall x \in U, U \in \mathcal{V}(x)$

On appelle système fondamental de voisinages de a (ou base de voisinages de a) tout sous ensemble $\mathcal{V}_0(a)$ de $\mathcal{V}(a)$ tel que, pour toute partie W de X on ait l'équivalence $W \in \mathcal{V}(a) \Leftrightarrow \exists U \in \mathcal{V}_0(a) U \subset W$.

Par exemple, l'ensemble des boules ouvertes (resp. fermées) de centre a et de rayon $1/n$ (n entier naturel strictement positif) est un système fondamental de voisinages de a .

THEOREME 1.2.6 (Séparation)

Soient a et a' deux points distincts de l'espace métrique (X, d) . Il existe un voisinage V de a et un voisinage V' de a' tels que $V \cap V' = \emptyset$. On dit que la topologie de X est séparée.

preuve

Prendre $V = B(a, r/2), V' = B(a', r/2)$ où $r = d(a, a')$.

1.2.4 Topologie d'un sous espace

Soit (X, d) un espace métrique et Y une partie non vide de X . Y est un espace métrique (comme sous espace de X , c.f. 1.1.3). Les ouverts (resp. les fermés) de Y sont les intersections avec Y des ouverts (resp. des fermés) de X .

Prouvons le pour les ouverts ; on sait que pour tout $a \in Y$ on a $B_Y(a, r) = B(a, r) \cap Y$. On en déduit facilement que si U est un ouvert de X alors $U \cap Y$ est un ouvert de Y . Montrons la réciproque ; soit W un ouvert de Y que l'on peut supposer non vide ; soit $x \in W$; il existe $r > 0$ tel que $B_Y(x, r) \subset W$. Notons $B_x = B(x, r)$ de sorte que $B_x \cap Y \subset W$. Soit $\Omega = \bigcup_{x \in Y} B_x$; c'est un ouvert de X et on a $W = \Omega \cap Y$. ■

De même, si $a \in Y$ et si $W \subset Y$, W est un voisinage de a dans Y si et seulement si il existe un voisinage V de a dans X tel que $W = V \cap Y$.

1.2.5 Topologie produit

Soient p un entier, $p > 1$, $(X_i, d_i)_{1 \leq i \leq p}$ p espaces métriques et $X = X_1 \times \cdots \times X_p$. Les trois distances usuelles (§1.1.8) sur X sont équivalentes donc définissent la même topologie, appelée topologie produit. Soit $\Omega \subset X$ et $a \in \Omega$. Ω est un voisinage de a pour la topologie produit si et seulement si il existe $r > 0$ tel que $B_{D_\infty}(a, r) \subset \Omega$, ce qui revient à dire que $B_1(a_1, r) \times \cdots \times B_p(a_p, r) \subset \Omega$, ou encore qu'il existe des voisinages V_i de a_i dans X_i tels que $V_1 \times \cdots \times V_p \subset \Omega$. Comme la notion de voisinage dans X_i ne dépend que de la topologie de X_i , on voit que la notion de voisinage dans X ne dépend que de ces topologies, ce qui justifie le nom de topologie produit donnée à la topologie de X .

Plus généralement, une notion qui ne dépend que de la topologie de X (et non de la distance définissant cette topologie) est dite topologique. Le fait pour un sous ensemble de X d'être fermé, par exemple, est une notion topologique. Par contre, le fait pour un sous ensemble d'être borné dépend de la distance (cf exercice 1 du §1.2.1). On dit que c'est une notion métrique.

Soient $F_i \subset E_i$ pour $i = 1, \dots, p$ et $F = F_1 \times \cdots \times F_p \subset E$. Il résulte de 1.1.8 que F considéré comme sous espace de E a la même topologie que F considéré comme produit des F_i .

1.3 Sous ensembles d'un espace métrique

1.3.1 Intérieur

DEFINITION 1.3.1

Soient (X, d) un espace métrique, $A \subset X$. On dit qu'un point $a \in X$ est intérieur à A si A est un voisinage de a . Ce qui revient à dire qu'il existe un $r > 0$ tel que $B(a, r) \subset A$.

On appelle intérieur de A et on note $\overset{\circ}{A}$ l'ensemble des points intérieurs à A .

PROPOSITION 1.3.1

L'intérieur de A est le plus grand ouvert (au sens de l'inclusion) de X contenu dans A . En particulier A est ouvert ssi $A = \overset{\circ}{A}$.

preuve

Soit a intérieur à A et $r > 0$ tel que $B(a, r) \subset A$; la boule $B(a, r)$ est un ouvert donc voisinage de chacun de ses points. Donc tout x de $B(a, r)$ est intérieur à A , soit $B(a, r) \subset \overset{\circ}{A}$. Ceci prouve que que l'intérieur de A est un ouvert de X . Si U est un ouvert de X contenu dans A , par définition tout point de U est intérieur à A donc $U \subset \overset{\circ}{A}$. $\overset{\circ}{A}$ est donc le plus grand ouvert contenu dans A .

Attention! Ecrire "le plus grand ouvert" ayant telle propriété (ici contenu dans A) nécessite de montrer que ce plus grand ouvert existe. La démonstration ci dessus prouve entre autres choses cette existence. On peut aussi montrer directement cette existence de la manière suivante: d'une part il existe des ouverts contenus dans A (l'ensemble vide) et d'autre part la réunion de tous les ouverts contenus dans A est encore un ouvert contenu dans A et c'est manifestement le plus grand (pour l'inclusion).

1.3.2 Adhérence

DEFINITION 1.3.2

Soient (X, d) un espace métrique et $A \subset X$. On dit que le point $a \in X$ est adhérent à A si tout voisinage de a rencontre A . On appelle adhérence de A et on note \overline{A} l'ensemble des points adhérents à A .

On a donc $a \in \overline{A} \iff (\forall V \in \mathcal{V}(a), V \cap A \neq \emptyset) \iff (\forall r > 0, \exists x \in A, d(a, x) < r)$.

PROPOSITION 1.3.2

L'adhérence de A est un fermé et c'est le plus petit fermé contenant A . En particulier A est fermé ssi $A = \overline{A}$.

preuve

Soit $U = X \setminus \overline{A}$. Il s'agit de montrer que U est ouvert. Soit $x \in U$. Par définition, x n'est pas adhérent à A donc il existe $r > 0$ tel que $B(x, r) \cap A = \emptyset$. Si y est un point quelconque de $B(x, r)$, $B(x, r)$ est un voisinage de y qui ne rencontre pas A . Par conséquent $y \notin \overline{A}$; autrement dit $B(x, r) \subset U$, ce qui prouve que U est ouvert donc que \overline{A} est fermé. Il est évident que $A \subset \overline{A}$. Donc \overline{A} est un fermé contenant A . Enfin, soit F un fermé contenant A . Si $x \notin F$, l'ouvert $X \setminus F$ est un voisinage de x qui ne rencontre pas A , donc $x \notin \overline{A}$ et par conséquent $\overline{A} \subset F$.

PROPOSITION 1.3.3

♠ Soit A une partie d'un espace métrique (X, d) . On a $\overline{A} = \{x \in X | d(x, A) = 0\}$.

Attention! L'adhérence de la boule ouverte $B(a, r)$ est contenue dans la boule fermée $B'(a, r)$. Mais dans un espace métrique quelconque on n'a pas l'égalité en général. Par exemple, soit $X = \mathbb{Z}$ muni de la distance induite par celle de \mathbb{R} . La boule ouverte de centre 0 et de rayon 1 est le singleton $\{0\}$ qui est fermé (comme tout singleton dans un espace métrique). Son adhérence est donc elle même. Mais la boule fermée de centre 0 de rayon 1 est $B'(0, 1) = \{-1, 0, 1\} \neq \overline{B(0, 1)} = \{0\}$.

PROPOSITION 1.3.4

♠ Soient (X, d) un espace métrique, $A \subset X$, $B = X \setminus A$. On a $\overset{\circ}{B} = X \setminus \overline{A}$ et $\overline{B} = X \setminus \overset{\circ}{A}$

PROPOSITION 1.3.5

Soient A et B deux parties d'un espace métrique (X, d) . On a :

$$(1) \quad \overline{A \cap B} = \overset{\circ}{A} \cap \overset{\circ}{B} \quad \text{et} \quad \overset{\circ}{A} \cup \overset{\circ}{B} \subset \overline{A \cup B}$$

$$(2) \quad \overline{A \cap B} \subset \overline{A} \cap \overline{B} \quad \text{et} \quad \overline{A \cup B} = \overline{A} \cup \overline{B}$$

preuve

(1) L'intérieur de $A \cap B$ est un ouvert contenu dans A donc dans l'intérieur de A et dans B donc dans l'intérieur de B ; par conséquent $\overline{A \cap B} \subset \overset{\circ}{A} \cap \overset{\circ}{B}$. Inversement, $\overset{\circ}{A} \cap \overset{\circ}{B}$ est un ouvert (intersection de deux ouverts) contenu dans A et dans B donc dans $A \cap B$. D'où $\overset{\circ}{A} \cap \overset{\circ}{B} \subset \overline{A \cap B}$.

De même, l'intérieur de A (resp. de B) est un ouvert contenu dans $A \cup B$; donc $\overset{\circ}{A} \cup \overset{\circ}{B} \subset \overline{A \cup B}$

(2) résulte de la proposition précédente, ou peut faire l'objet d'une démonstration directe.

Remarque

On n'a pas, en général, l'égalité entre $\overset{\circ}{A} \cup \overset{\circ}{B}$ et $\overset{\circ}{A \cup B}$ ni entre $\overline{A \cap B}$ et $\overline{A} \cap \overline{B}$. Un exemple est donné par $X = \mathbb{R}$, $A = \mathbb{Q}$ et $B = \mathbb{R} \setminus \mathbb{Q}$. L'intérieur de A et celui de B sont vides, alors que l'intérieur de $A \cup B$ est \mathbb{R} . D'autre part, $\overline{A \cap B} = \emptyset$, alors que $\overline{A} = \overline{B} = \mathbb{R}$.

DEFINITION 1.3.3

Soit (X, d) un espace métrique, $A \subset X$. On appelle frontière de A , et on note $Fr(A)$ ou ∂A l'ensemble $\overline{A} \cap \overline{X - A}$.

Un point x est dans la frontière de A si et seulement si tout voisinage V de x contient des points appartenant à A et des points n'appartenant pas à A . La frontière de A est un sous ensemble fermé de X .

DEFINITION 1.3.4

Soit (X, d) un espace métrique, $A \subset X$. On dit que A est partout dense dans X si $\overline{A} = X$.

Plus généralement, si $B \subset A \subset X$, on dit que B est dense dans A si $A \subset \overline{B}$.

A est partout dense dans X si et seulement si tout ouvert non vide de X rencontre A .

EXEMPLE 1.3.1

\mathbb{Q} et $\mathbb{R} \setminus \mathbb{Q}$ sont denses dans \mathbb{R} .

DEFINITION 1.3.5

Soit (X, d) un espace métrique, $A \subset X$. Un point a de X est un point d'accumulation de A si tout voisinage de a rencontre l'ensemble $A \setminus \{a\}$. Un point d'accumulation de A est donc nécessairement adhérent à A .

Un point a de A est dit isolé dans A si il existe un voisinage V de A tel que $V \cap A = \{a\}$.

- ♠ Soit $a \in X$. a est un point isolé de A si et seulement si $a \in A$ et a n'est pas un point d'accumulation de A .
- ♠ a est un point d'accumulation de A si et seulement si pour tout voisinage V de a , l'ensemble $V \cap A$ est infini.
- ♠ L'ensemble A' des points d'accumulation de A est un ensemble fermé et on a $\overline{A} = A \cup A'$.

1.4 Suites dans un espace métrique

1.4.1 Définition, suites extraites

Si X est un ensemble quelconque, une suite d'éléments de X est une application $u : \mathbb{N} \rightarrow X$. On note $u_n = u(n)$ ou aussi $u = (u_n)_{n \geq 0}$.

Etant donnée une suite u d'éléments de X , une sous suite (ou suite extraite) de u est une suite v de la forme $v = u \circ \varphi$ où φ est une application $\mathbb{N} \rightarrow \mathbb{N}$, strictement croissante. φ est donc une suite strictement croissante d'entiers naturels.

Si on note $n_k = \varphi(k)$, on a $v_k = u_{\varphi(k)} = u_{n_k}$. Toute sous suite d'une sous suite de u est une sous suite de u .

Remarque importante Si $\varphi : \mathbb{N} \rightarrow \mathbb{N}$ est une application strictement croissante, on a pour tout $k \in \mathbb{N}$, $\varphi(k) \geq k$.

EXEMPLE 1.4.1

Si u est une suite, les suites $(v_p = u_{2p})_{p \geq 0}$ et $(w_p = u_{2p+1})_{p \geq 0}$ sont des suites extraites de u .

1.4.2 Limites, valeurs d'adhérence

DEFINITION 1.4.1

Soit (X, d) un espace métrique, u une suite d'éléments de X et $\lambda \in X$. On dit que la suite u converge vers λ si pour tout voisinage V de λ il existe un entier N tel que $u_n \in V$ pour $n \geq N$. Il revient au même de dire que u_n appartient à V sauf pour

un nombre fini d'indices.

On dit que la suite u est convergente si il existe un élément λ , appelé limite de u tel que u converge vers λ .

Donc λ est limite de u si et seulement si

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall n \in \mathbb{N}, n \geq N \Rightarrow d(u_n, \lambda) < \varepsilon$$

♠ Si une suite est convergente, sa limite est unique. (utiliser le théorème 1.2.6). On peut donc écrire $\lambda = \lim u_n$.

♠ Si une suite u converge vers λ , toute sous suite de u converge aussi vers λ .

♠ Toute suite convergente est bornée.

♠ Si les suites (u_{2p}) et (u_{2p+1}) ont une même limite λ , la suite u converge vers λ .

DEFINITION 1.4.2

Soit (X, d) un espace métrique, u une suite d'éléments de X et $\lambda \in X$. On dit que λ est valeur d'adhérence de la suite u si pour tout voisinage V de λ et tout entier naturel N il existe un entier $n \geq N$ tel que $u_n \in V$.

Cela revient à dire que pour tout voisinage V de λ , il existe une infinité d'entiers n tels que $u_n \in V$.

λ est valeur d'adhérence de la suite u si et seulement si

$$\forall \varepsilon > 0, \forall N \in \mathbb{N}, \exists n \in \mathbb{N}, n \geq N \text{ et } d(u_n, \lambda) < \varepsilon$$

Attention! Etant donnée une suite u , on peut considérer l'ensemble A_0 des termes de la suite. Il faut prendre garde à ne pas confondre un point adhérent à A_0 et une valeur d'adhérence de la suite u . Cette dernière notion fait intervenir l'application $u : \mathbb{N} \rightarrow X$ alors que la première ne fait intervenir que l'ensemble $u(\mathbb{N})$. On a en fait :

PROPOSITION 1.4.1

Soit (X, d) un espace métrique, u une suite d'éléments de X . Posons, pour p entier, $A_p = \{u_n; n \geq p\}$. L'ensemble des valeurs d'adhérence de la suite u est $\bigcap_{n \geq 0} \overline{A_n}$.

En effet, λ est valeur d'adhérence de u si et seulement si pour tout n et tout V voisinage de λ on a $V \cap A_n \neq \emptyset$, c'est à dire si et seulement si pour tout n , $\lambda \in \overline{A_n}$.

Exercice

Soit λ une valeur d'adhérence de la suite (u_n) . On suppose que $\lambda \notin A_0 = \{u_n, n \geq 0\}$. Montrer que λ est un point d'accumulation de A_0 . Inversement, si λ est un point d'accumulation de A_0 (appartenant ou non à A_0), λ est une valeur d'adhérence de la suite.

PROPOSITION 1.4.2

Soit (X, d) un espace métrique, u une suite d'éléments de X et $\lambda \in X$. λ est valeur d'adhérence de la suite u si et seulement si il existe une sous suite de la suite u qui converge vers λ .

preuve

Soit λ une valeur d'adhérence de la suite u . On a donc :

$$(P) \quad \forall \varepsilon > 0, \forall N \in \mathbb{N}, \exists n \in \mathbb{N}, (n \geq N \text{ et } d(u_n, \lambda) < \varepsilon)$$

On va construire par récurrence, une suite strictement croissante d'entiers (n_k) telle que $d(u_{n_k}, \lambda) < \frac{1}{k+1}$ pour tout k . En appliquant la propriété (P) avec $\varepsilon = \frac{1}{k+1}$, et $N = 0$, on obtient un entier n_0 tel que $d(u_{n_0}, \lambda) < 1$. Supposons déterminés des entiers n_k pour $0 \leq k \leq p$ tels que $n_0 < n_1 < \dots < n_p$ et $d(u_{n_k}, \lambda) < \frac{1}{k+1}$ pour $0 \leq k \leq p$. On applique (P) avec $\varepsilon = \frac{1}{p+2}$ et $N = n_p + 1$. On obtient un entier $n_{p+1} > n_p$ tel que $d(u_{n_{p+1}}, \lambda) < \frac{1}{p+2}$. La suite $(u_n)_{n \geq 0}$ ainsi construite est extraite de la suite u et converge vers λ .

♠ Réciproquement, si il existe une suite extraite de u qui converge vers λ , λ est valeur d'adhérence de u .

1.4.3 Suites de Cauchy

DEFINITION 1.4.3

Soit (X, d) un espace métrique, u une suite d'éléments de X . On dit que la suite u est une suite de Cauchy si elle vérifie la propriété suivante:

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall p, q \in \mathbb{N}, (p \geq N \text{ et } q \geq N) \Rightarrow d(u_p, u_q) < \varepsilon$$

ce qui s'écrit aussi, de manière équivalente et plus facile à manipuler

$$\forall \varepsilon > 0, \exists N \in \mathbb{N}, \forall p, q \in \mathbb{N}, p \geq N \Rightarrow d(u_p, u_{p+q}) < \varepsilon$$

Toute suite convergente est de Cauchy.

Toute suite de Cauchy est bornée.

DEFINITION 1.4.4

On dit que l'espace métrique (X, d) est complet si toute suite de Cauchy d'éléments de X est convergente.

Les propriétés des espaces complets seront étudiées ultérieurement.

1.4.4 Caractérisation de l'adhérence par les suites

THEOREME 1.4.1

Soit (X, d) un espace métrique, A une partie de X et $a \in X$. Le point a est adhérent à A si et seulement si il existe une suite (x_n) de points de A qui converge vers a .

preuve

Supposons $a \in \bar{A}$. Pour tout entier n , l'ensemble $B(a, \frac{1}{n+1}) \cap A$ est non vide. Choisissons un point x_n dans cet ensemble. On obtient ainsi une suite de points de A telle que $d(x_n, a) < \frac{1}{n+1}$, donc qui converge vers a . Réciproquement, supposons qu'il existe une suite (x_n) de points de A convergeant vers a . Pour tout $r > 0$, il existe un entier N tel que $d(x_n, a) < r$ pour $n \geq N$. On a donc $x_N \in A \cap B(a, r)$ ce qui prouve que cet ensemble est non vide. Ceci étant valable pour tout $r > 0$, on a $a \in \bar{A}$.

COROLLAIRE 1.4.1

Soit (X, d) un espace métrique, A une partie de X . A est une partie fermée de X si et seulement si toute suite de points de A qui converge dans X a sa limite dans A .

COROLLAIRE 1.4.2

Soit (X, d) un espace métrique, A une partie de X et (x_n) une suite de points de A . Toute valeur d'adhérence de la suite (x_n) appartient à l'adhérence de A .

1.5 Limites

DEFINITION 1.5.1

Soient (X, d) et (Y, δ) deux espaces métriques. Soient A une partie de X , $f : A \rightarrow Y$ une application, a un point de X adhérent à A et λ un élément de Y . On dit que λ est limite de f en a si pour tout voisinage W de λ dans Y , il existe un voisinage V de a dans X tel que $f(V \cap A) \subset W$. Ce qui s'écrit aussi:

$$\forall \varepsilon > 0, \exists \eta > 0, \forall x \in A, d(x, a) < \eta \Rightarrow \delta(f(x), \lambda) < \varepsilon$$

La limite de f en a si elle existe est unique ce qui justifie l'écriture $\lambda = \lim_{x \rightarrow a} f(x)$.

Remarque

Cette définition appliquée à $X = \bar{\mathbb{R}}$, $A = \mathbb{N}$, et $a = \infty$ qui est bien adhérent à \mathbb{N} lorsqu'on munit $\bar{\mathbb{R}}$ de la distance définie exemple 1.1.1 2 redonne la définition de la limite d'une suite.

Supposons que Y soit un produit de p espaces métriques (Y_i, δ_i) , muni d'une des distances définissant la topologie produit. L'application f s'écrit $f = (f_1, \dots, f_p)$, où $f_i : A \rightarrow Y_i$. Si $\lambda = (\lambda_1, \dots, \lambda_p) \in Y$, on a $\lambda = \lim_{x \rightarrow a} f(x)$ si et seulement si, pour tout i entre 1 et p , $\lambda_i = \lim_{x \rightarrow a} f_i(x)$.

THEOREME 1.5.1 (Composition)

Soient X, Y et Z trois espaces métriques, A une partie de X , B une partie de Y , a un point de X adhérent à A , b un point de Y et c un point de Z . Soient $f : A \rightarrow Y$ et $g : B \rightarrow Z$ deux applications. On suppose que $f(A) \subset B$ et que $\lim_{x \rightarrow a} f(x) = b$. Alors b est adhérent à B . Si en outre $\lim_{y \rightarrow b} g(y) = c$, on a $\lim_{x \rightarrow a} g \circ f(x) = c$.

preuve

- Soit V un voisinage de b . Il existe un voisinage U de a tel que $f(U \cap A) \subset V$. Comme $f(A) \subset B$, on a en fait $f(U \cap A) \subset V \cap B$. Puisque a est adhérent à A , $U \cap A$ est non vide, donc $V \cap B$ est non vide, donc $b \in \overline{B}$.
- Supposons $\lim_{y \rightarrow b} g(y) = c$; soit W un voisinage de c . Il existe un voisinage V de b tel que $g(V \cap B) \subset W$. Si U est un voisinage de a tel que $f(U \cap A) \subset V$, on a $g \circ f(U \cap A) \subset W$.

Si f et g sont des applications de $A \subset X$ dans \mathbb{R} ou \mathbb{C} les théorèmes usuels sur la limite d'une somme, d'un produit, d'un quotient restent valables.

THEOREME 1.5.2 (Utilisation de suites)

Soient (X, d) et (Y, δ) deux espaces métriques, A une partie de X , $f : A \rightarrow Y$ une application et a un point de X adhérent à A . Alors f a une limite quand x tend vers a si et seulement si pour toute suite (x_n) de points de A convergeant vers a , la suite $(f(x_n))$ converge dans Y .

preuve

Si $f(x)$ tend vers λ quand x tend vers a , il est immédiat, par composition, que pour toute suite (x_n) de points de A convergeant vers a , la suite $(f(x_n))$ converge vers λ .

Montrons maintenant la réciproque. On suppose donc que pour toute suite (x_n) de points de A convergeant vers a , la suite $(f(x_n))$ converge dans Y .

- Soient (x_n) et (y_n) deux suites de points de A qui convergent vers a . Alors $\lim f(x_n) = \lim f(y_n)$.
En effet, définissons une suite (z_n) de points de A par $z_{2p} = x_p$ et $z_{2p+1} = y_p$. La suite (z_n) converge vers a donc la suite $(f(z_n))$ converge dans Y . Les suites de termes généraux $f(x_n) = f(z_{2n})$ et $f(y_n) = f(z_{2n+1})$ sont extraites de la suite $(f(z_n))$ donc convergent vers la même limite.
- Notons désormais λ cette limite commune, et montrons par l'absurde que $\lim_{x \rightarrow a} f(x) = \lambda$. Si ceci est en défaut, il existe un $\varepsilon_0 > 0$ tel que pour tout $\eta > 0$ il existe un $x \in A$ vérifiant $d(x, a) < \eta$ et $\delta(f(x), \lambda) \geq \varepsilon_0$. En appliquant ceci avec $\eta = \frac{1}{n+1}$, pour $n \in \mathbb{N}$, on obtient un point $x_n \in A$ tel que $d(x_n, a) < \frac{1}{n+1}$ et $\delta(f(x_n), \lambda) \geq \varepsilon_0$. (x_n) est une suite de points de A qui converge vers a et la suite $f(x_n)$ ne converge pas vers λ . L'hypothèse que f ne tend pas vers λ quand x tend vers a conduit à une contradiction, donc elle est fautive.

1.6 Continuité. Continuité uniforme

DEFINITION 1.6.1

Soient (X, d) et (Y, δ) deux espaces métriques et $f : X \rightarrow Y$. On dit que f est continue en $a \in X$ si $\lim_{x \rightarrow a} f(x) = f(a)$.

On dit que f est continue sur X si elle est continue en tout point de X .

Si f est lipschitzienne, elle est continue.

Soit g une application de Y dans un troisième espace métrique; si f est continue en a et g en $f(a)$, l'application $g \circ f$ est continue en a .

Soient (Y_i, δ_i) p espaces métriques ($i = 1, \dots, p$) et Y leur produit muni de l'une des trois distances usuelles définissant la topologie produit. Les projections $\pi_i : Y \rightarrow Y_i$ sont continues. Une application $f : X \rightarrow Y$ est continue en a si et seulement si les p applications $f_i = \pi_i \circ f : X \rightarrow Y_i$ sont continues en a .

THEOREME 1.6.1 (Caractérisation de la continuité en un point)

Soient (X, d) et (Y, δ) deux espaces métriques et $f : X \rightarrow Y$. f est continue en $a \in X$ si et seulement si pour toute suite (x_n) de points de X convergeant vers a , la suite $(f(x_n))$ converge vers $f(a)$.

Résulte immédiatement du théorème 1.5.2.

THEOREME 1.6.2 (Caractérisation de la continuité globale)

Soient (X, d) et (Y, δ) deux espaces métriques et $f : X \rightarrow Y$. Les propriétés suivantes sont équivalentes:

1. f est continue sur X .

2. L'image réciproque par f de tout ouvert de Y est un ouvert de X .

3. L'image réciproque par f de tout fermé de Y est un fermé de X .

preuve

- (1) \Rightarrow (2) Soit V un ouvert de Y et $U = f^{-1}(V)$. Soit $x_0 \in U$. V est un voisinage de $y_0 = f(x_0)$ et f est continue en x_0 , donc il existe un voisinage W de x_0 tel que $f(W) \subset V$. On en déduit $W \subset U$ ce qui prouve que U est un voisinage de x_0 . U est voisinage de chacun de ses points, donc est ouvert.
- (2) \Rightarrow (1) Soit $x_0 \in X$, $y_0 = f(x_0)$ et V un voisinage de y_0 . Il existe un ouvert ω de Y tel que $y_0 \in \omega \subset V$. Par hypothèse $U = f^{-1}(\omega)$ est un ouvert, contenant x_0 , donc un voisinage de x_0 qui vérifie $f(U) \subset V$, ce qui prouve la continuité de f en x_0 .
- (2) \Leftrightarrow (3) résulte de la relation $f^{-1}(Y \setminus B) = X \setminus f^{-1}(B)$ pour tout sous ensemble B de Y .

Exercice

Montrer l'équivalence : f continue sur X si et seulement si pour toute partie A de X $f(\overline{A}) \subset \overline{f(A)}$.

DEFINITION 1.6.2

Soient (X, d) et (Y, δ) deux espaces métriques et $f : X \rightarrow Y$. On dit que f est uniformément continue sur X si

$$\forall \varepsilon > 0, \exists \eta > 0, \forall x, x' \in X, d(x, x') < \eta \Rightarrow \delta(f(x), f(x')) < \varepsilon$$

Le point important étant bien entendu que le réel η ne dépend que de ε (et de f) mais pas du point x .

- ♠ Toute application lipschitzienne est uniformément continue.
- ♠ Une application uniformément continue est continue.
- ♠ La composée de deux applications uniformément continues est uniformément continue.
- ♠ Si (x_n) est une suite de Cauchy de X et si f est uniformément continue, la suite $(f(x_n))$ est de Cauchy de Y .

Remarque

La notion de continuité est une notion topologique ; celle de continuité uniforme n'est pas une notion topologique, mais une notion métrique qui dépend essentiellement des distances utilisées.

Exercice

Soit $Y = \mathbb{R}$ muni de la distance usuelle. Soit $X = \mathbb{R}_+^*$. On peut munir X de deux distances: la distance d_1 induite par celle de \mathbb{R} : $d_1(x, x') = |x - x'|$ et la distance d_2 définie par $d_2(x, x') = |x^2 - x'^2|$. Vérifier que ceci définit bien une distance sur X . Montrer que la topologie définie par cette distance est la topologie usuelle. Soit $f : X \rightarrow Y$ l'application $f(x) = x^2$. Montrer que si on munit X de la distance usuelle d_1 , l'application f n'est pas uniformément continue. Montrer que si on munit X de la distance d_2 , f est uniformément continue.

1.7 Homéomorphisme

DEFINITION 1.7.1

Soient (X, d) et (Y, δ) deux espaces métriques. On appelle homéomorphisme de X sur Y toute bijection $f : X \rightarrow Y$ telle que f et la bijection réciproque f^{-1} soient continues. Deux espaces métriques sont dits homéomorphes si il existe un homéomorphisme de l'un sur l'autre.

La composée de deux homéomorphismes est un homéomorphisme. Soit (X, d) un espace métrique. L'ensemble des homéomorphismes de X sur lui même, muni de la composition des applications est un groupe.

Soient d_1 et d_2 deux distances sur X ; d_1 et d_2 définissent la même topologie sur X si et seulement si l'application identité $Id_X : (X, d_1) \rightarrow (X, d_2)$ est un homéomorphisme.

Soit f est un homéomorphisme de (X, d) sur (Y, δ) .

L'application f_* qui à une partie A de X associe la partie $f_*(A) = \{f(x) ; x \in A\}$ (notée habituellement $f(A)$) induit une bijection de l'ensemble $\mathcal{O}(X, d)$ des ouverts de X sur l'ensemble $\mathcal{O}(Y, \delta)$ des ouverts de Y . De même pour l'ensemble des fermés.

2

Espaces vectoriels normés

Dans tout ce qui suit \mathbb{K} désigne le corps \mathbb{R} des nombres réels ou celui \mathbb{C} des nombres complexes.

2.1 Espaces vectoriels normés

2.1.1 Norme et distance associée

DEFINITION 2.1.1

Soit E un \mathbb{K} -espace vectoriel. On appelle norme sur E une application $n : E \rightarrow \mathbb{R}$ vérifiant

1. $\forall x \in E, n(x) \geq 0$ et $n(x) = 0 \Leftrightarrow x = 0_E$
2. $\forall \lambda \in \mathbb{K}, \forall x \in E, n(\lambda x) = |\lambda| n(x)$
3. $\forall x, y \in E, n(x + y) \leq n(x) + n(y)$. (Inégalité triangulaire).

Un \mathbb{K} -espace vectoriel normé (\mathbb{K} -evn) est un couple (E, n) où E est un \mathbb{K} -ev et n une norme sur E . Soient (E, n) un \mathbb{K} -evn et F un sous espace vectoriel de E . La restriction n_F de n à F est une norme sur F , que l'on notera encore n sauf mention expresse du contraire. (F, n) est un sous espace normé de (E, n) .

Soit n une norme sur le \mathbb{K} -ev E . L'application $d : E \times E \rightarrow \mathbb{R}$ définie par $d(x, y) = n(x - y)$ est une distance sur E , dite distance associée à n . Muni de cette distance E devient un espace métrique, et est donc muni d'une topologie, qui sera dite topologie de l'evn (E, n) . Si (F, n_F) est un sous espace normé de E , la distance sur F associée à la norme n_F est la distance induite sur F par la distance de E . Muni de cette distance, F est donc un sous espace métrique de E .

2.1.2 Normes équivalentes

DEFINITION 2.1.2

Deux normes n_1 et n_2 sur un même \mathbb{K} -ev E sont équivalentes si il existe deux réels a et b positifs tels que

$$\forall x \in E, n_1(x) \leq a n_2(x) \text{ et } n_2(x) \leq b n_1(x)$$

Il revient au même de dire qu'il existe des réels **strictement positifs** k et k' tels que $k n_1 \leq n_2 \leq k' n_1$.

Les normes n_1 et n_2 sont équivalentes si et seulement si les distances associées le sont.

2.1.3 Norme sur un produit

Soient (E_i, n_i) , $1 \leq i \leq p$, des \mathbb{K} -evn et $E = E_1 \times E_2 \times \dots \times E_p$ le \mathbb{K} -ev produit. Si $x = (x_1, \dots, x_p) \in E$, on définit $N_1(x) = \sum_{1 \leq i \leq p} n_i(x_i)$, $N_2(x) = \sqrt{\sum_{1 \leq i \leq p} n_i^2(x)}$ et $N_\infty(x) = \sup_{1 \leq i \leq p} n_i(x_i)$. N_1 , N_2 et N_∞ sont trois normes équivalentes dont les distances associées sont les distances D_1 , D_2 et D_∞ définies sur le produit E à partir des distances d_i associées aux

normes n_i . (cf 1.1.8). Les topologies définies par ces trois normes sur le produit sont donc les mêmes. E muni d'une de ces normes sera appelé produit des evn E .

Désormais on utilisera souvent le symbole $\| \cdot \|$ pour désigner une norme.

PROPOSITION 2.1.1

Soit $(E, \| \cdot \|)$ un \mathbb{K} -evn.

- L'application $E \rightarrow \mathbb{R}, x \rightarrow \|x\|$ est 1-lipschitzienne.
- Si on munit $E \times E$ d'une des normes produit, l'application $E \times E \rightarrow E, (x, y) \rightarrow x + y$ est lipschitzienne.
- Si on munit $\mathbb{K} \times E$ d'une des normes produit, l'application $\mathbb{K} \times E \rightarrow E, (\lambda, x) \rightarrow \lambda x$ est continue.

2.1.4 Exemples

1. Sur \mathbb{R} ou \mathbb{C} l'application valeur absolue (ou module) est une norme.
2. Sur \mathbb{R}^p (resp. \mathbb{C}^p) on définit les trois normes équivalentes

$$\|x\|_1 = \sum_{1 \leq i \leq p} |x_i|, \quad \|x\|_2 = \sqrt{\sum_{1 \leq i \leq p} |x_i|^2} \quad \text{et} \quad \|x\|_\infty = \max_{1 \leq i \leq p} |x_i| \quad (\text{où } x = (x_1, \dots, x_p))$$

La deuxième est la norme euclidienne (resp. hermitienne).

3. Soit X un ensemble non vide et $E = \mathcal{B}(X, \mathbb{C})$ l'espace vectoriel des fonctions définies sur X , à valeurs dans \mathbb{C} et bornées. L'application $f \rightarrow \|f\|_\infty = \sup_{x \in X} |f(x)|$ est une norme sur E appelée norme de la convergence uniforme.
4. Soit $X = [a, b]$ un intervalle fermé borné de \mathbb{R} ($a < b$) et $E = C([a, b], \mathbb{C})$ le \mathbb{C} -espace vectoriel des fonctions continues sur X à valeurs dans \mathbb{C} . Soit $f \in E$. Posons $\|f\|_1 = \int_a^b |f(t)| dt$, $\|f\|_2 = \sqrt{\int_a^b |f(t)|^2 dt}$ et $\|f\|_\infty = \sup_{a \leq t \leq b} |f(t)|$. On définit ainsi trois normes sur E et ces normes **ne sont pas équivalentes**. La première s'appelle norme de la convergence en moyenne, la seconde norme de la convergence quadratique et la dernière norme de la convergence uniforme.
5. Soit $l^1(\mathbb{N})$ l'ensemble des suites $x = (x_n)_{n \geq 0}$ de nombres complexes telles que la série $\sum |x_n|$ converge. On définit une norme sur $l^1(\mathbb{N})$ en posant $\|x\| = \sum_{n \geq 0} |x_n|$.

2.2 Applications linéaires continues

THEOREME 2.2.1

Soient E et F deux \mathbb{K} -evn et $f : E \rightarrow F$ linéaire. Les propriétés suivantes sont équivalentes :

1. f est continue.
2. f est continue en 0_E .
3. Il existe un réel $c \geq 0$ tel que $\forall x \in E \quad \|f(x)\|_F \leq c \|x\|_E$.
4. f est bornée sur la boule unité fermée de E .
5. f est bornée sur la sphère unité de E .

preuve

- (2) \Rightarrow (4): la continuité de f en 0_E implique l'existence d'un $r > 0$ tel que $\forall z \in E, \|z\|_E \leq r \Rightarrow \|f(z)\|_F \leq 1$. Si x est un élément de E tel que $\|x\|_E \leq 1, z = rx$ vérifie $\|z\|_E \leq r$, donc $\|f(z)\|_F \leq 1$ soit $\|f(x)\|_F \leq 1/r$.
- (5) \Rightarrow (3) si f est bornée sur la sphère unité de E il existe $k \geq 0$ tel que $\forall z \in E, \|z\|_E \leq 1 \Rightarrow \|f(z)\|_F \leq k$. Si x est un élément de E non nul, $z = \frac{x}{\|x\|_E}$ est de norme 1, donc $\|f(z)\|_F \leq k$ soit $\|f(x)\|_F \leq k \|x\|_E$. Cette inégalité subsiste si $x = 0$.

- Les implications (1) \Rightarrow (2), (4) \Rightarrow (5) sont triviales. Enfin, f étant linéaire, (3) implique que f est c lipschitzienne et en particulier continue. Donc (3) \Rightarrow (1) et le théorème est prouvé. ■

On notera $L_c(E, F)$ l'ensemble des applications linéaires continues de E dans F . C'est un \mathbb{K} -ev.

Soient $k > 0$, $f \in L_c(E, F)$, $B = \{x \in E ; \|x\|_E \leq 1\}$, et $S = \{x \in E ; \|x\|_E = 1\}$. Il résulte immédiatement de la démonstration précédente que:

$$\forall x \in B, \|f(x)\|_F \leq k \Rightarrow \forall x \in S, \|f(x)\|_F \leq k \Rightarrow \forall x \in E, \|f(x)\|_F \leq k \|x\|_E \Rightarrow \forall x \in B, \|f(x)\|_F \leq k$$

On obtient donc l'égalité des quatre nombres suivants:

$$\sup_{\|x\|_E=1} \|f(x)\|_F = \sup_{\|x\|_E \leq 1} \|f(x)\|_F = \sup_{x \neq 0} \frac{\|f(x)\|_F}{\|x\|_E} = \inf \{k \geq 0 ; \forall x \in E, \|f(x)\|_F \leq k \|x\|_E\}$$

THEOREME 2.2.2

L'application qui à $f \in L_c(E, F)$ associe $\|f\| = \sup_{\|x\|_E=1} \|f(x)\|_F$ est une norme sur $L_c(E, F)$. On a pour tout $x \in E$, $\|f(x)\|_F \leq \|f\| \cdot \|x\|_E$ et $\|f\|$ est la plus petite constante positive k telle que pour tout $x \in E$, $\|f(x)\|_F \leq k \cdot \|x\|_E$.

La deuxième affirmation résulte de ce qui précède ; la preuve de la première est laissée au lecteur. ■

THEOREME 2.2.3

Soient E, F, G , trois \mathbb{K} -evn et $f \in L_c(E, F)$, $g \in L_c(F, G)$; alors $g \circ f \in L_c(E, G)$ et $\|g \circ f\| \leq \|g\| \cdot \|f\|$.

preuve

Soit $x \in E$. On a $\|(g(f(x)))\|_G \leq \|g\| \cdot \|f(x)\|_F$ et $\|f(x)\|_F \leq \|f\| \cdot \|x\|_E$ donc $\|g \circ f(x)\|_G \leq \|g\| \cdot \|f\| \cdot \|x\|_E$, ce qui avec le théorème 2.2.2 prouve le résultat. ■

On en déduit que si $u \in L_c(E)$, on a pour tout $n \in \mathbb{N}^*$, $\|u^n\| \leq \|u\|^n$ (où $u^n = \underbrace{u \circ u \circ \dots \circ u}_{n \text{ fois}}$) et aussi $\|Id_E\| = 1$.

Application : équivalence des normes On sait que deux normes équivalentes définissent la même topologie. Nous allons maintenant prouver la réciproque (**réci-proque qui est fautive pour les distances**).

THEOREME 2.2.4

Deux normes N_1 et N_2 sur un \mathbb{K} -ev E sont équivalentes si et seulement si elles définissent la même topologie sur E .

preuve

(E, N_1) et (E, N_2) ont la même topologie si et seulement si $Id_E : (E, N_1) \rightarrow (E, N_2)$ est un homéomorphisme, ce qui revient à dire que cette application et son inverse: $Id_E : (E, N_2) \rightarrow (E, N_1)$ sont continues.

Or $Id_E : (E, N_1) \rightarrow (E, N_2)$ est linéaire; elle est continue ssi il existe une constante positive b telle que, pour tout $x \in E$ $N_2(x) \leq bN_1(x)$. De même, $Id_E : (E, N_2) \rightarrow (E, N_1)$ est continue si et seulement si il existe $c > 0$ telle que pour tout $x \in E$, $N_1(x) \leq cN_2(x)$. D'où le théorème.

2.3 Applications multilinéaires

THEOREME 2.3.1

Soient E_1, \dots, E_p, F ($p + 1$) \mathbb{K} -evn et $\Phi : E_1 \times E_2 \times \dots \times E_p \rightarrow F$ une application p -linéaire. On munit le produit $E_1 \times E_2 \times \dots \times E_p$ d'une quelconque des normes définissant la topologie produit. Les propriétés suivantes sont équivalentes:

1. Φ est continue.
2. Φ est continue en $(0, \dots, 0)$.
3. Il existe $K > 0$ tel que $\forall (x_1, \dots, x_p) \in E_1 \times E_2 \times \dots \times E_p, \|\Phi(x_1, \dots, x_p)\| \leq K \|x_1\|_{E_1} \dots \|x_p\|_{E_p}$.
4. Φ est bornée sur le produit $\prod_{i=1}^{i=p} B'_{E_i}(0, 1)$ des boules unité fermées des E_i .

♠ La preuve est analogue à celle du théorème 2.2.1

2.4 Exemples

1. Soit $E = \mathbb{K}^n$ muni de la norme $\|x\|_1 = \sum_{1 \leq i \leq n} |x_i|$ (où $x = (x_1, \dots, x_n)$) ; soit $B = (e_1, \dots, e_n)$ la base canonique de E . Soit $B^* = (\varepsilon_1, \dots, \varepsilon_n)$ la base de E^* duale de la base B . Soit $\varphi : E \rightarrow \mathbb{K}$ linéaire, et $(\lambda_i)_{1 \leq i \leq n}$ les coordonnées de φ dans la base B^* de sorte que $\varphi(x) = \sum_{1 \leq i \leq n} \lambda_i x_i$ pour tout $x \in E$.

On a $|\varphi(x)| \leq \left(\max_{1 \leq i \leq n} |\lambda_i| \right) \|x\|_1$.

Il en résulte que φ est continue et que $\|\varphi\| \leq \max_{1 \leq i \leq n} |\lambda_i|$.

Si k est tel que $|\lambda_k| = \max_{1 \leq i \leq n} |\lambda_i|$, on a $|\varphi(e_k)| = |\lambda_k| = \left(\max_{1 \leq i \leq n} |\lambda_i| \right) \|e_k\|$ de sorte que $\max_{1 \leq i \leq n} |\lambda_i| \leq \|\varphi\|$, d'où l'égalité.

Par conséquent, on a $E^* = L_c(E, \mathbb{K})$ et pour $\varphi = \sum_{1 \leq i \leq n} \lambda_i \varepsilon_i$ on a $\|\varphi\| = \|(\lambda_1, \dots, \lambda_n)\|_\infty$.

2. Soit $E = C([a, b], \mathbb{R})$ l'espace vectoriel des fonctions continues de l'intervalle $[a, b]$ dans \mathbb{R} muni de la norme de la convergence uniforme : $\|f\|_\infty = \sup_{t \in [a, b]} |f(t)|$. Pour $f \in E$, on note $I(f) = \int_a^b f(t) dt$. Alors I est une application linéaire continue de E dans \mathbb{R} et $\|I\| = b - a$.
3. Soient I un intervalle de \mathbb{R} , et E l'espace vectoriel des fonctions continues sur I et intégrables ; pour $f \in E$, on pose $\|f\|_1 = \int_I |f|$; $\|\cdot\|_1$ est une norme sur E et pour cette norme l'application $f \rightarrow \int_I f$ est linéaire continue de norme 1.
4. Soit $(E, \langle \cdot, \cdot \rangle)$ un espace euclidien et $J : E \rightarrow E^*$ l'isomorphisme canonique: $J(x) = \langle x, \cdot \rangle$. Si on munit E de la norme euclidienne, toute application linéaire de E dans \mathbb{R} est continue et l'isomorphisme J est continu, de norme 1. En fait c'est une isométrie de E sur son dual : pour tout $x \in E$, on a $\|J(x)\|_{E^*} = \|J(x)\| = \|x\|_E$.

3

Compacité

3.1 Définitions

DEFINITION 3.1.1

On dit qu'un espace métrique (X, d) est compact si il possède la propriété suivante (Bolzano Weierstrass) :

(BW) Toute suite de points de X possède une valeur d'adhérence.

Une partie A d'un espace métrique (X, d) est dite compacte si le sous-espace métrique (A, d) est compact.

Donc A est compacte si et seulement si de toute suite de points de A , on peut extraire une sous-suite convergente dans A (c'est à dire une sous-suite convergente dans X dont la limite est dans A).

Exemples

1. Tout intervalle fermé borné de \mathbb{R} est compact.
2. Soient $(x_n)_{n \geq 0}$ une suite convergente de points d'un espace métrique (Y, d) , $\lambda = \lim x_n$ et $X = \{x_n, n \geq 0\} \cup \{\lambda\}$. X muni de la distance induite par celle de Y est compact.

preuve

Soit $(u_k)_{k \geq 0}$ une suite de points de X . Distinguons deux cas.

- L'ensemble $\{u_k ; k \geq 0\}$ est fini. Il existe alors une suite extraite de la suite $(u_k)_{k \geq 0}$ qui est une suite constante.
- L'ensemble $\{u_k ; k \geq 0\}$ est infini ; il existe alors une suite extraite $v_j = u_{n_j}$ de la suite $(u_k)_{k \geq 0}$ dont tous les termes sont distincts. Montrons que cette suite extraite converge vers λ . Soit V un voisinage de λ . Par hypothèse, il existe un entier N tel que pour tout $n \geq N$ on ait $x_n \in V$. Puisque tous les v_j sont distincts, il existe un entier K tel que pour $j \geq K \Rightarrow v_j \notin \{x_0, \dots, x_{N-1}\}$. Alors, pour un tel j , v_j est soit λ , soit un des x_n avec $n \geq N$, donc $j \geq K \Rightarrow v_j \in V$.

On a donc montré que dans tous les cas la suite $(u_k)_{k \geq 0}$ avait une valeur d'adhérence dans Y . ■

Du dernier exemple on déduit, en utilisant 1.5.2 le résultat suivant important dans la pratique.

PROPOSITION 3.1.1

Soient (X, d) et (X', d') deux espaces métriques, $f : X \rightarrow X'$ une application. f est continue si et seulement si sa restriction à tout compact de X est continue.

PROPOSITION 3.1.2

Tout espace métrique homéomorphe à un espace compact est lui même compact.

3.2 Propriétés

THEOREME 3.2.1

Toute partie compacte d'un espace métrique (X, d) est fermée et bornée.

preuve

Soient Y une partie compacte de X et $a \in X$.

- Supposons Y non borné ; alors, pour tout entier n , il existe un $y_n \in Y$ tel que $d(a, y_n) \geq n$. Toute suite extraite $(y_{n_k})_{k \geq 0}$ vérifie alors $\forall k, d(a, y_{n_k}) \geq n_k \geq k$ donc est non bornée et ne peut converger. Contradiction. Donc Y est borné.
- Montrons ensuite que Y est fermé: soit $z \in \overline{Y}$; il existe une suite de points (y_n) de Y qui converge vers z . Y étant compact, cette suite possède une sous-suite $(y_{n_k})_{k \geq 0}$ convergeant dans Y ; or cette suite converge nécessairement vers z , donc $z \in Y$, ce qui prouve que Y est fermé. ■

THEOREME 3.2.2

Toute partie fermée d'un espace métrique compact (X, d) est compacte.

preuve

Soit Y une partie fermée de X compact. Soit (y_n) une suite de points de Y . C'est une suite de points de X compact, donc on peut en extraire une sous-suite $(y_{n_k})_{k \geq 0}$ convergeant vers un $z \in X$. Mais pour tout $k, y_{n_k} \in Y$, donc $z = \lim_{k \rightarrow \infty} y_{n_k} \in \overline{Y}$, et Y étant fermé, $\overline{Y} = Y$. Donc de toute suite de points de Y on peut extraire une sous-suite convergeant vers un point de Y et Y est compact. ■

COROLLAIRE 3.2.1

Une intersection de compacts est compacte.

THEOREME 3.2.3

Soient $(X_i, d_i), 1 \leq i \leq p, p$ espaces métriques compacts. L'espace métrique produit $X = \prod_{i=1}^{i=p} X_i$ est compact. (Rappelons qu'il est muni d'une des trois distance D_1, D_2 ou D_∞).

preuve (pour $p = 2$)

Notons Y et Z les deux espaces compacts et $X = Y \times Z$. Soit $(x_n) = (y_n, z_n)$ une suite de points de X . Puisque Y est compact, il existe une application $\varphi : \mathbb{N} \rightarrow \mathbb{N}$ strictement croissante telle que la suite $y \circ \varphi = (y_{\varphi(n)})_{n \geq 0}$ converge vers un élément $a \in Y$. Puisque Z est compact, de la suite $z \circ \varphi$ on peut extraire une sous-suite qui converge dans Z : il existe $\psi : \mathbb{N} \rightarrow \mathbb{N}$ strictement croissante telle que la suite $z \circ \varphi \circ \psi$ converge vers un élément $b \in Z$. La suite $y \circ \varphi \circ \psi$ extraite de la suite $y \circ \varphi$ converge encore vers a . Par conséquent, la suite $x \circ \varphi \circ \psi$ extraite de la suite x converge vers $(a, b) \in X$. ■

COROLLAIRE 3.2.2

Les compacts de \mathbb{R}^n (resp. de \mathbb{C}^n) (muni de l'une des trois normes usuelles) sont les fermés bornés.

preuve

Compte tenu du théorème 3.2.1 il reste à montrer que les fermés bornés sont compacts. On munit \mathbb{R}^n de la norme $\| \cdot \|_\infty$. Soit F un fermé borné de \mathbb{R}^n . Il existe $a > 0$ tel que $F \subset B'(0, a) = [-a, a]^n$. D'après 1.2.5 le sous espace $[-a, a]^n$ de \mathbb{R}^n est identique à l'espace produit n fois des $[-a, a]$. On sait que $[-a, a]$ est compact, donc d'après le théorème 3.2.3, il en est de même de $[-a, a]^n$. F est fermé contenu dans un compact donc est compact (théorème 3.2.2). Enfin, \mathbb{C}^n muni de la norme $\| \cdot \|_2$ est isométrique à \mathbb{R}^{2n} muni de la norme $\| \cdot \|_2$.

En particulier $\mathbb{D}' = \{z \in \mathbb{C}; |z| \leq 1\}$ est compact.

THEOREME 3.2.4

Soit $(K_n)_{n \geq 0}$ une suite décroissante de compacts non vides d'un espace métrique (X, d) . $\bigcap_{n \geq 0} K_n$ est un compact non vide.

preuve

On sait qu'une intersection de compacts est compacte. Il reste à montrer que l'intersection considérée est non vide. Choisissons pour chaque n un point $x_n \in K_n$. La suite (x_n) est une suite de points du compact K_0 donc elle admet une valeur d'adhérence λ dans K_0 . Soit $k \rightarrow x_{n_k}$ une suite extraite convergeant vers λ . Pour tout entier p , on a pour $k \geq p, n_k \geq n_p \geq p$ donc $x_{n_k} \in K_p$. On en déduit $\lambda \in \overline{K_p} = K_p$ pour tout p , donc $\lambda \in \bigcap_{n \geq 0} K_n$. ■

3.3 Fonctions continues sur un compact

THEOREME 3.3.1

"L'image d'un compact par une application continue est compacte". Plus précisément, si $f : X \rightarrow Y$ est une application continue entre deux espaces métriques (X, d) et (Y, δ) , et si K est un compact de X , alors $f(K)$ est un compact de Y .

preuve

Soit (y_n) une suite de points de $f(K)$. Pour chaque n on peut choisir un antécédent par f , $x_n \in K$ de y_n . K est compact, donc il existe une suite extraite $(x_{n_k})_{k \geq 0}$ de la suite (x_n) qui converge vers un élément $a \in K$. Alors, f étant continue au point a , on a $\lim_{k \rightarrow \infty} f(x_{n_k}) = f(a) \in f(K)$. Donc la suite (y_{n_k}) extraite de la suite (y_n) converge vers un élément de $f(K)$. ■

THEOREME 3.3.2

Une bijection continue d'un espace métrique compact (X, d) sur un espace métrique (Y, δ) est un homéomorphisme de X sur Y (et Y est compact).

preuve

Soit $f : X \rightarrow Y$ une bijection continue de X sur Y . Il s'agit de montrer que $g = f^{-1}$ est continue. On va montrer que l'image réciproque par g de tout fermé F de Y est un fermé de X . On a $g^{-1}(F) = f(F)$; or F est fermé dans Y compact, donc est compact (théorème 3.2.2). Donc $f(F)$ est compact (théorème 3.3.1) donc fermé (théorème 3.2.1).

THEOREME 3.3.3

"Toute fonction numérique continue sur un compact (non vide) est bornée et atteint ses bornes".

Plus précisément, soient (X, d) un espace métrique et $f : X \rightarrow \mathbb{R}$ une application continue. Soit K un compact non vide de X . Alors f est bornée sur K et atteint ses bornes, i.e. :

$$\exists a, b \in K, f(a) = \inf_{x \in K} f(x), f(b) = \sup_{x \in K} f(x)$$

preuve

$f(K)$ est un compact non vide de \mathbb{R} donc un fermé borné non vide; $f(K)$ est borné non vide donc possède une borne supérieure M et une borne inférieure m ; $f(K)$ étant fermé, ces bornes appartiennent à $f(K)$. ■

THEOREME 3.3.4

"Toute fonction continue sur un espace compact est uniformément continue."

Si (X, d) est un espace métrique compact et si f est une application continue de X dans un autre espace métrique (Y, δ) f est uniformément continue sur X .

preuve

On suppose le contraire : il existe $\varepsilon_0 > 0$ tel que $\forall \eta > 0, \exists x, y \in X, d(x, y) < \eta$ et $\delta(f(x), f(y)) \geq \varepsilon_0$. Appliquons ceci avec $\eta = 1/n$, pour $n \in \mathbb{N}^*$. Il existe donc $x_n, y_n \in X$ tels que $d(x_n, y_n) \leq 1/n$ et $\delta(f(x_n), f(y_n)) \geq \varepsilon_0$. De la suite (x_n) on peut extraire une sous-suite $(x_{n_k})_{k \geq 0}$ convergeant vers un élément $a \in X$. Comme $\forall k, d(x_{n_k}, y_{n_k}) \leq \frac{1}{n_k} \leq \frac{1}{k}$, on a aussi $\lim_{k \rightarrow \infty} y_{n_k} = a$. Comme f est continue en a , les suites $(f(x_{n_k}))$ et $(f(y_{n_k}))$ convergent toutes les deux vers $f(a)$, donc $\delta(f(x_{n_k}), f(y_{n_k}))$ tend vers 0, ce qui est contradictoire avec $\delta(f(x_{n_k}), f(y_{n_k})) \geq \varepsilon_0$ pour tout k . ■

3.4 Application aux espaces vectoriels normés de dimension finie

THEOREME 3.4.1

Soit E un \mathbb{K} -ev de dimension finie. ($\mathbb{K} = \mathbb{R}$ ou \mathbb{C}). Alors

1. Toutes les normes sur E sont équivalentes. E possède donc une unique topologie d'evn.
2. Pour cette topologie, les parties compactes de E sont les parties fermées bornées. De toute suite bornée de E on peut extraire une sous-suite convergente
3. Si $(F, \| \cdot \|)$ est un autre evn, toute application linéaire de E dans F est continue.

preuve

1. Préliminaire

On rappelle qu'il a été prouvé (corollaire 3.2.2) que dans \mathbb{K}^n muni de la norme $\|\cdot\|_\infty$, les boules fermées sont compactes, et que par conséquent, les compacts sont les parties fermées bornées.

2. Première étape

On va montrer que toute norme p sur \mathbb{K}^n est équivalente à la norme $\|\cdot\|_\infty$

- L'application $p : \mathbb{K}^n \rightarrow \mathbb{R}$ est continue lorsqu'on munit \mathbb{K}^n de la norme $\|\cdot\|_\infty$.
En effet, soient $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n)$ et $(e_i)_{1 \leq i \leq n}$ est la base canonique de \mathbb{K}^n ;
on a $0 \leq p(x - y) \leq \sum_{1 \leq i \leq n} |x_i - y_i| p(e_i) \leq C \|x - y\|_\infty$ où $C = \sum_{1 \leq i \leq n} p(e_i)$ est une constante.
- L'ensemble $S = \{x \in \mathbb{K}^n ; \|x\|_\infty = 1\}$ est un fermé borné de $(\mathbb{K}^n, \|\cdot\|_\infty)$ donc est un compact de $(\mathbb{K}^n, \|\cdot\|_\infty)$.
- L'application $p : \mathbb{K}^n \rightarrow \mathbb{R}$ continue sur le compact S est bornée et atteint ses bornes. Il existe donc des réels a et b et un élément $x_1 \in S$ tels que $a = \inf_{x \in S} p(x) = p(x_1) > 0$ et $b = \sup_{x \in S} p(x)$. On a donc $0 < a \leq b$. Pour tout x non nul de \mathbb{K}^n , on a $a \leq p(x / \|x\|_\infty) \leq b$ soit $a \|x\|_\infty \leq p(x) \leq b \|x\|_\infty$; cette inégalité reste valable si $x = 0$, d'où le résultat.

3. Deuxième étape

Deux normes quelconques sur \mathbb{K}^n sont équivalentes.

Cela résulte de la première étape et du fait que l'équivalence des normes est une relation transitive.

4. Troisième étape

Deux normes quelconques sur E sont équivalentes. Soit $(\varepsilon_i)_{1 \leq i \leq n}$ une base de E . Soient N et N' deux normes sur E .

Soit $x = (x_1, \dots, x_n) \in \mathbb{K}^n$. On pose $p(x) = N\left(\sum_{1 \leq i \leq n} x_i \varepsilon_i\right)$ et $p'(x) = N'\left(\sum_{1 \leq i \leq n} x_i \varepsilon_i\right)$. Comme $x \rightarrow \sum_{1 \leq i \leq n} x_i \varepsilon_i$

est un isomorphisme d'espaces vectoriels de E sur \mathbb{K}^n , on vérifie facilement que p et p' sont deux normes sur \mathbb{K}^n . Il existe donc $0 < a \leq b$ tels que $ap \leq p' \leq bp$. On en déduit $\forall v \in E$, $aN(v) \leq N'(v) \leq bN(v)$ ce qui achève la preuve du 1) du théorème.

5. Démonstration des points 2 et 3

Les points (2) et (3) sont vrais dans \mathbb{K}^n muni de $\|\cdot\|_\infty$. Si p est une norme sur \mathbb{K}^n elle est équivalente à $\|\cdot\|_\infty$, donc (2) et (3) sont vrais dans (\mathbb{K}^n, p) .

Soient N une norme sur E , $(\varepsilon_i)_{1 \leq i \leq n}$ une base de E et p la norme sur \mathbb{K}^n définie par $p(x) = N\left(\sum_{1 \leq i \leq n} x_i \varepsilon_i\right)$ (cette norme dépend du choix de la base ε de E). L'application $h : x \rightarrow \sum_{1 \leq i \leq n} x_i \varepsilon_i$ de \mathbb{K}^n sur E est une isométrie par construction ; elle est donc continue ainsi que h^{-1} . C'est donc un homéomorphisme de \mathbb{K}^n sur E . Une partie K de E est compacte (respectivement fermée, bornée) dans E si et seulement si $h^{-1}(K)$ est compacte (respectivement fermée, bornée) dans \mathbb{K}^n muni de p . D'où la caractérisation des compacts de E . D'autre part, une application $\varphi : E \rightarrow F$ est continue si et seulement si $\varphi \circ h$ l'est. Or si φ est linéaire, $\varphi \circ h : \mathbb{K}^n \rightarrow F$ est aussi linéaire, donc continue.

La preuve du théorème suivant est laissée au lecteur.

THEOREME 3.4.2

Soient E_1, \dots, E_p \mathbb{K} -espaces vectoriels de dimension finie, et F un espace vectoriel normé quelconque. Toute application p -linéaire $\Phi : E_1 \times \dots \times E_p \rightarrow F$ est continue (pour l'unique topologie d'evn du produit).

3.5 Application: théorème de d'Alembert-Gauss

THEOREME 3.5.1

Tout polynôme à coefficients complexes de degré au moins égal à 1 possède une racine dans \mathbb{C} .

preuve

Soit $P(z) = a_n z^n + \dots + a_1 z + a_0$ un polynôme à coefficients dans \mathbb{C} , de degré $n \geq 1$. ($a_n \neq 0$). La démonstration se fait en deux étapes: dans un premier temps, on va prouver l'existence d'un $z_0 \in \mathbb{C}$ tel que la fonction $z \rightarrow |P(z)|$ ait un minimum

en z_0 . Ensuite, on prouvera que si z_0 est un minimum pour la fonction $z \rightarrow |P(z)|$, nécessairement $P(z_0) = 0$. Dans cette deuxième partie, nous utiliserons le fait que tout nombre complexe admet, pour tout entier $p \geq 2$ une racine p -ième. Ce résultat se prouve à l'aide de la forme trigonométrique des nombres complexes, autrement dit au moyen de l'exponentielle complexe.

1. Première étape: existence d'un minimum de la fonction $|P|$.

- Ecrivons pour $z \neq 0$

$$P(z) = a_n z^n \left(1 + \sum_{k=0}^{n-1} \frac{a_k}{a_n} \frac{1}{z^{n-k}} \right)$$

Le terme entre parenthèses tend vers 1 quand $|z|$ tend vers l'infini et par conséquent $|P(z)|$ tend vers l'infini avec $|z|$. Il existe donc un $R > 0$ tel que

$$|z| > R \Rightarrow |P(z)| > |P(0)|$$

- Soit K le disque fermé de centre 0 et de rayon R . K est compact, et $z \rightarrow |P(z)|$ est continue donc cette fonction atteint son minimum sur K en un point $z_0 \in K$. Pour tout $z \in \mathbb{C}$, on a $|P(z)| \geq |P(z_0)|$. En effet, ou bien $z \in K$ et cela résulte de la définition de z_0 , ou bien $z \notin K$ et alors $|z| > R$, donc $|P(z)| > |P(0)| \geq |P(z_0)|$ car $0 \in K$.

2. Deuxième étape: $P(z_0) = 0$.

Nous allons supposer $P(z_0) \neq 0$ et montrer que cela conduit à une contradiction.

- On se ramène au cas $z_0 = 0$ et $P(z_0) = 1$.

Considérons $Q(w) = \frac{1}{P(z_0)} P(z_0 + w)$. Q est un polynôme de même degré que P et on voit que

$$|P(z_0)| = \inf_{z \in \mathbb{C}} |P(z)| \Rightarrow 1 = Q(0) = \inf_{w \in \mathbb{C}} |Q(w)|$$

- On peut écrire

$$Q(z) = 1 + b_p z^p + \sum_{k=p+1}^{k=n} b_k z^k = 1 + b_p z^p \left(1 + \sum_{k=p+1}^{k=n} \frac{b_k}{b_p} z^{k-p} \right) \text{ avec } b_p \neq 0 \text{ et } p \geq 1$$

Soit ζ une racine p -ième de $-b_p$ de sorte que Q s'écrit $Q(z) = 1 - (\zeta z)^p h(z)$ avec $\lim_{z \rightarrow 0} h(z) = 1$.

- On termine avec une technique de développement limité : considérons, pour $t > 0$:

$$f(t) = Q\left(\frac{t}{\zeta}\right)$$

Il vient

$$f(t) = 1 - t^p h_1(t) \text{ avec } \lim_{t \rightarrow 0} h_1(t) = 1$$

On a :

$$|f(t)|^2 = 1 - 2t^p \operatorname{Re}(h_1(t)) + t^{2p} |h_1(t)|^2$$

Puisque $p \geq 1$, le troisième terme est négligeable devant le second lorsque t tend vers 0, de sorte que $|f(t)|^2 - 1$ est équivalent, quand t tend vers 0 à $-2t^p$ (car $\operatorname{Re}(h_1(t))$ tend vers 1 quand t tend vers 0). Il en résulte que pour t suffisamment petit $|f(t)|^2 - 1 < 0$, ce qui contredit le fait que le minimum du module de Q est 1.

L'hypothèse $P(z_0) \neq 0$ conduit à une contradiction, donc $P(z_0) = 0$ et le théorème est prouvé. ■

3.6 Complément : caractérisation de Borel Lebesgue des compacts

NB: Ce paragraphe est hors du programme officiel du concours.

Soit (X, d) un espace métrique. Un recouvrement ouvert de X est une famille d'ouverts $(U_i)_{i \in I}$, (indexée par un ensemble I quelconque) telle que $X = \bigcup_{i \in I} U_i$

DEFINITION 3.6.1

On dit qu'un espace métrique (X, d) possède la propriété de Borel-Lebesgue (B.L.) si de tout recouvrement ouvert $(U_i)_{i \in I}$, de X on peut extraire un recouvrement fini, i.e. il existe $J \subset I$, J fini, tel que $X = \bigcup_{i \in J} U_i$.

En passant aux complémentaires, on voit que la propriété (BL) est équivalente à la suivante (BL') :

Soit $(F_i)_{i \in I}$ une famille de fermés de X . Si pour tout sous-ensemble fini J de I l'intersection $\bigcap_{i \in J} F_i$ est non vide, alors $\bigcap_{i \in I} F_i$ est non vide.

THEOREME 3.6.1

Soit (X, d) un espace métrique. X est compact ssi il possède la propriété de Borel-Lebesgue.

preuve

1. Si X possède la propriété (BL), il est compact.

Soit (a_n) une suite d'éléments de X . Soit $A_n = \{a_k ; k \geq n\}$. Si J est une partie finie de \mathbb{N} , de plus grand élément m , on a $\bigcap_{i \in J} A_i = A_m \neq \emptyset$ car pour tout k on a $A_{k+1} \subset A_k$ donc $A_{k+1} \subset A_k$. D'après (BL)' l'intersection $\bigcap_{n \geq 0} A_n$ est non vide, donc la suite (a_n) possède une valeur d'adhérence. X est compact.

2. Si X est compact, il possède la propriété (BL). Soit $\mathcal{U} = (U_i)_{i \in I}$ un recouvrement ouvert de X .

(a) Nous allons d'abord montrer qu'il existe un nombre $\varepsilon > 0$ (nombre de Lebesgue du recouvrement \mathcal{U}) tel que pour tout x de X , il existe un indice $i \in I$ tel que $B(x, \varepsilon) \subset U_i$.

Supposons le contraire. Alors, pour tout $\varepsilon > 0$ il existe $x \in X$ tel que $\forall i \ B(x, \varepsilon) \not\subset U_i$. Appliquons ceci avec $\varepsilon = 1/n, n \in \mathbb{N}^*$. On obtient un point x_n tel que pour tout $i, B(x_n, \frac{1}{n}) \not\subset U_i$. De la suite (x_n) on peut extraire une suite (x_{n_k}) convergente. Soit λ sa limite.

Puisque \mathcal{U} est un recouvrement de X , il existe $j \in I$ tel que $\lambda \in U_j$. Puisque U_j est ouvert, il existe $r > 0$ tel que $B(\lambda, r) \subset U_j$. La suite $(x_{n_k})_{k \in \mathbb{N}}$ converge vers λ . Choisissons un entier K tel que $d(x_{n_K}, \lambda) < \frac{r}{2}$ et $n_K > \frac{2}{r}$. On a alors

$$B\left(x_{n_K}, \frac{1}{n_K}\right) \subset B\left(\lambda, \frac{1}{n_K} + d(\lambda, x_{n_K})\right) \subset B(\lambda, r) \subset U_j$$

ce qui est contraire à l'hypothèse. Le résultat annoncé est donc prouvé par l'absurde.

(b) Fin de la démonstration

Supposons qu'il n'existe pas de recouvrement ouvert fini de X extrait de \mathcal{U} . Soit $x_0 \in X$. Il existe un indice $i_0 \in I$ tel que $B(x_0, \varepsilon) \subset U_{i_0}$. Par hypothèse, $U_{i_0} \neq X$, donc il existe $x_1 \in X \setminus U_{i_0}$. Il existe un indice i_1 tel que $B(x_1, \varepsilon) \subset U_{i_1}$; par hypothèse $U_{i_0} \cup U_{i_1} \neq X$. De proche en proche, on construit une suite (x_n) de points de X et une suite (i_n) d'éléments de I tels que $B(x_k, \varepsilon) \subset U_{i_k}$ et $x_k \notin \bigcup_{0 \leq j < k} U_{i_j}$ pour tout k entier. Il en résulte

que pour tout couple (p, q) d'entiers distincts on a $d(x_p, x_q) \geq \varepsilon$. Donc, si $(x_{n_k})_{k \geq 0}$ est une suite extraite de la suite (x_n) on a encore $k \neq m \Rightarrow d(x_{n_k}, x_{n_m}) \geq \varepsilon$. La suite $(x_{n_k})_{k \geq 0}$ n'est pas de Cauchy, donc ne converge pas. On a donc construit une suite (x_n) n'admettant aucune valeur d'adhérence, ce qui entre en contradiction avec la compacité de X . L'hypothèse "il n'existe pas de recouvrement fini extrait de \mathcal{U} " conduit à une contradiction, donc cette hypothèse est fautive et X possède la propriété (BL).

4

Connexité

4.1 Définition et propriétés élémentaires

THEOREME 4.1.1

Soit A une partie d'un espace métrique X . Les propriétés suivantes sont équivalentes :

1. Les seules parties à la fois ouvertes et fermées de A sont \emptyset et A .
2. Il n'existe pas de partition de A en deux ouverts de A .
3. Il n'existe pas de partition de A en deux fermés de A .
4. Toute fonction continue sur A à valeurs dans $\{0, 1\}$ est constante.

DEFINITION 4.1.1

Une partie A qui vérifie l'une de ces propriétés (et donc les quatre) est dite connexe.

Attention! Dans cette définition, les ouverts et fermés sont des ouverts ou fermés de A pour la topologie induite. La propriété (1) se traduit donc de la manière suivante : soit Y une partie de A telle qu'il existe un ouvert O de X et un fermé F de X tels que $Y = O \cap A = F \cap A$; alors $Y = A$ ou $Y = \emptyset$

De même (2) peut s'écrire ainsi: soient O_1 et O_2 deux ouverts de X tels que $A \subset O_1 \cup O_2$ et $A \cap O_1 \cap O_2 = \emptyset$. Alors on a $A \cap O_1 = \emptyset$ ou $A \cap O_2 = \emptyset$.

preuve

Les équivalences entre (1), (2) et (3) sont triviales. Notons $E = \{0, 1\}$ muni de la distance induite par celle de \mathbb{R} . $\{0\}$ et $\{1\}$ sont fermés dans E . Supposons (3) vérifiée; soit $\varphi : A \rightarrow E$ continue ; chacun des deux ensembles $\varphi^{-1}(0)$ et $\varphi^{-1}(1)$ est fermé dans A et $A = \varphi^{-1}(0) \cup \varphi^{-1}(1)$; par (3) l'un des deux ensembles $\varphi^{-1}(0)$, $\varphi^{-1}(1)$ est vide et φ est constante. Réciproquement,

supposons (4) ; soit $\{F_0, F_1\}$ une partition de A en deux fermés. Soit φ la fonction indicatrice de F_1 : $\varphi(x) = \begin{cases} 0 & \text{si } x \in F_0 \\ 1 & \text{si } x \in F_1 \end{cases}$.

Les fermés de E sont \emptyset , $\{0\}$, $\{1\}$ et E ; leurs images réciproques par φ sont \emptyset , F_0 , F_1 , et A tous fermés dans A donc φ est continue. Contradiction. ■

THEOREME 4.1.2

Soient A une partie connexe de X et B une partie de X telle que $A \subset B \subset \bar{A}$. B est connexe.

preuve

Soit $\varphi : B \rightarrow \{0, 1\}$ continue ; A étant connexe, la restriction φ_A de φ à A est constante, par exemple égale à 0. Tout point b de B est limite d'une suite de points (a_n) de A et φ étant continue on a $\varphi(b) = \lim_{n \rightarrow \infty} \varphi(a_n) = 0$, donc φ est constante sur B .

THEOREME 4.1.3

Soit $(A_i)_{i \in I}$ une famille de parties connexes de X d'intersection non vide. La réunion $\bigcup_{i \in I} A_i$ est connexe.

preuve

Soient $A = \bigcup_{i \in I} A_i$, $a \in \bigcap_{i \in I} A_i$ et $\varphi : A \rightarrow \{0, 1\}$ continue ; soit $x \in A$; il existe un $i \in I$ tel que $x \in A_i$; A_i étant connexe, la restriction de φ à A_i est constante, donc $\varphi(x) = \varphi(a)$; φ est constante donc A est connexe. ■

THEOREME 4.1.4 (Image continue d'un connexe)

Soient A une partie connexe de X et $h : A \rightarrow Y$ une application continue de A dans un autre espace métrique Y . Alors $h(A)$ est une partie connexe de Y .

preuve

Soit $\varphi : h(A) \rightarrow \{0, 1\}$ continue. La fonction $\varphi \circ h$ est continue sur A connexe, à valeurs dans $\{0, 1\}$, donc est constante; si c est sa valeur on a pour tout $y = h(x) \in h(A)$, $\varphi(y) = \varphi \circ h(x) = c$, donc φ est constante. ■

4.2 Connexes de \mathbb{R}

THEOREME 4.2.1

Les parties connexes de \mathbb{R} sont les intervalles.

preuve

Soit A une partie de \mathbb{R} . On va montrer : A connexe $\Rightarrow A$ convexe $\Rightarrow A$ est un intervalle $\Rightarrow A$ connexe.

- A connexe $\Rightarrow A$ convexe
Supposons A non convexe ; alors il existe a et b dans A et $x_0 \in]a, b[$ tels que $x_0 \notin A$. Les ensembles $] - \infty, x_0[\cap A$ et $]x_0, \infty[\cap A$ constituent une partition de A en deux ouverts. Contradiction.
- A convexe $\Rightarrow A$ est un intervalle
On peut supposer A non vide ; soit $x_0 \in A$ et $\beta = \sup(A)$ (la borne supérieure est prise dans $\overline{\mathbb{R}}$: on peut avoir $\beta = +\infty$). Si $y \in]x_0, \beta[$, il existe un $x_1 \in A$ tel que $y < x_1 < \beta$ par définition de la borne supérieure. A étant convexe, $[x_0, x_1] \subset A$, donc $y \in A$. Il en résulte que $]x_0, \beta[\subset A$; donc $A \cap]x_0, +\infty[$ est soit égal à $]x_0, \beta[$, soit égal à $]x_0, \beta]$ (avec dans ce cas β fini). Donc $A \cap]x_0, +\infty[$ est un intervalle. De même pour $A \cap]-\infty, x_0]$. Donc A réunion de deux intervalles ayant un point commun est un intervalle.
- A est un intervalle $\Rightarrow A$ connexe
Supposons A non connexe. Il existe deux fermés de \mathbb{R} , F_1 et F_2 tels que $A \subset F_1 \cup F_2$, $A \cap F_1 \cap F_2 = \emptyset$, $F_1 \cap A \neq \emptyset$ et $F_2 \cap A \neq \emptyset$. Soient $a_i \in F_i \cap A$, $i = 1, 2$ avec, par exemple $a_1 < a_2$. On a $[a_1, a_2] \subset A$ car A est un intervalle. Soient $K_i = [a_1, a_2] \cap F_i$, $i = 1, 2$. On a $K_1 \cup K_2 = [a_1, a_2]$. K_1 est un fermé borné non vide de \mathbb{R} donc $M = \sup(K_1) \in K_1$. On a $K_1 \cap K_2 = \emptyset$ et $a_2 \in K_2$ donc $M < a_2$ et $]M, a_2] \subset K_2$ ce qui implique $M \in \overline{K_2} = K_2$ (car K_2 est un fermé de \mathbb{R}). Contradiction avec $K_1 \cap K_2 = \emptyset$. Donc A est connexe.

4.3 Connexité par arcs

DEFINITION 4.3.1

Soit A une partie d'un espace métrique (X, d) . On dit que A est connexe par arcs si pour tout couple de points (x, y) de A il existe une fonction continue $\gamma : [0, 1] \rightarrow A$ telle que $\gamma(0) = x$ et $\gamma(1) = y$.

THEOREME 4.3.1

Toute partie connexe par arcs d'un espace métrique (X, d) est connexe.

preuve

Soit $A \subset X$ connexe par arcs. On peut supposer $A \neq \emptyset$; soient $a \in A$ et $\varphi : A \rightarrow \{0, 1\}$ continue. Soit $x \in A$ et $\gamma : [0, 1] \rightarrow A$ continue telle que $\gamma(0) = a$ et $\gamma(1) = x$. La fonction $\varphi \circ \gamma : [0, 1] \rightarrow \{0, 1\}$ est continue, donc constante puisque $[0, 1]$ est connexe : $\varphi \circ \gamma(1) = \varphi \circ \gamma(0)$ soit $\varphi(x) = \varphi(a)$; φ est constante donc A est connexe.

COROLLAIRE 4.3.1

Toute partie convexe d'un espace vectoriel normé est connexe. Toute partie convexe d'un espace affine réel de dimension finie est connexe.

THEOREME 4.3.2

Soit E un \mathbb{K} -evn et U un ouvert de E . Alors U est connexe si et seulement si il est connexe par arcs.

preuve

Lemme

Soit A une partie d'un espace métrique X et ρ la relation binaire sur A définie par $(x\rho y) \Leftrightarrow \exists \gamma : [0, 1] \rightarrow A$, γ continue, $\gamma(0) = x$, $\gamma(1) = y$. ρ est une relation d'équivalence sur A .

- Preuve du lemme :

Nous noterons $\gamma_{x,y}$ une fonction $\gamma : [0, 1] \rightarrow A$ continue telle que $\gamma(0) = x$ et $\gamma(1) = y$.

La relation ρ est réflexive (pour x dans A , prendre $\gamma_{x,x}(t) = x$).

Elle est symétrique : si $x\rho y$ on a $y\rho x$ (prendre $\gamma_{y,x}(t) = \gamma_{x,y}(1-t)$).

Enfin, elle est transitive: supposons $x\rho y$ et $y\rho z$; soit h définie par $h(t) = \begin{cases} \gamma_{x,y}(2t) & \text{si } 0 \leq t \leq 1/2 \\ \gamma_{y,z}(2t-1) & \text{si } 1/2 < t \leq 1 \end{cases}$

h est une application continue de I dans A et vérifie $h(0) = x$ et $h(1) = z$.

- Preuve du théorème : Il s'agit de montrer que U ouvert connexe implique U connexe par arcs. On peut supposer U non vide. Soit ρ la relation d'équivalence sur U définie dans le lemme ci dessus et C une classe d'équivalence. C est un ouvert de E ; en effet, soit $x \in C$ et soit $r > 0$ tel que $B(x, r) \subset U$. Toute boule dans un evn est convexe donc connexe par arcs. Tous les points de $B(x, r)$ sont donc liés à x par ρ . Donc $B(x, r) \subset C$. Soit $a \in U$ et C la classe d'équivalence de a . Le complémentaire de C dans U est une réunion de classes d'équivalence, donc une réunion d'ouverts de E ; c'est donc un ouvert de E ; il en résulte que C est fermé dans E donc dans U . Or U est connexe, et C est non vide, ouvert et fermé dans U donc $C = U$; tout point de U est lié à a par ρ donc U est connexe par arcs. ■

4.4 Applications

4.4.1 Théorème des valeurs intermédiaires

THEOREME 4.4.1

Soit (X, d) un espace métrique connexe et $f : X \rightarrow \mathbb{R}$ continue. $f(X)$ est un intervalle.

C'est une conséquence immédiate des théorèmes 4.1.4 et 4.2.1. En particulier, si $f : [a, b] \rightarrow \mathbb{R}$ est continue et vérifie $f(a)f(b) \leq 0$ alors il existe $c \in [a, b]$ tel que $f(c) = 0$.

4.4.2 Homéomorphismes d'intervalles

THEOREME 4.4.2

Soit I un intervalle de \mathbb{R} et $f : I \rightarrow \mathbb{R}$ continue et strictement croissante. Alors $J = f(I)$ est un intervalle.

1. Si $I = [a, b]$, $J = [f(a), f(b)]$.
2. Si $I = [a, b[$ avec $-\infty < a < b \leq +\infty$, $J = [f(a), \lim_{x \rightarrow b} f(x)[$.
3. Si $I =]a, b]$ avec $-\infty \leq a < b < +\infty$, $J =]\lim_{x \rightarrow a} f(x), f(b)]$.
4. Si $I =]a, b[$ avec $-\infty \leq a < b \leq +\infty$, $J =]\lim_{x \rightarrow a} f(x), \lim_{x \rightarrow b} f(x)[$.

L'application f est un homéomorphisme de I sur J et l'homéomorphisme réciproque f^{-1} est strictement croissant..

Enoncé analogue, *mutatis mutandis*, dans le cas où f est strictement décroissante.

preuve

D'après le théorème des valeurs intermédiaires, J est un intervalle. Le cas $I = [a, b]$ est facile. Considérons par exemple le cas $I = [a, b[$. Puisque f est croissante, elle a une limite (finie ou infinie) quand x tend vers b , soit β . Soit $x \in I$. Choisissons $y \in I$, $y > x$. On a $f(a) \leq f(x) < f(y) \leq \beta$, donc $f(x) \in [f(a), \beta[$ d'où $J \subset [f(a), \beta[$. Prouvons l'inclusion réciproque : soit $z \in [f(a), \beta[$. Puisque $\beta = \lim_{x \rightarrow b} f(x)$, il existe $y \in I$ tel que $z < f(y)$. Par le théorème des valeurs intermédiaires, $[f(a), f(y)] \subset f([a, y]) \subset J$, donc $z \in J$.

Montrons maintenant la continuité de f^{-1} sur J . Soit $[u, v] \subset J$ et $a = f^{-1}(u)$, $b = f^{-1}(v)$; la restriction de f à $[a, b]$ est une bijection continue du compact $[a, b]$ sur $[u, v]$; c'est donc un homéomorphisme (théorème 3.3.2). La bijection réciproque qui n'est autre que la restriction de f^{-1} à $[u, v]$ est donc continue sur $[u, v]$. Or tout compact de J est inclus dans un intervalle de la forme $[u, v]$; donc la restriction de f^{-1} à tout compact de J est continue. Il en résulte que f^{-1} est continue sur J (proposition 3.1.1).

Le taux de variation de f^{-1} entre les points u, v de J est l'inverse de celui de f entre les points $f^{-1}(u), f^{-1}(v)$ de I donc il est strictement positif, ce qui prouve que f^{-1} est strictement croissante. ■

Exercice Démontrer la continuité de f^{-1} sur J de manière élémentaire, c'est à dire sans faire intervenir les théorèmes utilisant la compacité.

THEOREME 4.4.3

Soient I un intervalle de \mathbb{R} et $f : I \rightarrow \mathbb{R}$ continue et injective. f est strictement monotone.

preuve

Soient $\Delta = \{(x, y) \in I \times I \mid y > x\}$ et $T : \Delta \rightarrow \mathbb{R}$ définie par $T(x, y) = \frac{f(y) - f(x)}{y - x}$. Δ est convexe donc connexe. L'application T est continue sur Δ (car f l'est) et ne s'annule pas car f est injective. Donc (thm des valeurs intermédiaires) T garde un signe constant sur Δ ce qui prouve que f est strictement monotone. ■

4.4.3 Structure des ouverts de \mathbb{R}

THEOREME 4.4.4

Soit Ω un ouvert de \mathbb{R} . Ω est la réunion d'une famille finie ou dénombrable d'intervalles ouverts deux à deux disjoints.

preuve

On suppose Ω non vide. On définit une relation binaire ρ sur Ω par $\forall x, y \in \Omega$, $x \rho y \Leftrightarrow ([x, y] \subset \Omega \text{ ou } [y, x] \subset \Omega)$; c'est une relation d'équivalence. Soit C une classe d'équivalence. Par définition, si $x, y \in C$, avec $x < y$, on a $[x, y] \subset C$, autrement dit C est convexe, donc c'est un intervalle. Soit $x_0 \in C$; Ω étant ouvert, il existe $a > 0$ tel que $J =]x_0 - a, x_0 + a[\subset \Omega$. Si $x \in J$, on a $x_0 \rho x$ donc $x \in C$; on a donc $J \subset C$ ce qui prouve que C est ouvert. C'est donc un intervalle ouvert.

Soit $E = \Omega / \rho$ l'ensemble des classes d'équivalence de ρ ; les éléments de E forment une partition de Ω en intervalles ouverts. Chacun de ces intervalles I contient un rationnel r_I . Puisque les intervalles sont deux à deux disjoints, ces rationnels sont deux à deux distincts. On dispose ainsi d'une application injective $E \rightarrow \mathbb{Q}$. Comme \mathbb{Q} est dénombrable, E est fini ou dénombrable.

4.5 Complément: composantes connexes

NB : Cette notion ne figure plus au programme.

Soit (X, d) un espace métrique, A une partie quelconque non vide de X . Considérons la relation binaire ρ sur A définie par $x \rho y$ si et seulement si il existe une partie connexe C de A qui contient à la fois x et y ; ρ est une relation d'équivalence. (la transitivité découle du théorème 4.1.3).

DEFINITION 4.5.1

Les composantes connexes de A sont les classes d'équivalence de ρ .

PROPOSITION 4.5.1

Soit A une partie non vide d'un espace métrique X . Si B est une partie non vide et connexe de A , B est contenu dans une composante connexe de A et une seule.

Les composantes connexes de A sont connexes et fermées dans A .

preuve

Notons $\Gamma(x)$ la composante connexe de A contenant l'élément x de A .

- Soit $x_0 \in B$. Pour tout $y \in B$, on a $x_0 \rho y$ par définition de ρ , donc $B \subset \Gamma(x_0)$. L'unicité est triviale.
- Soit $x \in A$ et $\varphi : \Gamma(x) \rightarrow \{0, 1\}$ continue. Si $y \in \Gamma(x)$, il existe une partie connexe C_y de A qui contient x et y ; φ restreinte à C_y est constante et vaut $\varphi(x)$; on a donc $\forall y \in \Gamma(x)$, $\varphi(y) = \varphi(x)$ donc φ est constante. Ceci prouve que $\Gamma(x)$ est connexe.
- On a $\Gamma(x) \subset \overline{\Gamma(x)} \cap A \subset \overline{\Gamma(x)}$. Puisque $\Gamma(x)$ est connexe, il en est de même $\overline{\Gamma(x)} \cap A$ (thm 4.1.3). Comme cet ensemble contient x et est contenu dans A , il est contenu dans $\Gamma(x)$ ce qui prouve que $\overline{\Gamma(x)} \cap A = \Gamma(x)$, donc que $\Gamma(x)$ est un fermé relatif de A . ■

EXEMPLE 4.5.1

- Les composantes connexes de \mathbb{Q} sont les singletons.
- Si A est un ouvert Ω de \mathbb{R} la relation ρ est la relation définie au § 4.4.3. Les composantes connexes de Ω sont donc les intervalles ouverts cités dans le théorème 4.4.4.

5

Espaces Complets

5.1 Définitions et propriétés élémentaires

DEFINITION 5.1.1

On dit qu'une suite de points $(x_n)_{n \geq 0}$ d'un espace métrique (X, d) est une suite de Cauchy si elle vérifie:

$$\forall \varepsilon > 0, \exists N = N(\varepsilon), \forall p, q \in \mathbb{N} \quad p \geq N \Rightarrow d(x_{p+q}, x_p) \leq \varepsilon$$

Propriétés élémentaires

Soit (X, d) un espace métrique.

1. ♠ La suite (x_n) de points de X est de Cauchy ssi il existe une suite réelle (ε_n) de limite nulle telle que pour tout n et $p \in \mathbb{N}$, $d(x_{n+p}, x_n) \leq \varepsilon_n$.
2. ♠ Toute suite convergente de points de X est de Cauchy.
3. ♠ Toute suite de Cauchy dans X est bornée.
4. Toute suite de Cauchy dans X qui possède une valeur d'adhérence λ n'en possède qu'une et converge vers cette valeur d'adhérence.
En effet, soit $\varepsilon > 0$. Puisque λ est valeur d'adhérence, il existe un entier $n \geq N(\varepsilon/2)$ tel que $d(x_n, \lambda) \leq \varepsilon/2$. Si p est supérieur ou égal à n , on a $d(x_p, \lambda) \leq d(x_p, x_n) + d(x_n, \lambda) \leq \varepsilon$.
5. ♠ Soit $(x_n)_{n \geq 0}$ une suite de points de X et $A_n = \{x_k \mid k \geq n\}$. La suite $(x_n)_{n \geq 0}$ est de Cauchy si et seulement si $\text{diam}(A_n)$ tend vers 0 quand n tend vers l'infini.
6. ♠ Soient d_1 et d_2 deux distances fortement équivalentes ($ad_1 \leq d_2 \leq bd_1$, $a > 0$) sur un même ensemble X et $(x_n)_{n \geq 0}$ une suite de points de X . $(x_n)_{n \geq 0}$ est de Cauchy dans (X, d_1) ssi elle est de Cauchy dans (X, d_2) .
7. ♠ Soient (X, d) et (X', d') deux espaces métriques. Soient $Y = X \times X'$ muni de l'une des distances δ définissant la topologie produit. Une suite $(y_n) = (x_n, x'_n)$ de points de Y est de Cauchy dans (Y, δ) ssi les deux suites (x_n) et (x'_n) sont de Cauchy dans X et X' respectivement.

DEFINITION 5.1.2

On dit qu'un espace métrique (X, d) est complet si toute suite de Cauchy de points de X est convergente. Si (X, d) est un espace métrique, une partie A de X est dite complète si l'espace métrique (A, d_A) (où d_A est la distance induite par d sur A) est complet.

5.2 Propriétés des espaces complets

THEOREME 5.2.1

Toute partie complète d'un espace métrique (X, d) est fermée.

preuve

Soient A une partie complète de X et $a \in \overline{A}$. Il existe une suite $(x_n) \in A^{\mathbb{N}}$ qui converge vers a . Cette suite est de Cauchy ; A étant complet, elle converge dans A donc $a \in A$ et $\overline{A} = A$. A est fermé.

THEOREME 5.2.2

Toute partie fermée d'un espace métrique complet est complète.

preuve

Soit F fermé de X complet ; soit $(x_n) \in F^{\mathbb{N}}$ une suite de Cauchy d'éléments de F . X étant complet, elle converge dans X vers une limite a . On a $a \in \overline{F} = F$, donc toute suite de Cauchy de F converge dans F . F est complet.

THEOREME 5.2.3

Tout espace métrique compact est complet.

En effet toute suite dans un espace compact possède une valeur d'adhérence. La conclusion résulte de 5.1 assertion 4.

THEOREME 5.2.4

Le produit de deux espaces métriques complets est un espace métrique complet.

Résulte de 5.1 assertion 7.

COROLLAIRE 5.2.1

\mathbb{R}^n et \mathbb{C}^n sont complets.

Tout espace vectoriel normé de dimension finie est complet.

THEOREME 5.2.5

Soient (X, d) un espace métrique complet et (F_n) une suite décroissante de fermés non vides de X , dont le diamètre tend vers 0. L'intersection des F_n est un singleton.

preuve

Choisissons, pour chaque n un point $x_n \in F_n$. Soit $A_n = \{x_k \mid k \geq n\}$; on a $A_n \subset F_n$, donc $\text{diam}(A_n) \rightarrow 0$ et la suite (x_n) est de Cauchy. Soit a sa limite. Pour tout $p \geq n$, $x_p \in F_p \subset F_n$, donc $a \in \overline{F_n} = F_n$; a appartient à l'intersection des F_n qui est donc non vide. Si a et b sont dans cette intersection, on a $d(a, b) \leq \text{diam}(F_n)$ pour tout n donc $d(a, b) = 0$ soit $a = b$.

Attention! La condition sur le diamètre est indispensable. Une suite décroissante de fermés non vides dans un espace métrique complet peut avoir une intersection vide, comme le prouve la suite $F_n = [n, +\infty[$ dans \mathbb{R} .

THEOREME 5.2.6 (Critère de Cauchy pour les fonctions)

Soient A une partie d'un espace métrique (X, d) , $a \in \overline{A}$, et f une application de A dans un espace métrique complet Y ; alors f a une limite quand x tend vers a si et seulement si :

$$(C) \quad \forall \varepsilon > 0, \exists V \in \mathcal{V}(A) \text{ tel que } \text{diam}(f(V \cap A)) < \varepsilon$$

preuve

Dans le sens non trivial, on prend une suite arbitraire de points $(x_n) \in A^{\mathbb{N}}$ qui converge vers a . Grâce à la condition (C), on voit facilement que la suite $(f(x_n))$ est de Cauchy dans Y complet, donc converge. On sait que ce fait implique que f a une limite quand x tend vers a .

5.3 Théorème du point fixe

THEOREME 5.3.1

Soient A une partie complète non vide d'un espace métrique (X, d) et $f : A \rightarrow A$ une application contractante (c'est à dire k -lipschitzienne avec $k < 1$).

Alors f a un point fixe $a \in A$ et un seul.

Pour tout $x_0 \in A$, la suite (x_n) de premier terme x_0 telle que $\forall n \ x_{n+1} = f(x_n)$ converge vers a et

$$\forall n \ d(x_n, a) \leq \frac{k^n}{1-k} d(x_0, f(x_0))$$

preuve

Soit $x_0 \in A$ et, pour tout n , $x_{n+1} = f(x_n)$. La suite (x_n) est une suite bien définie d'éléments de A . Par hypothèse, on a $d(x_{n+1}, x_n) \leq kd(x_n, x_{n-1})$ pour tout $n \geq 1$ donc $\forall n \ d(x_{n+1}, x_n) \leq k^n d(x_1, x_0)$. Soit alors p entier, $p \geq 1$. L'inégalité triangulaire donne:

$$(*) \quad d(x_{n+p}, x_n) \leq \sum_{j=1}^{j=p} d(x_{n+j}, x_{n+j-1}) \leq \sum_{j=1}^{j=p} k^{n+j-1} d(x_1, x_0) \leq \frac{k^n}{1-k} d(x_1, x_0)$$

Comme k est inférieur à 1, k^n tend vers 0, donc la suite (x_n) est de Cauchy. Elle converge donc vers un élément $a \in A$. Comme f est continue en a , de l'égalité valable pour tout $n \ x_{n+1} = f(x_n)$, on déduit, en passant à la limite, $a = f(a)$, donc a est un point fixe de f .

En faisant tendre p vers l'infini dans (*), on obtient $d(x_n, a) \leq \frac{k^n}{1-k} d(x_1, x_0)$.

Enfin, supposons $f(a) = a$ et $f(b) = b$. On a $d(a, b) = d(f(a), f(b)) \leq kd(a, b)$ soit $(1-k)d(a, b) \leq 0$, ce qui implique $d(a, b) = 0$ soit $a = b$. ■

5.4 Espaces de Banach

DEFINITION 5.4.1

On appelle espace de Banach un \mathbb{K} -espace vectoriel normé complet. En particulier, tout \mathbb{K} -ev de dimension finie est un espace de Banach.

On appelle algèbre de Banach sur le corps \mathbb{K} une \mathbb{K} -algèbre associative unitaire A , munie d'une norme $\| \cdot \|_A$ vérifiant $\|1_A\|_A = 1$ et $\forall u, v \in A, \|uv\|_A \leq \|u\|_A \cdot \|v\|_A$, et telle que l'espace normé $(A, \| \cdot \|_A)$ soit complet.

THEOREME 5.4.1

Soient E un espace de Banach, et F un sous-espace vectoriel de E . F est un espace de Banach si et seulement si F est fermé dans E .

Conséquence immédiate des théorèmes 5.2.1 et 5.2.2.

THEOREME 5.4.2

Soit E un \mathbb{K} -evn, et F un sous-espace vectoriel de E de dimension finie. F est fermé dans E .

preuve

F muni de la norme induite par celle de E est un espace vectoriel normé de dimension finie, donc complet. Le résultat découle du théorème 5.2.1.

THEOREME 5.4.3 (Un théorème de prolongement)

Soit G un espace vectoriel normé, F un sous espace vectoriel de G partout dense dans G et $u : F \rightarrow H$ une application linéaire continue de F dans un espace de Banach H . u se prolonge de manière unique en une application linéaire continue $U : G \rightarrow H$. De plus, $\|U\| = \|u\|$.

preuve

- Soit $x \in G$. Par hypothèse, il existe une suite (x_n) de points de F qui converge vers x .
On a $\|u(x_p) - u(x_q)\|_H \leq \|u\| \cdot \|x_p - x_q\|_G$. Il en résulte que la suite $(u(x_n))$ est de Cauchy dans H complet, donc elle converge. D'après le théorème 1.5.2, la fonction $u : F \rightarrow H$ admet une limite en x . Notons là $U(x)$. Soit $x \in F$ et (x_n) la suite constante telle que $x_n = x$ pour tout n . On a $U(x) = \lim u(x_n) = u(x)$. Donc U prolonge u .
- Soient $x, y \in G$, $\lambda \in \mathbb{K}$, (x_n) et (y_n) deux suites de points de F convergeant respectivement vers x et y . On a $U(x + \lambda y) = \lim u(x_n + \lambda y_n) = \lim (u(x_n) + \lambda u(y_n)) = U(x) + \lambda U(y)$. U est linéaire.

- Avec les mêmes notations, on a $\|u(x_n)\|_H \leq \|u\| \cdot \|x_n\|_G$. Par continuité de la norme, $\|u(x_n)\|_H$ tend vers $\|U(x)\|_H$ quand n tend vers l'infini. On obtient donc en passant à la limite, $\|U(x)\|_H \leq \|u\| \cdot \|x\|_G$ ce qui prouve que U est continue et que $\|U\| \leq \|u\|$. Comme U prolonge u on a (le vérifier!) $\|u\| \leq \|U\|$ d'où l'égalité. ■

Exercice

En s'inspirant de la preuve du théorème ci dessus, prouver le résultat suivant :

Soient (X, d) et (X', d') deux espaces métriques, Y une partie de X dense dans X et $h : Y \rightarrow X'$ une application uniformément continue. Si l'espace (X', d') est complet, h admet un unique prolongement continu $H : X \rightarrow X'$ et ce prolongement H est uniformément continu.

DEFINITION 5.4.2

Soient $(E, \| \cdot \|)$ un \mathbb{K} -evn et (u_n) une suite d'éléments de E . On dit que la série $\sum_{n \geq 0} u_n$ converge si la suite des sommes

partielles $S_n = \sum_{k=0}^{n-1} u_k$ a une limite dans E .

On dit que la série de terme général u_n est absolument convergente si la série numérique $\sum \|u_n\|$ converge.

THEOREME 5.4.4

Toute série absolument convergente d'éléments d'un espace de Banach E est convergente dans E .

preuve

Soit $\varepsilon > 0$. Il existe un $N \in \mathbb{N}$ tel que, pour tout $p \geq N$, $q \geq 0$ on ait $\sum_{k=p+1}^{p+q} \|u_k\| \leq \varepsilon$. Alors $\left\| \sum_{k=p+1}^{p+q} u_k \right\| \leq \varepsilon$ et donc la suite

des sommes partielles $S_n = \sum_{k=0}^{n-1} u_k$ est de Cauchy. E étant complet, elle converge.

Attention! Ce résultat est faux dans un \mathbb{K} -evn qui n'est pas complet. En fait, on peut démontrer le résultat suivant : un \mathbb{K} -evn E est complet si et seulement si toute série absolument convergente d'éléments de E est convergente. Autrement dit, si E n'est pas complet, on peut trouver une série absolument convergente d'éléments de E qui n'est pas convergente. (Voir l'exercice 2 ci dessous)

Exercice 1

Soit $E = \mathbb{R}[X]$ le \mathbb{R} -ev des polynômes à coefficients réels. On pose, pour $P \in E$, $\|P\| = \sup_{0 \leq x \leq 1} |P(x)|$. Vérifier que l'on définit

ainsi une norme sur E . Soit $u_n(X) = X^n/n!$; montrer que la série de terme général u_n est absolument convergente, mais ne converge pas dans $(E, \| \cdot \|)$.

Exercice 2

Soit $(E, \| \cdot \|)$ un \mathbb{K} -espace vectoriel normé tel que toute série absolument convergente d'éléments de E est convergente.

1. Soit (x_n) une suite de Cauchy d'éléments de E . Montrer qu'il existe une suite extraite (x_{n_k}) telle que, pour tout k , $d(x_{n_{k+1}}, x_{n_k}) \leq 2^{-k}$.
2. En déduire que la suite (x_n) possède une valeur d'adhérence puis que E est complet.

6

Suites et séries de fonctions à valeurs dans un espace vectoriel normé

6.1 Convergence simple et convergence uniforme d'une suite de fonctions

Soient X un ensemble, $(E, \|\cdot\|)$ un \mathbb{K} -evn, $(u_n)_{n \geq 0}$ une suite d'applications définies sur X à valeurs dans E et u une application de X dans E .

DEFINITION 6.1.1

On dit que la suite (u_n) converge simplement vers u si pour tout x de E la suite $(u_n(x))$ converge vers $u(x)$. Ce qui signifie:

$$(CVS) \quad \forall x \in X, \forall \varepsilon > 0, \exists N = N(x, \varepsilon) \in \mathbb{N}, \forall n \in \mathbb{N}, n \geq N \Rightarrow \|u_n(x) - u(x)\| \leq \varepsilon$$

Il est essentiel de bien noter que le rang N à partir duquel l'inégalité $\|u_n(x) - u(x)\| \leq \varepsilon$ est vérifiée dépend de ε mais aussi de x . La convergence sera dite uniforme si on peut choisir le N indépendant de x , plus précisément :

DEFINITION 6.1.2

Une suite (u_n) de fonctions définies sur un ensemble X à valeurs dans un \mathbb{K} -evn E est uniformément convergente (sur X) vers une fonction $u : X \rightarrow E$ si pour tout $\varepsilon > 0$ il existe un entier $N(\varepsilon)$ tel que, pour tout $x \in X$ et tout entier $n \geq N$, $\|u_n(x) - u(x)\| \leq \varepsilon$. Ce qui s'écrit

$$(CVU) \quad \forall \varepsilon > 0, \exists N = N(\varepsilon) \in \mathbb{N}, \forall x \in X, \forall n \in \mathbb{N}, n \geq N \Rightarrow \|u_n(x) - u(x)\| \leq \varepsilon$$

On définit de la même manière la convergence uniforme sur un sous ensemble A de X . Il est évident que toute suite uniformément convergente vers u est simplement convergente vers u .

REGLE PRATIQUE

La suite (u_n) d'applications de X dans E converge uniformément vers $u : X \rightarrow E$ si et seulement si il existe un entier n_0 et une suite numérique $(a_n)_{n \geq n_0}$ de limite nulle telle que pour tout $n \geq n_0, \forall x \in X, \|u_n(x) - u(x)\| \leq a_n$.

DEFINITION 6.1.3 (Critère de Cauchy pour la convergence uniforme)

Soient X un ensemble, E un \mathbb{K} -evn et (u_n) une suite de fonctions définies sur X et à valeurs dans E . On dit que cette suite vérifie le critère de Cauchy pour la convergence uniforme si on a :

$$(C) \quad \forall \varepsilon > 0, \exists N = N(\varepsilon), \forall p, q \in \mathbb{N}, p \geq N \Rightarrow (\forall x \in X \|u_{p+q}(x) - u_p(x)\| \leq \varepsilon)$$

Si la suite (u_n) converge uniformément vers u , elle vérifie le critère de Cauchy uniforme. La réciproque est vraie si E est complet :

PROPOSITION 6.1.1

Soient X un ensemble, E un espace **de Banach** et (u_n) une suite de fonctions définies sur X à valeurs dans E et satisfaisant le critère de Cauchy uniforme. La suite (u_n) converge uniformément vers une application $u : X \rightarrow E$.

preuve

D'après (C), pour tout x de X , la suite $(u_n(x))$ est de Cauchy dans E complet, donc converge vers un élément de E , soit $u(x)$. On définit ainsi une application $u : X \rightarrow E$. Soit $\varepsilon > 0$ et N défini par la condition (C). En faisant tendre q vers l'infini dans (C), on obtient $\|u(x) - u_p(x)\| \leq \varepsilon$ pour tout x de X et tout $p \geq N$. ■

Remarque

Toutes les définitions précédentes ont un sens en remplaçant E par un espace métrique (Y, δ) , et les conditions de la forme $\|u(x) - u_p(x)\| \leq \varepsilon$ par $\delta(u(x), u_p(x)) \leq \varepsilon$.

6.2 Conservation de la continuité

THEOREME 6.2.1

Soient (X, d) un espace métrique, E un evn et (u_n) une suite de fonctions définies sur X et à valeurs dans E et vérifiant les propriétés suivantes :

1. Pour tout n , la fonction u_n est continue en x_0 (resp. est continue sur X).
2. La suite (u_n) converge uniformément sur X vers une fonction $u : X \rightarrow E$.

Alors, la fonction u est continue en x_0 (resp. est continue sur X).

preuve

Soient $x_0 \in X$ et $\varepsilon > 0$. La convergence uniforme de u_n vers u implique l'existence d'un entier N tel qu'on ait, pour tout $x \in X$: $\|u_n(x) - u(x)\| < \varepsilon/3$ dès que $n \geq N$. La continuité au point x_0 de la fonction u_N assure l'existence d'un voisinage V de x_0 tel que $x \in V \Rightarrow \|u_N(x) - u_N(x_0)\| \leq \varepsilon/3$. Alors, pour tout $x \in V$ on a

$$\|u(x) - u(x_0)\| \leq \|u(x) - u_N(x)\| + \|u_N(x) - u_N(x_0)\| + \|u_N(x_0) - u(x_0)\| \leq \varepsilon$$

Donc u est continue en x_0 . ■

Remarques

Dans l'énoncé du théorème, on peut remplacer la condition 1. par : " il existe un rang n_0 tel que pour tout $n \geq n_0$, u_n est continue sur X ".

Le théorème reste vrai si les fonctions u_n sont à valeurs dans un espace métrique (Y, δ) .

DEFINITION 6.2.1

Soient (u_n) une suite de fonctions définies sur l'espace métrique (X, d) à valeurs dans l'evn E et $u : X \rightarrow E$. On dit que la suite (u_n) converge uniformément sur tout compact vers u si, pour tout compact K de X , la suite $(u_n|_K)$ des restrictions de u_n à K converge uniformément vers la restriction de u à K .

THEOREME 6.2.2

Soient (X, d) un espace métrique, E un evn et (u_n) une suite de fonctions continues définies sur X , à valeurs dans E et convergeant uniformément sur tout compact vers $u : X \rightarrow E$. La fonction u est continue sur X .

preuve

D'après le théorème précédent, pour tout compact K de X la restriction de u à K est continue. X étant un espace métrique, on sait que cela implique la continuité de u sur X . (proposition 3.1.1) .■

6.3 Echange de limites

THEOREME 6.3.1

Soient (X, d) un espace métrique, $A \subset X$, $a \in \bar{A}$ et E un espace de Banach.

Soit (f_n) une suite d'applications de A dans E vérifiant les propriétés suivantes :

- (1) Pour tout n , $f_n(x)$ a une limite λ_n quand x tend vers a : $\lambda_n = \lim_{x \rightarrow a} f_n(x)$.
- (2) La suite (f_n) converge uniformément vers une fonction $f : A \rightarrow E$.

Alors,

1. La suite (λ_n) a une limite λ .
2. La fonction f a une limite quand x tend vers a .
3. $\lambda = \lim_{x \rightarrow a} f(x)$.

Autrement dit, on a $\lim_{x \rightarrow a} \left(\lim_{n \rightarrow \infty} f_n(x) \right) = \lim_{n \rightarrow \infty} \left(\lim_{x \rightarrow a} f_n(x) \right)$.

preuve

1. Soit $\varepsilon > 0$. Il existe un entier $N(\varepsilon)$, tel que $\forall p, q \in \mathbb{N}, p \geq N(\varepsilon) \Rightarrow \|f_{p+q}(x) - f_p(x)\| \leq \varepsilon$ pour tout $x \in A$. En faisant tendre x vers a pour p, q fixés, on obtient $\|\lambda_{p+q} - \lambda_p\| \leq \varepsilon$ dès que $p \geq N(\varepsilon)$. La suite (λ_n) est de Cauchy dans E complet, donc converge vers un élément λ de E .
2. Soit $p = N(\varepsilon/3)$. En faisant tendre q vers l'infini dans les inégalités ci dessus, on obtient $\|\lambda - \lambda_p\| \leq \varepsilon/3$ et $\|f(x) - f_p(x)\| \leq \varepsilon/3$ pour tout $x \in A$. Puisque $\lambda_p = \lim_{x \rightarrow a} f_p(x)$, il existe un voisinage V de a tel que pour $x \in V \cap A$ on ait $\|f_p(x) - \lambda_p\| \leq \varepsilon/3$. On en déduit que pour $x \in V \cap A$, on a $\|f(x) - \lambda\| \leq \varepsilon$.

Remarques

- a) Si $a \in A$, on a nécessairement $\lambda_n = f_n(a)$, les f_n sont continues en a et le théorème n'est autre que le théorème 6.2.1 et l'hypothèse E complet est inutile.
- b) La preuve ci dessus montre que la complétude de E n'a servi que dans le point 1. Donc si on suppose (dans les hypothèses du théorème) que la suite (λ_n) converge dans E vers une limite λ , il n'y a pas besoin de supposer E complet. En fait ce cas est un cas particulier du précédent, obtenu en remplaçant A par $A' = A \cup \{a\}$ et les f_n par leurs prolongements par continuité à A' obtenu en posant $f_n(a) = \lambda_n$.

6.4 Séries de fonctions à valeurs dans un Banach

Soient X un ensemble, E un evn, et u_n une suite d'applications de X vers E . On pose, pour tout n , $S_n = \sum_{0 \leq k \leq n} u_k$.

DEFINITION 6.4.1

On dit que la série $\sum u_n$ converge simplement si pour tout $x \in X$, la série $\sum u_n(x)$ converge, c'est à dire si la suite S_n converge simplement vers une fonction $S : X \rightarrow E$. On dit que la série $\sum u_n$ converge uniformément si la suite (S_n) converge uniformément vers S .

Donc, $\sum u_n$ converge uniformément ssi 1) $\sum u_n$ converge simplement et 2) la suite $R_n = \sum_{k \geq n+1} u_k$, qui est donc définie, converge uniformément vers 0.

DEFINITION 6.4.2

On dit que la série converge absolument simplement si la série $\sum \|u_n(x)\|$ converge pour tout x .

Si E est complet, la convergence absolue simple de la série entraîne la convergence simple.

DEFINITION 6.4.3

On dit que la série $\sum u_n$ converge normalement si les fonctions u_n sont bornées et si la série numérique de terme général $\|u_n\|_\infty = \sup_{x \in X} \|u_n(x)\|$ converge. On définit de la même manière la convergence normale sur un sous-ensemble A de X en considérant les restrictions de u_n à A .

REGLE PRATIQUE

La série $\sum u_n$ est normalement convergente ssi il existe une suite numérique $(a_n)_{n \geq 0}$ telle que :

1. La série $\sum_{n \geq 0} a_n$ converge.
2. $\forall x \in X, \forall n \in \mathbb{N}, \|u_n(x)\| \leq a_n$.

THEOREME 6.4.1

Soient E un espace de Banach et $u_n : X \rightarrow E$ une suite d'applications bornées d'un ensemble X dans E . Si la série $\sum u_n$ est normalement convergente sur X , elle est uniformément convergente sur X .

preuve

La convergence normale implique, de manière évidente, la convergence absolue simple, donc, E étant complet, la convergence simple. Soient $R_n = \sum_{k \geq n+1} u_k : X \rightarrow E$ le reste d'ordre n de la série et $\rho_n = \sum_{k \geq n+1} \sup_{x \in A} \|u_k(x)\|$. La suite (ρ_n) tend vers 0 et pour tout $x \in X, \|R_n(x)\| \leq \rho_n$ ce qui assure la convergence uniforme sur X de la suite (R_n) vers 0. ■

Les résultats du paragraphe précédent se traduisent pour les séries par les théorèmes suivants.

THEOREME 6.4.2

Soient (X, d) un espace métrique, E un \mathbb{K} -espace vectoriel normé, $u_n : X \rightarrow E$ une suite de fonctions continues telles que la série $\sum u_n$ converge uniformément. La somme $S = \sum_{k \geq 0} u_k$ est continue sur X .

Attention! Ici, on doit supposer que **toutes** les u_n sont continues pour être sûr de la continuité des sommes partielles.

COROLLAIRE 6.4.1

Soit (X, d) un espace métrique, E **un espace de Banach**, $u_n : X \rightarrow E$ une suite de fonctions continues telles que la série $\sum u_n$ soit normalement convergente sur X . La somme $S = \sum_{k \geq 0} u_k$ existe et est continue sur X .

THEOREME 6.4.3

Soient (X, d) un espace métrique, E un \mathbb{K} -espace vectoriel normé, $u_n : X \rightarrow E$ une suite de fonctions continues telles que la série $\sum u_n$ converge uniformément sur tout compact de X . La somme $S = \sum_{k \geq 0} u_k$ est continue sur X .

COROLLAIRE 6.4.2

Soit (X, d) un espace métrique, E **un espace de Banach**, $u_n : X \rightarrow E$ une suite de fonctions continues telles que la série $\sum u_n$ soit normalement convergente sur tout compact de X . La somme $S = \sum_{k \geq 0} u_k$ existe et est continue sur X .

7

Exemples d'espaces de Banach

7.1 L'espace $\mathcal{B}(X, E)$

Soient X un ensemble et E un \mathbb{K} -evn ; on notera comme d'habitude E^X l'espace vectoriel des fonctions définies sur X à valeurs dans E . $\mathcal{B}(X, E)$ désignera le sous espace de E^X formé des applications bornées sur X .

THEOREME 7.1.1

L'application $f \rightarrow \|f\|_\infty = \sup_{x \in E} \|f(x)\|$ est une norme sur $\mathcal{B}(X, E)$, appelée norme de la convergence uniforme.

Si E est un espace de Banach, $\mathcal{B}(X, E)$ muni de la norme de la convergence uniforme est un espace de Banach.

preuve

Le fait que $\|\cdot\|_\infty$ soit une norme est facile. Montrons la complétude. Soit (f_n) une suite de Cauchy dans $(\mathcal{B}(X, E), \|\cdot\|_\infty)$. On a donc :

$$\forall \varepsilon > 0, \exists N = N(\varepsilon) \in \mathbb{N}, \forall p, q \in \mathbb{N}, p \geq N \Rightarrow \|f_{p+q} - f_p\|_\infty \leq \varepsilon$$

Soit $\varepsilon > 0$. Pour tout $x \in X$ et tout $p \geq N(\varepsilon)$, $\|f_{p+q}(x) - f_p(x)\| \leq \varepsilon$. La suite $(f_n(x))$ est de Cauchy dans E complet, donc convergente ; notons $f(x)$ sa limite. En faisant tendre q vers l'infini dans l'inégalité précédente, on obtient que pour tout x et tout $p \geq N(\varepsilon)$, $\|f(x) - f_p(x)\| \leq \varepsilon$. On en déduit d'une part que $f_{N(\varepsilon)} - f \in \mathcal{B}(X, E)$, donc que $f \in \mathcal{B}(X, E)$, et d'autre part que $\|f - f_p\|_\infty \leq \varepsilon$ pour $p \geq N(\varepsilon)$, ce qui prouve que (f_n) converge vers f dans $(\mathcal{B}(X, E), \|\cdot\|_\infty)$. ■

Remarque

Si (f_n) est une suite d'éléments de $\mathcal{B}(X, E)$ et $f \in \mathcal{B}(X, E)$, f_n converge vers f dans $(\mathcal{B}(X, E), \|\cdot\|_\infty)$ si et seulement si la suite f_n converge uniformément vers f . Mais si on se donne une suite de fonctions (u_n) , $u_n \in E^X$ convergeant uniformément vers $u \in E^X$, les fonctions u_n et u ne sont pas nécessairement bornées. Par contre la différence $u_n - u \in \mathcal{B}(X, E)$ à partir d'un certain rang et tend vers 0 dans cet espace muni de la norme $\|\cdot\|_\infty$.

Dans les cas $E = \mathbb{K}$ la multiplication des fonctions (qui à f et g associe fg telle que $(fg)(x) = f(x)g(x)$) munit l'espace $\mathcal{B}(X, \mathbb{K})$ d'une structure de \mathbb{K} -algèbre. Pour $f, g \in \mathcal{B}(X, \mathbb{K})$ on a $\|fg\|_\infty \leq \|f\|_\infty \cdot \|g\|_\infty$ et aussi $\|1\|_\infty = 1$ en désignant par 1 la fonction constante égale à 1. C'est donc une algèbre de Banach.

Exemple

On désigne par $l^\infty(\mathbb{K})$ le sous-espace vectoriel de $\mathbb{K}^{\mathbb{N}}$ formé des suites bornées. Si $u = (u_n)_{n \geq 0} \in l^\infty(\mathbb{K})$, on pose $\|u\|_\infty = \sup_{n \geq 0} |u_n|$. L'espace $l^\infty(\mathbb{K})$ muni de la norme $\|\cdot\|_\infty$ est un espace de Banach.

On désigne par $c_0(\mathbb{K})$ l'espace des suites d'éléments de \mathbb{K} qui convergent vers 0. C'est un sous-espace vectoriel fermé de $l^\infty(\mathbb{K})$ muni de $\|\cdot\|_\infty$ (cela résulte du théorème d'échange des limites 6.3.1) donc, muni de cette norme, c'est un espace de Banach.

7.2 L'espace $\mathcal{CB}(X, E)$

THEOREME 7.2.1

Soit (X, d) un espace métrique et E un espace de Banach. L'espace $\mathcal{CB}(X, E)$ des applications continues et bornées de X dans E muni de la norme de la convergence uniforme est un espace de Banach.

preuve

$(\mathcal{B}(X, E), \|\cdot\|_\infty)$ étant complet, pour montrer que $(\mathcal{CB}(X, E), \|\cdot\|_\infty)$ est un espace de Banach, il suffit de montrer que c'est un sous espace fermé de $(\mathcal{B}(X, E), \|\cdot\|_\infty)$. C'est clairement un sous espace vectoriel. Montrons qu'il est fermé. Soit $u \in \mathcal{B}(X, E)$ un élément adhérent à $\mathcal{CB}(X, E)$. Il existe une suite (u_n) d'éléments de $\mathcal{CB}(X, E)$ qui converge vers u dans $\mathcal{B}(X, E)$, i.e. qui converge uniformément vers u . D'après le théorème 6.2.1 u est continue, donc $u \in \mathcal{CB}(X, E)$. ■

$\mathcal{CB}(X, \mathbb{K})$ avec la multiplication usuelle des fonctions est une algèbre de Banach.

COROLLAIRE 7.2.1

Soit (X, d) un espace métrique compact et E un espace de Banach. L'espace $\mathcal{C}(X, E)$ des applications continues de X dans E muni de la norme de la convergence uniforme est un espace de Banach.

7.3 L'espace $L_c(E, F)$

THEOREME 7.3.1

Soient E et F deux \mathbb{K} -evn, F étant complet. L'espace $L_c(E, F)$ des applications linéaires continues de E dans F muni de la norme subordonnée aux normes de E et F est un espace de Banach.

preuve

Soient (f_n) une suite de Cauchy d'éléments de $L_c(E, F)$ et B la boule unité fermée de E .

- Pour tout x de E l'inégalité $\|f_{n+p}(x) - f_n(x)\|_F \leq \|f_{n+p} - f_n\| \cdot \|x\|_E$ montre que la suite $(f_n(x))_{n \geq 0}$ est de Cauchy dans F complet, donc elle converge vers un élément $f(x) \in F$.
- Si $x, y \in E$ et $\lambda \in \mathbb{K}$, on a $f(\lambda x + y) = \lim f_n(\lambda x + y) = \lim (\lambda f_n(x) + f_n(y)) = \lambda f(x) + f(y)$; f est linéaire.
- Soit $\varepsilon > 0$ et N entier tel que $\|f_{n+p} - f_n\| \leq \varepsilon$ pour tout $n \geq N$, $p \geq 0$. Par définition de la norme dans $L_c(E, F)$, on a pour tout $x \in B$ $\|f_{n+p}(x) - f_n(x)\|_F \leq \|f_{n+p} - f_n\| \leq \varepsilon$. En passant à la limite, quand p tend vers l'infini, n étant fixé, on obtient $\|(f - f_n)(x)\| \leq \varepsilon$ pour tout $n \geq N$ et tout x de B . On en déduit d'abord que l'application linéaire $(f - f_N)$ est continue, donc aussi que f est continue, puis que $\|f - f_n\| \leq \varepsilon$ pour $n \geq N$ ce qui prouve la convergence de la suite (f_n) vers f dans $L_c(E, F)$. ■

COROLLAIRE 7.3.1

Soit E un \mathbb{K} -evn ; le dual topologique $E' = L_c(E, \mathbb{K})$ est un espace de Banach.

Soit E un espace de Banach. Notons $L_c(E) = L_c(E, E)$ l'ensemble des endomorphismes continus de E . Muni des opérations usuelles, c'est une algèbre associative unitaire, dont l'unité est Id_E . La norme subordonnée à celle de E vérifie $\|Id_E\| = 1$ et $\forall u, v \in L_c(E, F)$, $\|uv\| \leq \|u\| \|v\|$. Enfin, d'après le paragraphe précédent, c'est un espace complet. On a prouvé :

THEOREME 7.3.2

L'algèbre des endomorphismes continus d'un espace de Banach E muni de la norme subordonnée à celle de E est une algèbre de Banach.

7.4 L'espace $l^1(\mathbb{K})$

DEFINITION 7.4.1

On désigne par $l^1(\mathbb{K})$ l'ensemble formé des suites $(u_n)_{n \geq 0} \in \mathbb{K}^{\mathbb{N}}$ telles que la série $\sum_{n \geq 0} |u_n|$ converge.

C'est un \mathbb{K} -ev et l'application $u \rightarrow \|u\|_1 = \sum_{k \geq 0} |u_k|$ est une norme sur $l^1(\mathbb{K})$.

THEOREME 7.4.1

L'espace $l^1(\mathbb{K})$ muni de la norme $\|\cdot\|_1$ est complet.

preuve

Soit $(u^{(p)})_{p \geq 0}$ une suite de Cauchy d'éléments de l^1 .

Chaque $u^{(p)}$ est donc une suite $u^{(p)} = (u_n^{(p)})_{n \geq 0}$ telle que la série $\sum_{n \geq 0} |u_n^{(p)}|$ converge.

L'hypothèse que la suite $(u^{(p)})$ est de Cauchy se traduit par : pour tout $\varepsilon > 0$, il existe un entier $P(\varepsilon)$ tel que, pour tout $p \geq P(\varepsilon)$ et tout $q \geq 0$ on ait

$$\|u^{(p+q)} - u^{(p)}\|_1 = \sum_{n=0}^{\infty} |u_n^{(p+q)} - u_n^{(p)}| \leq \varepsilon$$

Fixons ε . Pour tout $n \in \mathbb{N}$, $p \geq P(\varepsilon)$, $q \geq 0$, on a $|u_n^{(p+q)} - u_n^{(p)}| \leq \varepsilon$, ce qui prouve que la suite $(u_n^{(p)})_{p \geq 0}$ est (n fixé) de Cauchy dans \mathbb{K} . Soit λ_n sa limite et $\lambda = (\lambda_n)_{n \geq 0}$.

On va montrer que $\lambda \in l^1$ et que la suite $(u^{(p)})_{p \geq 0}$ converge dans l^1 vers λ .

Soient $N \in \mathbb{N}$ fixé et $p \geq P(\varepsilon)$; pour tout $q \geq 0$ on a $\sum_{n=0}^{n=N} |u_n^{(p+q)} - u_n^{(p)}| \leq \|u^{(p+q)} - u^{(p)}\|_1 \leq \varepsilon$. Si on fait tendre q vers l'infini, on obtient

$$(*) \quad \sum_{n=0}^{n=N} |\lambda_n - u_n^{(p)}| \leq \varepsilon$$

Pour tout $p \geq P(\varepsilon)$ fixé, cette inégalité a lieu pour tout entier N ; la suite des sommes partielles de la série à termes positifs $x_n = |\lambda_n - u_n^{(p)}|$ est bornée. Donc cette série converge. Il en résulte que $(\lambda - u^{(P(\varepsilon))}) \in l^1$ donc que $\lambda \in l^1$.

Ensuite, en faisant tendre N vers l'infini dans l'inégalité (*) on obtient $\sum_{n=0}^{\infty} |\lambda_n - u_n^{(p)}| \leq \varepsilon$, soit $\|\lambda - u^{(p)}\|_1 \leq \varepsilon$, et ceci pour tout $p \geq P(\varepsilon)$, ce qui prouve que la suite $(u^{(p)})$ converge vers λ dans l^1 . ■

8

Approximation des fonctions à valeurs dans un espace vectoriel de dimension finie

8.1 Introduction et notations

Dans tout ce chapitre, $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} . E désigne un \mathbb{K} -espace vectoriel de dimension finie. On sait que toutes les normes sur E sont équivalentes. On choisira à chaque fois que l'on en aura besoin une norme sur E qui sera notée $\| \cdot \|$.

Toute application linéaire de E dans un autre \mathbb{K} -espace vectoriel est continue ; en particulier, toutes les formes linéaires sur E sont continues. Si $\mathcal{B} = (e_1, \dots, e_n)$ est une base de E et si $(\varepsilon_1, \dots, \varepsilon_n)$ est la base duale, les applications ε_i sont continues.

Si (X, d) est un espace métrique et si $f : X \rightarrow E$ est une application qui s'écrit en coordonnées $f(x) = f_1(x)e_1 + \dots + f_n(x)e_n$, f est continue ssi les n applications $f_i : X \rightarrow \mathbb{K}$ sont continues. Les $f_i = \varepsilon_i \circ f$ sont les applications coordonnées de f dans la base \mathcal{B} . On a un résultat similaire pour l'existence d'une limite en un point.

Soit X un ensemble non vide et E^X l'ensemble formé de toutes les applications de X dans E . Si $f, g \in E^X$ et $\lambda \in \mathbb{K}$ on définit $f + g$ et λf par $(f + g)(x) = f(x) + g(x)$ et $(\lambda f)(x) = \lambda \cdot f(x)$. Muni de ces opérations E^X a une structure de \mathbb{K} -espace vectoriel.

8.2 Théorèmes d'approximation

8.2.1 Subdivisions

Soit $J = [a, b]$ un intervalle compact de \mathbb{R} . Une subdivision σ de J est un sous ensemble fini de J contenant les extrémités a et b ; on notera une telle subdivision $\sigma = \{x_0 = a, x_1, \dots, x_n = b\}$. On supposera toujours, le plus souvent implicitement, que $x_0 < x_1 < \dots < x_n$. Une subdivision τ est dite plus fine que la subdivision σ si $\sigma \subset \tau$.

8.2.2 Fonctions en escalier

Une fonction $\varphi : J \rightarrow E$ est dite en escalier si il existe une subdivision $\sigma = \{x_0 = a, x_1, \dots, x_n = b\}$ telle que φ soit constante sur chacun des intervalles $]x_i, x_{i+1}[$ pour $0 \leq i \leq n - 1$. Une telle subdivision σ sera dite adaptée à φ . Si τ est une subdivision plus fine que σ , τ est aussi adaptée à φ .

Soient φ, ψ deux fonctions en escalier, σ, τ des subdivisions adaptées respectivement à φ et ψ . La subdivision $\sigma \cup \tau$ est adaptée à la fois à φ et à ψ . On en déduit que l'ensemble $\mathcal{E}(J, E)$ des fonctions en escalier sur J à valeurs dans E est un sous espace vectoriel de E^J . Si $\varphi \in \mathcal{E}(J, E)$ alors la fonction $x \rightarrow \|\varphi(x)\|$ appartient à $\mathcal{E}(J, \mathbb{R})$. Si $\varphi \in \mathcal{E}(J, E)$ et $h \in \mathcal{E}(J, \mathbb{K})$, la fonction $x \rightarrow h(x)\varphi(x)$ est en escalier. En particulier, $\mathcal{E}(J, \mathbb{K})$ muni en outre de la multiplication usuelle des fonctions est une \mathbb{K} -algèbre.

8.2.3 Fonctions continues par morceaux

Une fonction $f : J \rightarrow E$ est dite continue par morceaux sur J si il existe une subdivision $\sigma = \{x_0 = a, x_1, \dots, x_n = b\}$ de J et des fonctions continues $f_i : [x_i, x_{i+1}] \rightarrow E$, $0 \leq i \leq n - 1$ telle que f coïncide avec f_i sur l'intervalle ouvert $]x_i, x_{i+1}[$.

Il revient au même de dire que f est continue sur $J \setminus \sigma$ et qu'elle admet en chaque x_i des limites finies à droite (sauf pour $i = n$) et à gauche (sauf pour $i = 0$).

Toute fonction continue sur J (respectivement en escalier sur J) est continue par morceaux.

L'ensemble $C_M(J, E)$ des fonctions continues par morceaux sur J à valeurs dans E est un sous espace vectoriel de E^J qui contient comme sous espace $\mathcal{E}(J, E)$ et l'espace vectoriel $C^0(J, E)$ des fonctions continues sur J à valeurs dans E .

Si $f \in C_M(J, E)$, la fonction $x \rightarrow \|f(x)\|$ appartient à $C_M(J, \mathbb{R})$.

$C_M(J, \mathbb{K})$ est une \mathbb{K} -algèbre dont $\mathcal{E}(J, \mathbb{K})$ et $C^0(J, \mathbb{K})$ sont des sous algèbres.

On notera, lorsqu'il n'y a pas d'ambiguïtés C_M, C^0, \mathcal{E} pour $C_M(J, E), C^0(J, E), \mathcal{E}(J, E)$

Attention! Une fonction en escalier ou une fonction continue par morceaux est par définition définie sur un intervalle fermé borné de \mathbb{R} .

PROPOSITION 8.2.1

♠ Toute fonction continue par morceaux sur J est la somme d'une fonction continue et d'une fonction en escalier. Cette décomposition n'est pas unique. Autrement dit on a $C_M(J, E) = C^0(J, E) + \mathcal{E}(J, E)$ la somme n'étant pas directe.

♠ Toute fonction continue par morceaux sur J est bornée.

L'expression $\|f\|_\infty = \sup_{t \in J} \|f(t)\|$ définit une norme sur $C_M(J, E)$ (norme de la convergence uniforme).

Remarque et exercice Cette norme sur $C_M(J, E)$ dépend de la norme choisie sur E . Montrer que si on remplace la norme $\|\cdot\|$ par une norme $\|\cdot\|'$ les normes obtenues sur $C_M(J, E)$ sont équivalentes.

THEOREME 8.2.1 (Approximation des fonctions continues par morceaux par des fonctions en escalier)

Pour toute $f \in C_M([a, b], E)$ et tout $\varepsilon > 0$ il existe une fonction $\varphi \in \mathcal{E}([a, b], E)$ telle que $\|f - \varphi\|_\infty \leq \varepsilon$, i.e. telle que $\|f(x) - \varphi(x)\| \leq \varepsilon$ pour tout $x \in J$.

Ce théorème admet les deux énoncés équivalents suivants:

(1) \mathcal{E} est dense dans C_M pour la topologie définie par la norme $\|\cdot\|_\infty$.

(2) Pour toute $f \in C_M$, il existe une suite $(\varphi_n)_{n \geq 0}$ de fonctions en escalier convergeant uniformément vers f .

preuve

Soit f continue par morceaux sur J . Il existe une fonction en escalier ψ_1 et une fonction continue g telles que $f = g + \psi_1$. Il suffit donc de montrer qu'il existe une fonction en escalier ψ telle que $\|g - \psi\|_\infty \leq \varepsilon$. La fonction $\varphi = \psi + \psi_1$ vérifiera alors $\|f - \varphi\|_\infty \leq \varepsilon$.

g est continue sur $J = [a, b]$ compact donc uniformément continue ; il existe un $\eta > 0$ tel qu'on ait $\|g(u) - g(v)\| \leq \varepsilon$ pour toute paire de points $u, v \in J$ vérifiant $|u - v| < \eta$. Soit $n \in \mathbb{N}^*$ tel que $(b - a)/n < \eta$ et pour $k = 0, \dots, n$, $x_k = a + k \frac{b-a}{n}$. Soit ψ la fonction en escalier définie sur J par $\psi(x) = g(x_k)$ si $x_k \leq x < x_{k+1}$ et $\psi(b) = g(b)$. On vérifie facilement que pour tout $x \in J$ on a $\|g(x) - \psi(x)\| \leq \varepsilon$. ■

8.2.4 Fonctions continues et affines par morceaux

Nous dirons qu'une fonction $F : [a, b] \rightarrow E$ est affine par morceaux si il existe une subdivision $\sigma = \{x_0 = a, x_1, \dots, x_n = b\}$ de J et des fonctions affines $A_k, 0 \leq k \leq n - 1$ telles que $F(x) = A_k(x)$ pour $x_k < x < x_{k+1}$. (Les fonctions A_k sont donc de la forme $A_k(x) = xu_k + v_k$, avec $u_k, v_k \in E$). Si en outre la fonction F est continue, on a évidemment $F(x) = A_k(x)$ pour $x_k \leq x \leq x_{k+1}$ et donc nécessairement $A_k(x_k) = A_{k-1}(x_k)$ pour $1 \leq k \leq n - 1$.

THEOREME 8.2.2

Soit $f : [a, b] \rightarrow E$ continue. Pour tout $\varepsilon > 0$ il existe une fonction $F : [a, b] \rightarrow E$ affine par morceaux et continue telle que $\|f - F\|_\infty \leq \varepsilon$.

preuve

Par continuité uniforme de f il existe un $n > 0$ tel que $\|f(v) - f(u)\| \leq \varepsilon$ pour $|v - u| \leq (b - a)/n$. On pose comme ci dessus $x_k = a + k \frac{b-a}{n}$. Il existe alors une unique fonction F affine par morceaux et continue sur J telle que $F(x_k) = f(x_k)$ pour $0 \leq k \leq n$. Soit $x \in I$. Si x est l'un des x_k on a $F(x) = f(x)$. Sinon, il existe un unique k tel que $x_k < x < x_{k+1}$. On a

alors $\|f(x) - f(x_k)\| \leq \varepsilon$ et $\|f(x) - f(x_{k+1})\| \leq \varepsilon$. Il existe un unique $t \in]0, 1[$ tel que $x = tx_k + (1-t)x_{k+1}$. Puisque F est affine sur $[x_k, x_{k+1}]$ on a $F(x) = tF(x_k) + (1-t)F(x_{k+1}) = tf(x_k) + (1-t)f(x_{k+1})$ et donc $\|f(x) - F(x)\| = \|t(f(x) - f(x_k)) + (1-t)(f(x) - f(x_{k+1}))\| \leq t\varepsilon + (1-t)\varepsilon = \varepsilon$. ■

8.2.5 Fonctions polynômes

THEOREME 8.2.3 (Théorème de Weierstrass)

Soit $J = [a, b]$ un intervalle fermé borné de \mathbb{R} et $f : [a, b] \rightarrow \mathbb{R}$ continue. Pour tout $\varepsilon > 0$, il existe un polynôme $P \in \mathbb{R}[X]$ tel que $\forall x \in [a, b]$, $|f(x) - P(x)| \leq \varepsilon$.

Ce théorème est admis par le programme. Donnons deux énoncés équivalents:

- Pour toute fonction $f : [a, b] \rightarrow \mathbb{R}$ continue, il existe une suite P_n de polynômes tels que la suite des fonctions polynômes P_n converge uniformément sur $[a, b]$ vers f .
- Le sous espace vectoriel des fonctions polynômes sur $[a, b]$ est dense dans l'espace vectoriel $C^0([a, b], \mathbb{R})$ des fonctions continues sur $[a, b]$, à valeurs réelles, muni de la norme $\|\cdot\|_\infty$ de la convergence uniforme.

Attention! L'énoncé analogue est faux si on remplace $[a, b]$ par un intervalle non compact. Par exemple, la fonction $x \rightarrow e^x$ n'est pas approchée uniformément par des polynômes sur $[0, +\infty[$. En effet, si on suppose $|e^x - P(x)| \leq \varepsilon$ pour tout $x \geq 0$, on obtient en divisant par e^x , $|1 - e^{-x}P(x)| \leq e^{-x}\varepsilon$ pour tout $x \geq 0$ d'où une contradiction en faisant tendre x vers l'infini.

Nous dirons qu'une fonction $P : J \rightarrow E$ est polynomiale si il existe une base $\mathcal{B} = (e_1, \dots, e_n)$ de E telle que les fonctions coordonnées de P dans cette base soient des fonctions polynômes. Il en est alors ainsi dans toute base de E (le vérifier!). Du théorème de Weierstrass on déduit immédiatement :

COROLLAIRE 8.2.1

Soit $f : [a, b] \rightarrow E$ continue ; pour tout $\varepsilon > 0$ il existe une fonction polynomiale $P : [a, b] \rightarrow E$ telle que $\|f - P\|_\infty \leq \varepsilon$.

8.3 Complément : polynômes de Bernstein

On se propose de donner une preuve du théorème de Weierstrass.

Soit $I = [0, 1]$ et $f : I \rightarrow \mathbb{R}$ continue. On définit la suite des polynômes de Bernstein de f , $B_n(f)$ par

$$B_n(f)(x) = \sum_{k=0}^{k=n} C_n^k f\left(\frac{k}{n}\right) x^k (1-x)^{n-k}$$

THEOREME 8.3.1

Sous les hypothèses ci dessus, la suite $B_n(f)$ converge uniformément vers f sur $[0, 1]$.

preuve

Soit $\varepsilon > 0$. On veut prouver l'existence d'un entier N tel qu'on ait, pour tout entier $n \geq N$

$$(C) \quad \forall x \in I, |B_n(f)(x) - f(x)| \leq \varepsilon$$

Comme $B_n(f)(0) = f(0)$ et $B_n(f)(1) = f(1)$, l'inégalité figurant dans (C) est valable pour tout n , pour $x = 0$ ou 1. On peut donc se limiter à $x \in]0, 1[$.

La fonction f est continue sur I intervalle compact, donc uniformément continue. Soit $\eta > 0$ tel que

$$\forall u, v \in I, |v - u| \leq \eta \Rightarrow |f(v) - f(u)| \leq \frac{\varepsilon}{2} \quad (8.1)$$

Fixons un x arbitraire dans $]0, 1[$. Soit X_n une variable aléatoire de loi binomiale de paramètres n et x : X_n prend ses valeurs dans $\{0, 1, \dots, n\}$ et $P(X_n = k) = C_n^k x^k (1-x)^{n-k}$. On a alors $B_n(f)(x) = \sum_{0 \leq k \leq n} f(k/n) P(X_n = k)$. On voit donc que

$B_n(f)(x)$ est l'espérance de la variable aléatoire $Y_n = f(X_n/n)$.

Par conséquent, $B_n(f)(x) - f(x) = E(f(X_n/n) - f(x))$.

Soit A l'événement $|X_n/n - x| \geq \eta$. Si A n'est pas réalisé, on a d'après (8.1), $|Y_n - f(x)| \leq \varepsilon/2$ alors que, dans le cas contraire on a $|Y_n - f(x)| \leq 2\|f\|_\infty$. Introduisons la variable aléatoire Z égale à $2\|f\|_\infty$ sur A et à $\frac{\varepsilon}{2}$ sur le complémentaire de A . On a

$$|Y_n - f(x)| \leq Z$$

donc

$$|E(Y_n) - f(x)| \leq E(|Y_n - f(x)|) \leq E(Z) = (1 - P(A))\varepsilon/2 + 2\|f\|_\infty P(A) \leq \frac{\varepsilon}{2} + 2\|f\|_\infty P(A) \quad (8.2)$$

On majore $P(A)$ grâce à l'inégalité de Tchebycheff : A est l'événement $|X_n/n - E(X_n/n)| \geq \eta$ donc

$$P(A) \leq \frac{1}{\eta^2} \text{Var}\left(\frac{X_n}{n}\right) = \frac{nx(1-x)}{n^2\eta^2} \leq \frac{1}{4n\eta^2}$$

Choisissons $N \in \mathbb{N}$ tel que $N > \|f\|_\infty / \varepsilon\eta^2$. Cet entier N est indépendant de x et pour $n \geq N$ on a $P(A) \leq \frac{\varepsilon}{4\|f\|_\infty}$ donc compte tenu de (8.2) $|B_n(f)(x) - f(x)| = |E(Y_n) - f(x)| \leq \varepsilon$ d'où la conclusion. ■

On en déduit le théorème de Weierstrass dans le cas général au moyen d'un changement de variable affine. Si $g : [a, b] \rightarrow \mathbb{R}$ est continue, soit h définie par $h(t) = a + t(b - a)$. C'est une bijection affine de $[0, 1]$ sur $[a, b]$. La fonction $f = g \circ h$ est continue sur $[0, 1]$ donc la suite de polynômes $P_n = B_n(f)$ converge uniformément vers f sur $[0, 1]$. On en déduit facilement que la suite de polynômes $Q_n = P_n \circ h^{-1}$ converge uniformément sur $[a, b]$ vers g .

8.4 Complément : une autre démonstration du théorème de Weierstrass

On va donner une autre preuve du théorème de Weierstrass, utilisant le produit de convolution des fonctions. Une preuve du même type est utilisée dans la théorie des séries de Fourier pour prouver que les sommes de Fejer d'une fonction périodique continue convergent uniformément vers cette fonction.

1) On commence par le cas d'une fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ continue et nulle hors de l'intervalle $J = [-1/2, 1/2]$.

Soit, pour $n \in \mathbb{N}^*$, φ_n la fonction définie sur \mathbb{R} par

$$\varphi_n(x) = \begin{cases} \frac{1}{a_n}(1-x^2)^n & \text{si } |x| \leq 1 \\ 0 & \text{si } |x| > 1 \end{cases} \quad \text{où } a_n = \int_{-1}^1 (1-x^2)^n dx$$

et $f_n : \mathbb{R} \rightarrow \mathbb{R}$ la fonction définie par

$$f_n(x) = f * \varphi_n(x) = \int_{-\infty}^{+\infty} f(x-t)\varphi_n(t) dt = \int_{-1}^1 f(x-t)\varphi_n(t) dt$$

* Propriétés de la suite φ_n .

1. $\forall t \in \mathbb{R}, \varphi_n(t) \geq 0$.
2. $\int_{\mathbb{R}} \varphi_n(t) dt = \int_{-1}^1 \varphi_n(t) dt = 1$ par définition de a_n .
3. $a_n \geq \int_{-1}^1 |t|(1-t^2)^n dt = \int_0^1 (1-t^2)^n 2t dt = 1/(n+1)$.

* La suite de fonctions (f_n) converge uniformément vers f sur \mathbb{R} .

Soit $\varepsilon > 0$. f est continue sur \mathbb{R} et nulle hors de l'intervalle compact J , donc est uniformément continue. Soit $\eta > 0$ tel que $|x - y| \leq \eta \Rightarrow |f(y) - f(x)| \leq \varepsilon/2$. On a alors

$$\begin{aligned} |f_n(x) - f(x)| &= \left| \int_{-1}^1 [f(x-t) - f(x)] \varphi_n(t) dt \right| \quad \text{en utilisant 2.} \\ &\leq \int_{-1}^{-\eta} |f(x-t) - f(x)| \varphi_n(t) dt + \int_{\eta}^1 |f(x-t) - f(x)| \varphi_n(t) dt \\ &\quad + \int_{-\eta}^{\eta} |f(x-t) - f(x)| \varphi_n(t) dt \end{aligned}$$

Pour $|t| \geq \eta$ on a $\varphi_n(t) \leq (1 - \eta^2)^n / a_n \leq \rho_n := (n + 1)(1 - \eta^2)^n$ donc chacune des deux premières intégrales est majorée par $2 \|f\|_\infty \rho_n$. La dernière est majorée par $(\varepsilon/2) \int_{-\eta}^{\eta} \varphi_n(t) dt \leq \varepsilon/2$.

On obtient finalement pour tout $x \in \mathbb{R}$ $|f_n(x) - f(x)| \leq 4\rho_n \|f\|_\infty + \varepsilon/2$. Comme ρ_n tend vers 0, il existe un entier N tel que $n \geq N \Rightarrow 4\rho_n \|f\|_\infty \leq \varepsilon/2$. Pour un tel n , on a $|f_n(x) - f(x)| \leq \varepsilon$ pour tout $x \in \mathbb{R}$.

* Il existe un polynôme P_n tel que $f_n(x) = P_n(x)$ pour $x \in J$.

Le changement de variables $u = x - t$ donne

$$f_n(x) = \int_{-\infty}^{+\infty} f(x-t)\varphi_n(t) dt = \int_{-\infty}^{\infty} f(u)\varphi_n(x-u) du = \int_{-1/2}^{1/2} f(u)\varphi_n(x-u) du$$

$|x| \leq 1/2$, et $|u| \leq 1/2 \Rightarrow -1 \leq x-u \leq 1$ donc $\varphi_n(x-u) = [1 - (x-u)^2]^n$ d'où l'expression de f_n pour $|x| \leq 1/2$:

$$f_n(x) = \frac{1}{a_n} \int_{-1/2}^{1/2} f(u) (1 - (x-u)^2)^n du$$

En développant par la formule du binôme sous l'intégrale, on voit que $f_n(x)$ s'écrit bien comme un polynôme en x , soit P_n .

* Conclusion : on a trouvé une suite de polynômes (P_n) convergeant uniformément vers f sur J .

2) Passons au cas général : soit $g : [a, b] \rightarrow \mathbb{R}$ continue. Soient c, d tels que $c < a < b < d$. Soit G le prolongement de g à \mathbb{R} tel que $G(x) = 0$ pour $x \leq c$ et pour $x \geq d$, et G affine sur chacun des intervalles $[c, a]$ et $[b, d]$. La fonction G est continue sur \mathbb{R} .

Soit $h : \mathbb{R} \rightarrow \mathbb{R}$ définie par $h(u) = c + (d-c)(u + \frac{1}{2})$. C'est une bijection affine de \mathbb{R} sur lui-même qui envoie J sur $[c, d]$. La fonction $F = G \circ h$ est continue sur \mathbb{R} , nulle hors de J . Pour tout $\varepsilon > 0$, il existe un polynôme P tel que $|P(t) - F(t)| \leq \varepsilon$ pour tout $t \in J$. Alors $Q := P \circ h^{-1}$ est un polynôme et $|Q(x) - G(x)| \leq \varepsilon$ pour tout x dans $[c, d]$. A fortiori, $|Q(x) - g(x)| \leq \varepsilon$ pour tout $x \in [a, b]$.

9

Dérivation

9.1 Définition et propriétés élémentaires

On désigne toujours par E un \mathbb{K} -espace vectoriel de dimension finie, $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} .

Si $I \subset \mathbb{R}$ et $a \in \mathbb{R}$ on note $I - a = \{x - a ; x \in I\}$

DEFINITION 9.1.1

Soit I un intervalle non trivial de \mathbb{R} , $f : I \rightarrow E$.

Soit $a \in I$. On dit que f est dérivable au point a si la fonction $T_a : I \setminus \{a\} \rightarrow E$ définie par $T_a(x) = \frac{f(x) - f(a)}{x - a}$ a une limite quand x tend vers a . Cette limite si elle existe est unique et sera notée $f'(a)$.

f est dérivable sur I si f est dérivable en chaque point de I ; dans ce cas l'application $x \rightarrow f'(x)$ est appelée fonction dérivée de f .

On dit que f est de classe C^1 sur I si elle est dérivable sur I et si sa dérivée f' est continue sur I .

♠ Les propriétés suivantes se vérifient facilement :

1. Si (f_1, \dots, f_n) sont les applications coordonnées de f dans la base $\mathcal{B} = (e_1, \dots, e_n)$, f est dérivable en a (resp. sur I) ssi chacune des applications f_i est dérivable en a (resp. sur I) et on a alors $f'(a) = f'_1(a)e_1 + \dots + f'_n(a)e_n$.
2. f est dérivable en a ssi la fonction f admet en a un développement limité d'ordre 1, i.e. si il existe un élément $e \in E$ et une fonction $\varepsilon : I - a \rightarrow E$ vérifiant

$$f(x) = f(a) + (x - a)e + (x - a)\varepsilon(x - a) \quad \text{et} \quad \lim_{t \rightarrow 0} \varepsilon(t) = 0$$

On a alors $f'(a) = e$.

3. Toute fonction f dérivable en a est continue en a .

Comme pour les fonctions numériques, on définit par récurrence la notion de fonctions k fois dérivable, de classe C^k pour $k \geq 2$: f est k fois dérivable (resp. de classe C^k) si f est dérivable et si sa dérivée f' est $k - 1$ fois dérivable (resp. de classe C^{k-1}). f est de classe C^∞ si elle est de classe C^k pour tout k .

Les ensembles $D^k(J, E)$, $C^k(J, E)$, et $C^\infty(J, E)$ des fonctions k fois dérivables (resp. de classe C^k , C^∞) sont des \mathbb{K} -espaces vectoriels. La dérivation est une application linéaire de $C^k(J, E)$ dans $C^{k-1}(J, E)$ pour $k \geq 1$, de $D^k(J, E)$ dans $D^{k-1}(J, E)$ pour $k \geq 2$, de $D^1(J, E)$ dans E^J et de $C^\infty(J, E)$ dans lui même.

9.2 Cas des fonctions numériques (rappels)

On rappelle ci dessous quelques propriétés particulièrement importantes des fonctions réelles d'une variable réelle.

1. Soit I un intervalle de \mathbb{R} et $f : I \rightarrow \mathbb{R}$ dérivable. Si f possède un extremum local atteint en un point c **intérieur à I** , $f'(c) = 0$.

En effet, pour x assez voisin de c , le taux d'accroissement entre x et c a un certain signe à gauche de c et le signe opposé à droite. Sa limite est donc nulle.

- Théorème de Rolle** Soit $f : [a, b] \rightarrow \mathbb{R}$ continue sur $[a, b]$, dérivable sur $]a, b[$ et telle que $f(a) = f(b)$. Il existe $c \in]a, b[$ tel que $f'(c) = 0$.
En effet, f continue sur le compact $[a, b]$ atteint son minimum et son maximum. Si f est constante, tout point c convient ; sinon le maximum et le minimum sont atteints en des points distincts. Puisque $f(a) = f(b)$ l'un de ces points est intérieur à $]a, b[$. D'où le résultat en utilisant 1.
- Théorème des accroissements finis** Soit $f : [a, b] \rightarrow \mathbb{R}$ continue sur $[a, b]$, dérivable sur $]a, b[$. Il existe $c \in]a, b[$ tel que $f(b) - f(a) = (b - a)f'(c)$.
Il suffit, pour le voir, d'appliquer le théorème de Rolle à la fonction $x \rightarrow g(x) = f(x) - \frac{f(b)-f(a)}{b-a}x$.
- Si $f : I \rightarrow \mathbb{R}$ est dérivable sur l'intervalle I et si $\forall x \in I f'(x) = 0$, f est constante sur I .
- Si $f : I \rightarrow \mathbb{R}$ est dérivable sur l'intervalle I et si $\forall x \in I f'(x) \geq 0$ (resp. $f'(x) \leq 0$), f est croissante (resp. décroissante) sur I .
- Inégalité des accroissements finis** Soit $f : [a, b] \rightarrow \mathbb{R}$ continue sur $[a, b]$, dérivable sur $]a, b[$. Si il existe $k \geq 0$ tel que $\forall x \in]a, b[, |f'(x)| \leq k$, on a $|f(b) - f(a)| \leq k(b - a)$.
Les trois dernières propriétés résultent immédiatement du théorème des accroissements finis.

PROPOSITION 9.2.1

Soit E un \mathbb{K} -espace vectoriel de dimension finie, I un intervalle de \mathbb{R} $f : I \rightarrow E$ dérivable sur I telle que $f'(x) = 0$ pour tout $x \in I$. f est constante sur I .

preuve

On fixe une base de E . Si f' est identiquement nulle, il en est de même de chacune des applications coordonnées de f ; ces applications sont donc constantes, donc f aussi.

9.3 Opérations

9.3.1 Formule de Leibnitz

THEOREME 9.3.1

Soient E, F , et G trois \mathbb{K} -espaces vectoriels de dimension finie et $B : F \times G \rightarrow E$ une application bilinéaire. Soient I un intervalle non trivial de \mathbb{R} , $f : I \rightarrow F$, $g : I \rightarrow G$ deux applications et $\Phi : I \rightarrow E$ définie par $\Phi(x) = B(f(x), g(x))$.

- Si f et g sont dérivables en $x \in I$, Φ aussi et $\Phi'(x) = B(f'(x), g(x)) + B(f(x), g'(x))$.
- Si f et g sont de classe C^n sur I , Φ aussi et

$$\Phi^{(n)}(x) = \sum_{k=0}^{k=n} C_n^k B(f^{(k)}(x), g^{(n-k)}(x))$$

où par convention $f^{(0)} = f$.

preuve

- Par hypothèse, pour h assez petit pour que $x + h \in I$, on a

$$f(x + h) = f(x) + hu(h), \quad g(x + h) = g(x) + hv(h) \quad \text{avec} \quad \lim_{h \rightarrow 0} u(h) = f'(x) \quad \text{et} \quad \lim_{h \rightarrow 0} v(h) = g'(x)$$

On en déduit

$$\begin{aligned} \frac{\Phi(x + h) - \Phi(x)}{h} &= \frac{1}{h} (B(f(x) + hu(h), g(x) + hv(h)) - B(f(x), g(x))) \\ &= B(f(x), v(h)) + B(u(h), g(x)) + hB(u(h), v(h)) \end{aligned}$$

Les espaces F et G étant de dimension finie, B est continue.

Donc $B(f(x), v(h)) \xrightarrow{h \rightarrow 0} B(f(x), g'(x))$ $B(u(h), g(x)) \xrightarrow{h \rightarrow 0} B(f'(x), g(x))$ $hB(u(h), v(h)) \xrightarrow{h \rightarrow 0} 0$
d'où la conclusion.

2. On démontre par récurrence sur n la propriété

(HR $_n$) Si f et g sont de classe C^n sur I , Φ aussi et $\Phi^{(n)}(x) = \sum_{k=0}^{k=n} C_n^k B(f^{(k)}(x), g^{(n-k)}(x))$

La première partie montre (HR $_1$) car si f et g sont C^1 , la continuité de B implique la continuité de $\Phi' = B(f', g) + B(f, g')$. Soit $n \geq 1$ et montrons que (HR $_n$) \Rightarrow (HR $_{n+1}$). Soient donc f et g de classe C^{n+1} . Les fonctions f' et g sont de classe C^n , donc d'après l'hypothèse de récurrence, il en est de même de $B(f', g)$. De même pour $B(f, g')$. Donc Φ' est C^n et Φ est C^{n+1} . Enfin, on a $\Phi' = B(f', g) + B(f, g')$. Toujours d'après l'hypothèse de récurrence, il vient

$$\begin{aligned} \Phi^{(n+1)}(x) &= \sum_{k=0}^{k=n} C_n^k B(f^{(k+1)}(x), g^{(n-k)}(x)) + \sum_{k=0}^{k=n} C_n^k B(f^{(k)}(x), g^{(n+1-k)}(x)) \\ &= \sum_{m=1}^{m=n+1} C_n^{m-1} B(f^{(m)}(x), g^{(n-m+1)}(x)) + \sum_{m=0}^{m=n} C_n^m B(f^{(m)}(x), g^{(n+1-m)}(x)) \\ &= B(f(x), g^{(n+1)}(x)) + \sum_{m=1}^{m=n} (C_n^{m-1} + C_n^m) B(f^{(m)}(x), g^{(n-m+1)}(x)) + B(f^{(n+1)}(x), g(x)) \\ &= \sum_{m=0}^{m=n+1} C_{n+1}^m B(f^{(m)}(x), g^{(n+1-m)}(x)) \end{aligned}$$

ce qui achève la preuve.

Le point 1 du théorème se généralise bien entendu à des applications k linéaires.

APPLICATIONS

Ce théorème a de très nombreuses applications. En voici quelques unes.

1. En prenant pour B l'application $\mathbb{K} \times E \rightarrow E, (t, v) \rightarrow tv$ on obtient :

Soient $u : I \rightarrow E$ et $\lambda : I \rightarrow \mathbb{K}$ de classe C^1 . Alors la fonction $\lambda \cdot u : x \rightarrow \lambda(x)u(x)$ est C^1 et

$$(\lambda \cdot u)'(x) = \lambda'(x)u(x) + \lambda(x)u'(x)$$

2. Soit $(E, <, >)$ un espace euclidien et $B : E \times E \rightarrow E$ le produit scalaire. On obtient :

Soient $u, v : I \rightarrow E$ de classe C^1 . La fonction $\varphi : x \rightarrow \varphi(x) = \langle u(x), v(x) \rangle$ est de classe C^1 et

$$\varphi'(x) = \langle u'(x), v(x) \rangle + \langle u(x), v'(x) \rangle$$

On en déduit en particulier que si $u : I \rightarrow E$ est de classe C^1 et vérifie $\|u(x)\| = Cte$ (la norme étant la norme euclidienne), alors pour tout x les vecteurs $u(x)$ et $u'(x)$ sont orthogonaux.

3. Soit $(E, <, >)$ un espace euclidien de dimension 3 orienté et B le produit vectoriel. On obtient :

Soient $u, v : I \rightarrow E$ de classe C^1 . La fonction $u \wedge v : x \rightarrow u(x) \wedge v(x)$ est de classe C^1 et

$$(u \wedge v)'(x) = u'(x) \wedge v(x) + u(x) \wedge v'(x) \quad (\text{attention à l'ordre!})$$

4. Soit B l'application $M_n(\mathbb{K}) \times \mathbb{K}^n \rightarrow \mathbb{K}^n \quad (M, X) \rightarrow MX$. On obtient :

Soient $A : I \rightarrow M_n(\mathbb{K})$ et $V : I \rightarrow \mathbb{K}^n$ de classe C^1 . L'application $F : x \rightarrow A(x)V(x)$ est de classe C^1 et

$$F'(x) = A'(x)V(x) + A(x)V'(x)$$

On a des résultats analogues pour les divers produits matriciels possibles, et aussi en remplaçant \mathbb{K}^n par un \mathbb{K} -espace vectoriel de dimension finie E et $M_n(\mathbb{K})$ par $L(E)$.

5. Soit $M : I \rightarrow M_n(\mathbb{K})$ une application de classe C^1 . Notons $C_1(x), \dots, C_n(x)$ les colonnes de la matrice $M(x)$. C_1, \dots, C_n sont des applications de classe C^1 de I dans \mathbb{K}^n . La fonction $x \rightarrow D(x) = \det M(x)$ est de classe C^1 et

$$D'(x) = \det(C'_1(x), C_2(x) + \dots, C_n(x)) + \det(C_1(x), C'_2(x), \dots, C_n(x)) + \dots + \det(C_1(x), C_2(x), \dots, C'_n(x))$$

Il suffit d'appliquer le théorème (généralisé aux applications multilinéaires) à l'application n -linéaire $\det : (\mathbb{K}^n)^n \rightarrow \mathbb{K}$. On a un énoncé analogue en remplaçant les colonnes de la matrice par ses lignes.

9.3.2 Fonctions composées

THEOREME 9.3.2

Soient I et J deux intervalles non triviaux de \mathbb{R} , $u : I \rightarrow J$ et $f : J \rightarrow E$.

1. Si u est dérivable en $a \in I$ et f dérivable en $b = u(a) \in J$, $g = f \circ u$ est dérivable en a et $g'(a) = u'(a)f'(u(a))$.
2. Si f et g sont de classes C^n ($n \in \mathbb{N}^*$ ou $n = \infty$) il en est de même de $f \circ u$.

preuve

1. Soit $\theta : J \rightarrow E$ définie par $\theta(y) = \frac{f(y)-f(b)}{y-b}$ pour $y \neq b$ et $\theta(b) = f'(b)$. La dérivabilité de f en b implique la continuité de θ en b . On a alors, pour tout $x \in I$, $x \neq a$

$$\frac{g(x) - g(a)}{x - a} = \frac{u(x) - u(a)}{x - a} \theta(u(x))$$

En effet, l'égalité est triviale si $u(x) \neq u(a)$ et si $u(x) = u(a)$ les deux membres sont nuls. On obtient le résultat en prenant la limite quand x tend vers a .

2. Le résultat se prouve par récurrence sur $n \in \mathbb{N}$. Si u et f sont C^1 , il en est de même de g et $g' = u' \cdot (f' \circ u)$. Supposons prouvé que si u et f sont C^n alors g est C^n . Soient u et f de classe C^{n+1} . Alors f' est C^n , u aussi, donc par hypothèse de récurrence $f' \circ u$ aussi. Ensuite, u' est C^n , donc d'après le théorème de Leibnitz, le produit $u' \cdot (f' \circ u)$ aussi ; par conséquent g' est C^n et g est de classe C^{n+1} . ■

9.4 Inégalité des accroissements finis et conséquences

THEOREME 9.4.1

Soit E un \mathbb{K} -espace vectoriel de dimension finie, $\| \cdot \|$ une norme sur E . Soient $[a, b]$ un intervalle compact de \mathbb{R} , $f : [a, b] \rightarrow E$ continue sur $[a, b]$ et dérivable sur $]a, b[$. On suppose qu'il existe un réel positif k tel que, pour tout $x \in]a, b[$ on ait $\|f'(x)\| \leq k$. Alors $\|f(b) - f(a)\| \leq k(b - a)$.

preuve

On va prouver que, pour tout $\varepsilon > 0$ on a $\|f(b) - f(a)\| \leq (k + \varepsilon)(b - a) + \varepsilon$.

Le résultat en découlera en faisant tendre ε vers 0.

Soit $\varepsilon > 0$. Définissons $I_\varepsilon = \{x \in [a, b] \mid \|f(x) - f(a)\| \leq (k + \varepsilon)(x - a) + \varepsilon\}$.

- Soit φ définie par $\varphi(x) = \|f(x) - f(a)\| - (k + \varepsilon)(x - a) - \varepsilon$. I_ε est l'image réciproque du fermé $] -\infty, 0]$ par la fonction continue φ . C'est donc un fermé.
- On a $\varphi(a) = -\varepsilon < 0$. Par continuité de φ , il existe a' , $a < a' < b$ tel que $\varphi(x) < 0$ pour $a \leq x < a'$. Donc $[a, a' [\subset I_\varepsilon$
- I_ε est non vide et majoré. Soit $x_0 = \sup I_\varepsilon$. Comme I_ε est fermé, $x_0 \in I_\varepsilon$.
- Le théorème sera prouvé si on montre que $x_0 = b$. Supposons le contraire. On a $a < a' \leq x_0 < b$; par conséquent f est dérivable en x_0 . Il existe donc un $\eta > 0$ tel que $x_0 + \eta < b$ et tel que pour $x \in]x_0, x_0 + \eta[$ on ait

$$\|f(x) - f(x_0) - (x - x_0)f'(x_0)\| \leq \varepsilon(x - x_0) \quad \text{donc} \quad \|f(x) - f(x_0)\| \leq (\|f'(x_0)\| + \varepsilon)(x - x_0) \leq (k + \varepsilon)(x - x_0)$$

Comme $x_0 \in I_\varepsilon$ on a

$$\|f(x_0) - f(a)\| \leq (k + \varepsilon)(x_0 - a) + \varepsilon$$

donc,

$$\|f(x) - f(a)\| \leq \|f(x) - f(x_0)\| + \|f(x_0) - f(a)\| \leq (k + \varepsilon)(x - a) + \varepsilon$$

ce qui prouve que $x \in I_\varepsilon$. Or $x > x_0 = \sup I_\varepsilon$. L'hypothèse $x_0 \neq b$ conduit à une contradiction, donc $x_0 = b$. ■

On peut retrouver ainsi le résultat déjà prouvé (prop. 9.2.1) :

COROLLAIRE 9.4.1

Soit I un intervalle non trivial de \mathbb{R} et $f : I \rightarrow E$ une fonction. f est constante ssi f est dérivable de dérivée nulle.

preuve

Il est clair que si f est constante, elle est dérivable de dérivée nulle. Inversement, supposons $f'(x) = 0$ pour tout $x \in I$. Soit $a \in I$ fixé. Pour $x \in I$ on applique l'inégalité des accroissements finis sur l'intervalle d'extrémités a et x avec $k = 0$ pour conclure que $f(x) = f(a)$.

THEOREME 9.4.2 (Prolongement des fonctions de classe C^1)

Soient E un \mathbb{K} -espace vectoriel de dimension finie, $f :]a, b[\rightarrow E$ continue, dérivable. Si $\|f'(x)\|$ est borné pour $x \in]a, b[$, f admet une limite quand x tend vers a .

preuve

Soit (x_n) une suite de points de $]a, b[$ convergeant vers a . Si K est un majorant de la norme de f' on a d'après l'inégalité des accroissements finis $\|f(x_n) - f(x_{n+p})\| \leq K|x_n - x_{n+p}|$ ce qui prouve que la suite $(f(x_n))$ est de Cauchy. E étant complet, elle converge, d'où le résultat.(1.5.2).

Remarque

Dans le cas $E = \mathbb{R}$ le résultat peut se prouver de manière plus élémentaire ; en effet, dans ce cas la fonction $g(x) = f(x) - Kx$ est décroissante sur $]a, b[$ puisque sa dérivée est négative. D'autre part elle est bornée car, par exemple, $|g(x)| \leq |g(x) - g(b)| + |g(b)| \leq 2K(b - a) + |g(b)|$. Donc elle admet une limite finie quand x tend vers a .

COROLLAIRE 9.4.2

Soit $f :]a, b[\rightarrow E$ de classe C^1 . Si $L := \lim_{x \rightarrow a} f'(x)$ existe, f se prolonge par continuité en a , le prolongement F est de classe C^1 sur $[a, b]$ et $F'(a) = L$.

preuve

f' est continue sur $]a, b[$ et a une limite en a donc est bornée sur $]a, b[$. Donc f se prolonge par continuité en a . Soit F le prolongement.

Soit $\varepsilon > 0$. Il s'agit de montrer qu'il existe $\eta > 0$ tel que $a \leq x \leq a + \eta \Rightarrow \|F(x) - F(a) - L(x - a)\| \leq \varepsilon(x - a)$.

Considérons la fonction φ définie sur $]a, b[$ par $\varphi(x) = F(x) - F(a) - L(x - a)$. Cette fonction est continue sur $[a, b]$, dérivable sur $]a, b[$ et sa dérivée est $\varphi'(x) = f'(x) - L$. Par hypothèse, il existe $\eta > 0$ tel que $a < x \leq a + \eta \Rightarrow \|f'(x) - L\| \leq \varepsilon$. On a donc, pour $a \leq x \leq a + \eta$, $\|\varphi(x) - \varphi(a)\| \leq \varepsilon(x - a)$ soit $\|F(x) - F(a) - L(x - a)\| \leq \varepsilon(x - a)$.

Par une récurrence facile, on en déduit :

COROLLAIRE 9.4.3

Soit $f :]a, b[\rightarrow E$ de classe C^n . Si $\lim_{x \rightarrow a} f^{(n)}(x)$ existe, f se prolonge en a en une fonction de classe C^n sur $[a, b]$.

10

Intégration des fonctions continues par morceaux sur un intervalle compact

Soit E un espace vectoriel de dimension finie $n \geq 1$ sur le corps $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} . On munit E d'une norme $\| \cdot \|$. Soit d'autre part un intervalle compact $J = [a, b]$ de \mathbb{R} où $a < b$. Ce chapitre se propose de présenter la théorie de l'intégration des fonctions continues par morceaux sur J à valeurs dans E . Ceci englobe bien entendu le cas des fonctions numériques (en prenant $E = \mathbb{R}$ ou \mathbb{C} muni de la valeur absolue comme norme).

On utilisera les notations du chapitre 8. En particulier, $\mathcal{E}(J, E)$, $C_M(J, E)$, et $C^0(J, E)$ désignent respectivement le \mathbb{K} -espace vectoriel des fonctions en escalier, continues par morceaux, continues sur J , à valeurs dans E . On omettra J et E si il n'y a pas d'ambiguïtés.

10.1 Intégration des fonctions en escalier

Soit $\varphi \in \mathcal{E}(J, E)$ et $\sigma = \{x_0 = a, x_1, \dots, x_n = b\}$ une subdivision adaptée à φ . Soit, pour $0 \leq i \leq n-1$, c_i la valeur de φ sur $]x_i, x_{i+1}[$. Le nombre $I(\varphi, \sigma) := \sum_{0 \leq i \leq n-1} (x_{i+1} - x_i)c_i$ ne dépend pas de σ . En effet, si $t \in J$ on vérifie facilement que $I(\varphi, \sigma \cup \{t\}) = I(\varphi, \sigma)$, donc par récurrence que $I(\varphi, \sigma \cup \{t_1, \dots, t_p\}) = I(\varphi, \sigma)$, si t_1, \dots, t_p sont des points de J . Alors, si σ et σ' sont deux subdivisions adaptées à φ on a $I(\varphi, \sigma) = I(\varphi, \sigma \cup \sigma') = I(\varphi, \sigma')$. Cet élément de E est appelé intégrale de φ sur $[a, b]$ et est noté $\int_J \varphi$ ou $\int_a^b \varphi$ ou enfin $\int_a^b \varphi(x) dx$

♠ Les propriétés suivantes se vérifient sans difficultés :

1. L'application $I : \mathcal{E}(J, E) \rightarrow E : \varphi \rightarrow \int_J \varphi$ est une application linéaire de \mathcal{E} dans E : si $\varphi, \psi \in \mathcal{E}$ et $\lambda \in \mathbb{K}$ on a

$$\int_J (\lambda\varphi + \psi) = \lambda \int_J \varphi + \int_J \psi$$

2. Pour toute $\varphi \in \mathcal{E}(J, E)$ la fonction $t \rightarrow \|\varphi(t)\|$ appartient à $\mathcal{E}(J, \mathbb{R})$ et on a

$$\left\| \int_a^b \varphi(t) dt \right\| \leq \int_a^b \|\varphi(t)\| dt \leq (b-a) \|\varphi\|_\infty$$

Donc l'application linéaire I est continue lorsqu'on munit \mathcal{E} de la norme de la convergence uniforme.

3. Si $E = \mathbb{R}$, I est une forme linéaire positive, i.e. si φ est positive ($\forall x \in J, \varphi(x) \geq 0$), on a $\int_J \varphi \geq 0$.
4. (Relation de Chasles) Soit $\varphi \in \mathcal{E}([a, b], E)$ et $c \in]a, b[$; la restriction de φ à chacun des intervalles $[a, c]$ et $[c, b]$ est en escalier. En notant $\int_a^c \varphi = \int_a^c \varphi|_{[a, c]}$ on a

$$\int_a^b \varphi = \int_a^c \varphi + \int_c^b \varphi$$

10.2 Intégration des fonctions continues par morceaux

10.2.1 Construction

Soit $f \in C_M$; il existe une suite (φ_n) de fonctions en escalier qui converge uniformément vers f . La suite $(I(\varphi_n))$ est de Cauchy dans E car $\|I(\varphi_n) - I(\varphi_{n+p})\| \leq (b-a)\|\varphi_{n+p} - \varphi_n\|_\infty$; E est complet (car de dimension finie), donc cette suite converge vers un $\lambda \in E$. On vérifie ensuite que λ dépend seulement de f et non de la suite (φ_n) choisie. On note (provisoirement) $\hat{I}(f) := \lambda$. On vérifie ensuite que l'application \hat{I} est linéaire et qu'elle prolonge I . De l'inégalité $\|I(\varphi_n)\| \leq (b-a)\|\varphi_n\|_\infty$ valable pour tout n on déduit $\|\hat{I}(f)\| \leq (b-a)\|f\|_\infty$. L'application linéaire \hat{I} est donc continue. L'application \hat{I} prolongeant I on notera $\hat{I} = I$ et pour $f \in C_M$, $I(f) = \int_J f = \int_a^b f = \int_a^b f(x) dx$.

Remarque

La construction ci dessus est un cas particulier du théorème 5.4.3. On applique ce théorème avec $G = C_M$ muni de la norme de la convergence uniforme, $F = \mathcal{E}$ qui est bien dense dans C_M et $u = I$.

THEOREME 10.2.1

L'intégrale possède les propriétés suivantes:

1. L'application $I : f \rightarrow \int_J f$ est une application linéaire de $C_M(J, E)$ dans E : si $f, g \in C_M$ et $\lambda \in \mathbb{K}$ on a

$$\int_J (\lambda f + g) = \lambda \int_J f + \int_J g$$

2. Pour toute $f \in C_M(J, E)$ la fonction $t \rightarrow \|f(t)\|$ appartient à $C_M(J, \mathbb{R})$ et on a

$$\left\| \int_a^b f(t) dt \right\| \leq \int_a^b \|f(t)\| dt \leq (b-a)\|f\|_\infty$$

L'application linéaire I est continue lorsqu'on munit C_M de la norme de la convergence uniforme.

3. (Relation de Chasles) Soit $f \in C_M([a, b], E)$ et $c \in]a, b[$; la restriction de f à chacun des intervalles $[a, c]$ et $[c, b]$ est continue par morceaux. En notant $\int_a^c f = \int_a^c f|_{[a, c]}$ on a

$$\int_a^b f = \int_a^c f + \int_c^b f$$

♠ La preuve se fait à partir des propriétés (1), (2) et (4) du paragraphe précédent par passage à la limite.

10.3 Cas des fonctions numériques

Notons que l'ensemble $C_M(J, \mathbb{K})$ muni des opérations usuelles est une \mathbb{K} algèbre dont $C^0(J, \mathbb{K})$ est une sous algèbre. Si $f \in C_M(J, \mathbb{K})$ (resp. $C^0(J, \mathbb{K})$), il en est de même de la fonction $|f|$.

10.3.1 Positivité

THEOREME 10.3.1

L'application $I : C_M(J, \mathbb{R}) \rightarrow \mathbb{R}$ est une forme linéaire positive, i.e. si f est positive ($\forall x \in J, f(x) \geq 0$), on a $\int_J f \geq 0$.

preuve

Soit (φ_n) une suite de fonctions en escalier convergeant uniformément vers f ; posons $\psi_n = \sup(\varphi_n, 0)$, i.e. $\psi_n(x) = \varphi_n(x)$ si $\varphi_n(x) \geq 0$ et $\psi_n(x) = 0$ sinon. On a pour tout $x \in J$ $|f(x) - \psi_n(x)| \leq |f(x) - \varphi_n(x)|$ car $f(x) \geq 0$ donc la suite (ψ_n) converge aussi uniformément vers f . On a $\int_J \psi_n \geq 0$ donc $\int_J f = \lim_{n \rightarrow \infty} \int_J \psi_n \geq 0$.

COROLLAIRE 10.3.1

Soient $f, g \in C_M(E, \mathbb{R})$; si $\forall x \in J f(x) \leq g(x)$ on a $\int_J f \leq \int_J g$.

THEOREME 10.3.2

Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction positive et continue par morceaux telle que $\int_a^b f(t) dt = 0$; si $x \in [a, b]$ est un point où f est continue, on a $f(x) = 0$.

preuve

En effet, si on suppose f continue en x et $f(x) > 0$ il existe un intervalle I contenu dans $[a, b]$ et contenant x tel que $f(t) \geq f(x)/2$ pour tout $t \in I$; soit $\phi : [a, b] \rightarrow \mathbb{R}$ la fonction en escalier telle que $\phi(t) = f(x)/2$ pour $t \in I$ et $\phi(t) = 0$ si $t \notin I$. On a donc $\forall t \in [a, b]$ $\phi(t) \leq f(t)$ ce qui implique

$$0 < \frac{1}{2} f(x) \text{long}(I) = \int_a^b \phi(t) dt \leq \int_a^b f(t) dt$$

d'où la contradiction.

COROLLAIRE 10.3.2

Soit $f : [a, b] \rightarrow \mathbb{R}$ une fonction positive, **continue** et non identiquement nulle. On a $\int_a^b f(x) dx > 0$.

COROLLAIRE 10.3.3

Soit E un \mathbb{K} espace vectoriel normé de dimension finie. L'application $f \rightarrow \|f\|_1 := \int_a^b \|f(x)\| dx$ est une norme sur l'espace vectoriel $C^0([a, b], E)$.

Cette norme est appelée norme de la convergence en moyenne. On a pour toute $f \in C^0([a, b], E)$ $\|f\|_1 \leq (b-a) \|f\|_\infty$.

Exercice

Soit $f_n : [0, 1] \rightarrow \mathbb{R}$ vérifiant $f_n(0) = f_n(2/n) = f_n(1) = 0$, $f_n(1/n) = \sqrt{n}$ et affine sur chacun des intervalles $[0, 1/n]$, $[1/n, 2/n]$ et $[2/n, 1]$. Montrer que la suite (f_n) converge en moyenne vers la fonction nulle. En déduire que les normes $\|\cdot\|_1$ et $\|\cdot\|_\infty$ ne sont pas équivalentes.

10.3.2 Inégalité de Cauchy Schwarz

Soient $f, g : [a, b] \rightarrow \mathbb{C}$ continues par morceaux et \bar{f} l'application $t \rightarrow \overline{f(t)}$ (complexe conjugué de $f(t)$).

♠ L'application $(f, g) \rightarrow \int_a^b \bar{f}(t)g(t) dt$ est une forme hermitienne positive sur $C_M([a, b], E)$, dont la restriction à $C^0([a, b], E)$ est définie positive d'après le théorème 10.3.2. On en déduit :

THEOREME 10.3.3

1. (Inégalité de Cauchy Schwarz) Pour $f, g \in C_M([a, b], \mathbb{C})$

$$\left| \int_a^b \bar{f}(t)g(t) dt \right|^2 \leq \int_a^b |f(t)|^2 dt \cdot \int_a^b |g(t)|^2 dt$$

2. L'application $(f, g) \rightarrow \int_a^b \bar{f}(t)g(t) dt$ est un produit scalaire complexe sur $C^0([a, b], \mathbb{C})$ dont la norme associée est donnée par $\|f\|_2 := \sqrt{\int_a^b |f(t)|^2 dt}$.

3. De même, l'application $(f, g) \rightarrow \int_a^b f(t)g(t) dt$ est un produit scalaire réel sur $C^0([a, b], \mathbb{R})$ dont la norme associée est donnée par $\|f\|_2 := \sqrt{\int_a^b f(t)^2 dt}$.

La norme $\|\cdot\|_2$ est appelée norme de la convergence en moyenne quadratique. Elle vérifie $\|\cdot\|_2 \leq \sqrt{b-a} \|\cdot\|_\infty$.

Une preuve de l'inégalité de Cauchy-Schwarz pour une forme hermitienne définie positive est donnée au chapitre 11. Elle est bien entendu indépendante de toute notion d'intégrale!

10.3.3 Formules de la moyenne

THEOREME 10.3.4

Soient $f, g : [a, b] \rightarrow \mathbb{R}$; on suppose que f est continue sur $[a, b]$ et que g est continue par morceaux et de signe constant sur $[a, b]$. Alors, il existe $c \in [a, b]$ tel que $\int_a^b f(t)g(t) dt = f(c) \int_a^b g(t) dt$.

preuve

On peut supposer g positive. La fonction f continue sur l'intervalle fermé borné $[a, b]$ est bornée. Soit $m = \inf_{a \leq t \leq b} f(t)$ et $M = \sup_{a \leq t \leq b} f(t)$.

Puisque $g(t) \geq 0$, on a, pour tout $t \in [a, b]$ $mg(t) \leq f(t)g(t) \leq Mg(t)$ donc $m \int_a^b g \leq \int_a^b fg \leq M \int_a^b g$. Si $\int_a^b g = 0$, il en est donc de même de $\int_a^b fg$ et tout point $c \in [a, b]$ convient ; sinon, $\int_a^b g > 0$ et l'existence du point c résulte du théorème des valeurs intermédiaires appliqué à la fonction continue f et de l'inégalité $m \leq \int_a^b fg / \int_a^b g \leq M$. ■

COROLLAIRE 10.3.4

Soit $f : [a, b] \rightarrow \mathbb{R}$ continue ; il existe $c \in [a, b]$ tel que $\frac{1}{b-a} \int_a^b f(t) dt = f(c)$

10.4 Sommes de Riemann

DEFINITION 10.4.1

Une subdivision pointée d'un intervalle fermé borné $[a, b]$ est un couple (σ, τ) où $\sigma = (x_0 = a, x_1, \dots, x_n = b)$ est une subdivision de $[a, b]$ et $\tau = (t_0, t_1, \dots, t_{n-1})$ un n -uplet de points vérifiant $\forall i \in \mathbb{N}_{n-1} = \{0, 1, \dots, n-1\}$, $x_i \leq t_i \leq x_{i+1}$. Le pas de la subdivision est $p(\sigma) = \max_{0 \leq i \leq n-1} (x_{i+1} - x_i)$.

Soit $f : [a, b] \rightarrow E$. La somme de Riemann de f associée à la subdivision pointée (σ, τ) est $S(f, \sigma, \tau) = \sum_{i=0}^{n-1} (x_{i+1} - x_i) f(t_i)$

THEOREME 10.4.1

Soit $f : [a, b] \rightarrow E$ continue par morceaux ; pour tout $\varepsilon > 0$ il existe un $\eta > 0$ tel que pour toute subdivision pointée (σ, τ) de pas inférieur à η on ait

$$\left| \int_a^b f(t) dt - S(f, \sigma, \tau) \right| \leq \varepsilon$$

preuve

Notons que pour σ, τ fixées l'application $f \rightarrow S(f, \sigma, \tau)$ est linéaire et vérifie $\|S(f, \sigma, \tau)\| \leq (b-a) \|f\|_\infty$.

- a) On commence par prouver le théorème pour la fonction caractéristique χ d'un intervalle K quelconque, d'extrémités α et β , contenu dans $[a, b]$. Soient $\varepsilon > 0$, $\eta < \min(\varepsilon/2, \beta - \alpha)$ et (σ, τ) une subdivision pointée de pas inférieur à η . Soient i et j tels que $\alpha \in [x_i, x_{i+1}]$ et $\beta \in [x_j, x_{j+1}]$. Puisque $\beta - \alpha > p(\sigma)$ on a $i < j$. Il vient

$$\int_a^b \chi - S(\chi, \sigma, \tau) = \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} (\chi(x) - \chi(t_k)) dx = \int_{x_i}^{x_{i+1}} (\chi(x) - \chi(t_i)) dx + \int_{x_j}^{x_{j+1}} (\chi(x) - \chi(t_j)) dx$$

donc

$$\left| \int_a^b \chi - S(\chi, \sigma, \tau) \right| \leq (x_{i+1} - x_i) + (x_{j+1} - x_j) \leq 2\eta \leq \varepsilon$$

- b) On en déduit le théorème pour une fonction en escalier $\varphi : [a, b] \rightarrow E$. Une telle fonction s'écrit $\varphi(t) = \sum_{1 \leq k \leq m} \chi_k(t) a_k$ où les χ_k sont des fonctions caractéristiques d'intervalle et les a_k des éléments non nuls de E . D'après le a), pour chaque k , il existe un $\eta_k > 0$ tel que $\left| \int_a^b \chi_k - S(\chi_k, \sigma, \tau) \right| \leq \frac{\varepsilon}{m \|a_k\|}$ pour toute subdivision pointée de pas inférieur à η_k . Si $\eta = \min_{1 \leq k \leq m} \eta_k$, on a pour toute subdivision pointée de pas inférieur à η $\left\| \int_a^b \varphi - S(\varphi, \sigma, \tau) \right\| \leq \varepsilon$.

- c) Fin de la preuve: soit f continue par morceaux sur $[a, b]$; soit $\varepsilon > 0$ et $\varepsilon' = \varepsilon/3(b-a)$. Il existe une fonction en escalier φ telle que $\|\varphi - f\|_\infty \leq \varepsilon'$. Pour cette fonction φ il existe d'après le b) un $\eta > 0$ tel que $\left\| \int_a^b \varphi - S(\varphi, \sigma, \tau) \right\| \leq \varepsilon/3$ pour toute subdivision pointée telle que $p(\sigma) < \eta$. Pour une telle subdivision on a alors

$$\begin{aligned} \left\| \int_a^b f - S(f, \sigma, \tau) \right\| &\leq \left\| \int_a^b (f - \varphi) \right\| + \left\| \int_a^b \varphi - S(\varphi, \sigma, \tau) \right\| + \|S(\varphi - f, \sigma, \tau)\| \\ &\leq (b-a)\varepsilon' + \varepsilon/3 + (b-a) \|f - \varphi\|_\infty \leq \varepsilon \quad \blacksquare \end{aligned}$$

Exercice

Donner une autre démonstration (plus courte) du théorème dans le cas d'une fonction f continue sur $[a, b]$ (Utiliser la continuité uniforme de f).

Exemple usuel

Pour un entier $n \geq 1$ on considère la subdivision régulière de $[a, b]$ de pas $(b-a)/n$: $x_k = a + k \frac{b-a}{n}$; on choisit $t_k = x_k$ pour tout k entre 0 et $n-1$. On obtient

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} f \left(a + k \frac{b-a}{n} \right) = \frac{1}{b-a} \int_a^b f$$

Si on choisit $t_k = x_{k+1}$, on obtient

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=1}^n f \left(a + k \frac{b-a}{n} \right) = \frac{1}{b-a} \int_a^b f$$

10.5 Intégration et primitives

10.5.1 Propriétés d'une intégrale fonction de sa borne supérieure

DEFINITION 10.5.1

Soit I un intervalle de \mathbb{R} et $f : I \rightarrow E$. On dira que f est localement continue par morceaux si la restriction de f à tout intervalle fermé borné $J \subset I$ est continue par morceaux.

NB Cette définition pratique n'est pas standard.

Soit $f : I \rightarrow E$ localement continue par morceaux sur I et $a, b \in I$. Si $a < b$ l'intégrale $\int_a^b f$ est donc définie. Si $b < a$ on pose, par convention $\int_a^b f = -\int_b^a f$ et aussi $\int_a^a f = 0$. Avec ces conventions, la relation de Chasles a lieu sous sa forme la plus générale : si $a, b, c \in I$, on a $\int_a^b f = \int_a^c f + \int_c^b f$.

THEOREME 10.5.1

Soit $f : I \rightarrow E$ localement continue par morceaux, $a \in I$. Soit $F : I \rightarrow E$ définie par $F(x) = \int_a^x f(t) dt$.

- 1) F est continue sur I .
- 2) Si f est continue au point $x_0 \in I$, F est dérivable en x_0 et $F'(x_0) = f(x_0)$.
- 3) Si f est continue sur I , F est de classe C^1 sur I .

preuve

Soit $x_0 \in I$; on choisit deux points c et d de I comme suit : si x_0 n'est pas une extrémité de I on prend $c < x_0 < d$; si x_0 est l'extrémité inférieure de I on prend $c = x_0$ et $x_0 < d$; enfin si x_0 est l'extrémité supérieure de I on prend $c < x_0$ et $d = x_0$. $J = [c, d]$ ainsi défini est un voisinage de x_0 dans I . Sur l'intervalle J , f est continue par morceaux, donc bornée ; il existe $M > 0$ tel que $\forall x \in J, \|f(x)\| \leq M$. Pour $x \in J$ on a $\|F(x) - F(x_0)\| = \left\| \int_{x_0}^x f(t) dt \right\| \leq \left| \int_{x_0}^x \|f(t)\| dt \right| \leq M|x - x_0|$. D'où la continuité de F en x_0 . Soit $\varepsilon > 0$; si f est continue en x_0 , on peut choisir c et d comme ci dessus de sorte que $\|f(t) - f(x_0)\| \leq \varepsilon$ pour $c \leq t \leq d$. Alors, si $x \in [c, d]$, on a

$$\left\| \frac{F(x) - F(x_0)}{x - x_0} - f(x_0) \right\| = \frac{1}{|x - x_0|} \left\| \int_{x_0}^x (f(t) - f(x_0)) dt \right\| \leq \frac{1}{|x - x_0|} \left| \int_{x_0}^x \|f(t) - f(x_0)\| dt \right| \leq \varepsilon$$

d'où le point 2). Le 3) en découle. ■

COROLLAIRE 10.5.1

Toute fonction continue sur un intervalle I à valeurs dans un espace vectoriel de dimension finie E admet des primitives.

Il suffit de choisir $a \in I$ et de prendre $F(x) = \int_a^x f(t) dt$.

COROLLAIRE 10.5.2 (Théorème fondamental du calcul intégral)

Si $f : I \rightarrow E$ est une fonction continue et si F est une primitive de f sur I , on a, pour $a, b \in I$, $\int_a^b f(x) dx = F(b) - F(a)$ quantité notée traditionnellement $F(t) \Big|_{t=a}^{t=b}$

preuve

Soit F une primitive quelconque de f et G définie par $G(x) = \int_a^x f(t) dt$; G est une primitive de f d'après le corollaire précédent. La différence $G - F$ a une dérivée nulle sur I , donc est constante sur I . On en déduit $F(b) - F(a) = G(b) - G(a) = \int_a^b f(t) dt$. ■

COROLLAIRE 10.5.3

Soient I un intervalle non trivial de \mathbb{R} , $a \in I$ et $f : I \rightarrow E$ de classe C^1 . On a, pour tout $x \in I$, $f(x) = f(a) + \int_a^x f'(u) du$.

10.5.2 Applications

COROLLAIRE 10.5.4 (Intégration par parties)

Soient E un \mathbb{K} espace vectoriel de dimension finie, $u : [a, b] \rightarrow \mathbb{K}$ et $v : [a, b] \rightarrow E$ de classe C^1 . On a

$$\int_a^b u'(t)v(t) dt = u(b)v(b) - u(a)v(a) - \int_a^b u(t)v'(t) dt$$

Car la fonction continue $u'v + uv'$ admet la fonction uv comme primitive. ■

Ce corollaire n'est qu'un cas particulier du résultat suivant plus général:

COROLLAIRE 10.5.5

Soient E , F , et G trois \mathbb{K} espaces vectoriels de dimension finie, $B : E \times F \rightarrow G$ une application bilinéaire, $u : [a, b] \rightarrow E$ et $v : [a, b] \rightarrow F$ de classe C^1 . On a

$$\int_a^b B(u'(t), v(t)) dt = B(u(b), v(b)) - B(u(a), v(a)) - \int_a^b B(u(t), v'(t)) dt$$

La preuve est la même: la fonction continue $B(u', v) + B(u, v')$ admet la fonction $B(u, v)$ comme primitive. ■

COROLLAIRE 10.5.6 (Changement de variables)

Soit $f : [a, b] \rightarrow E$ continue et $\varphi : [\alpha, \beta] \rightarrow [a, b]$ de classe C^1 . On a

$$\int_{\alpha}^{\beta} \varphi'(t) f(\varphi(t)) dt = \int_{\varphi(\alpha)}^{\varphi(\beta)} f(u) du$$

En effet, si F est une primitive de f sur $[a, b]$, $F \circ \varphi$ est une primitive de $\varphi' \cdot f \circ \varphi$ sur $[\alpha, \beta]$. On voit alors que les deux membres sont égaux à $F(\varphi(\beta)) - F(\varphi(\alpha))$. ■

THEOREME 10.5.2 (Inégalité des accroissements finis)

Soient E un \mathbb{K} espace vectoriel normé de dimension finie, $f : [a, b] \rightarrow E$ et $g : [a, b] \rightarrow \mathbb{R}$ de classe C^1 vérifiant $\forall t \in [a, b], \|f'(t)\| \leq g'(t)$. On a

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

preuve

C'est immédiat : $\|f(b) - f(a)\| = \left\| \int_a^b f'(t) dt \right\| \leq \int_a^b \|f'(t)\| dt \leq \int_a^b g'(t) dt = g(b) - g(a)$. ■

THEOREME 10.5.3 (Formule de Taylor, reste intégrale)

Soient E un \mathbb{K} espace vectoriel de dimension finie et $f : [a, b] \rightarrow E$ de classe C^{n+1} . On a

$$f(b) = f(a) + (b-a)f'(a) + \dots + \frac{1}{n!} \int_a^b (b-t)^n f^{(n+1)}(t) dt$$

preuve

Soit, pour $0 \leq p \leq n$, $I_p = \frac{1}{p!} \int_a^b (b-t)^p f^{(p+1)}(t) dt$. Il vient en intégrant par parties (ce qui est possible puisque f est de classe C^{n+1}) $I_p = -\frac{(b-a)^p}{p!} f^{(p)}(a) + I_{p-1}$. En ajoutant de $p=0$ à $p=n$ on obtient le résultat voulu en tenant compte de l'égalité $I_0 = f(b) - f(a)$. ■

COROLLAIRE 10.5.7 (Inégalité de Taylor Lagrange)

Soient E un \mathbb{K} espace vectoriel de dimension finie, $f : [a, b] \rightarrow E$ de classe C^{n+1} et $M > 0$ tel que $\forall t \in [a, b]$, $\|f^{(n+1)}(t)\| \leq M$.
On a

$$\left\| f(b) - f(a) - \sum_{k=1}^{k=n} \frac{(b-a)^k}{k!} f^{(k)}(a) \right\| \leq M \frac{(b-a)^{n+1}}{(n+1)!}$$

11

Espaces préhilbertiens

11.1 Produit scalaire

11.1.1 Définitions

DEFINITION 11.1.1

Soit E un \mathbb{R} -ev ; on appelle produit scalaire sur E une forme bilinéaire symétrique définie positive sur E , c'est à dire une application $\langle | \rangle : E \times E \rightarrow \mathbb{R}$ vérifiant :

1. Pour tout $x \in E$ l'application $y \rightarrow \langle x | y \rangle$ de E dans \mathbb{R} est linéaire.
2. $\forall x, y \in E, \langle x | y \rangle = \langle y | x \rangle$
3. $\forall x \in E, x \neq 0 \Rightarrow \langle x | x \rangle > 0$.

DEFINITION 11.1.2

Soit E un \mathbb{C} -ev ; on appelle produit scalaire sur E une forme sesquilinéaire à symétrie hermitienne définie positive sur E , c'est à dire une application $\langle | \rangle : E \times E \rightarrow \mathbb{C}$ vérifiant :

1. Pour tout $x \in E$ l'application $y \rightarrow \langle x | y \rangle$ de E dans \mathbb{C} est linéaire.
2. $\forall x, y \in E, \langle x | y \rangle = \overline{\langle y | x \rangle}$
3. $\forall x \in E, x \neq 0 \Rightarrow \langle x | x \rangle > 0$.

Dans le cas d'un espace complexe, en utilisant (1) et (2), on voit que, pour y fixé dans E , l'application $x \rightarrow \langle x | y \rangle$ est semi-linéaire, c'est à dire qu'elle vérifie :

$$\forall x_1, x_2 \in E, \langle x_1 + x_2 | y \rangle = \langle x_1 | y \rangle + \langle x_2 | y \rangle$$
$$\forall x \in E, \forall \lambda \in \mathbb{C}, \langle \lambda x | y \rangle = \bar{\lambda} \langle x | y \rangle$$

Notons aussi que (2) implique que $\langle x | x \rangle$ est réel.

DEFINITION 11.1.3

On appelle espace préhilbertien tout couple $(E, \langle | \rangle)$ où E est un espace vectoriel sur \mathbb{R} ou \mathbb{C} et $\langle | \rangle$ un produit scalaire sur E . Si $\mathbb{K} = \mathbb{R}$ l'espace est dit préhilbertien réel, si $\mathbb{K} = \mathbb{C}$, préhilbertien complexe.

EXEMPLE 11.1.1

1. Un espace préhilbertien réel (resp. complexe) de dimension finie est un espace euclidien (resp. hermitien).
2. Soit $E = C^0([a, b], \mathbb{C})$ le \mathbb{C} -espace vectoriel des fonctions continues sur $[a, b]$ à valeurs dans \mathbb{C} . L'application $(f, g) \rightarrow \int_a^b \overline{f(t)}g(t) dt$ est un produit scalaire sur E .

3. Cet exemple est essentiel dans l'étude des séries de Fourier. Soit E l'ensemble des fonctions $f : \mathbb{R} \rightarrow \mathbb{C}$, 2π -périodiques, localement continues par morceaux (i.e. dont la restriction à tout intervalle compact de \mathbb{R} est continue par morceaux), et vérifiant $\forall x \in \mathbb{R}, f(x) = \frac{1}{2} \left(\lim_{\substack{y \rightarrow x \\ y > x}} f(y) + \lim_{\substack{y \rightarrow x \\ y < x}} f(y) \right)$.

Pour $f, g \in E$ on pose $\langle f | g \rangle = \frac{1}{2\pi} \int_0^{2\pi} \overline{f(t)} g(t) dt$. $(E, \langle | \rangle)$ est un espace préhilbertien complexe.

Si E est un espace préhilbertien et F un sev de E , la restriction du produit scalaire à $F \times F$ munit F d'une structure d'espace préhilbertien.

Dans toute la suite, \mathbb{K} désignera soit \mathbb{R} soit \mathbb{C} . Pour $\lambda \in \mathbb{K}$, $\bar{\lambda}$ désignera le complexe conjugué de λ si $\mathbb{K} = \mathbb{C}$ et le nombre λ lui-même si $\mathbb{K} = \mathbb{R}$. Avec ces conventions, la définition 2 suffit à définir un produit scalaire dans tous les cas.

11.1.2 Inégalité de Cauchy-Schwarz

THEOREME 11.1.1

Soit E un espace préhilbertien, x, y deux éléments de E . On a $|\langle x | y \rangle|^2 \leq \langle x | x \rangle \langle y | y \rangle$. L'égalité a lieu si et seulement si le système (x, y) est lié.

preuve

Pour tout $\lambda \in \mathbb{K}$, on a $\langle \lambda x + y | \lambda x + y \rangle \geq 0$, soit $\lambda \bar{\lambda} \langle x | x \rangle + \bar{\lambda} \langle x | y \rangle + \lambda \langle y | x \rangle + \langle y | y \rangle \geq 0$. Posons $\langle x | y \rangle = \rho u$ avec $\rho \geq 0, |u| = 1$ avec la convention $u = 1$ si $\langle x | y \rangle = 0$. Prenons λ de la forme $\lambda = tu$, avec t réel.

Il vient: $\forall t \in \mathbb{R}, t^2 \langle x | x \rangle + 2t\rho + \langle y | y \rangle \geq 0$. Si $x = 0$, l'inégalité à prouver est triviale. Dans le cas contraire, on a $\langle x | x \rangle > 0$ et l'inégalité à prouver s'obtient en écrivant que le trinôme du second degré $t^2 \langle x | x \rangle + 2t\rho + \langle y | y \rangle$ a un discriminant négatif ou nul.

Supposons qu'on ait l'égalité; si x ou y est nul, le système (x, y) est lié. Sinon, le trinôme ci dessus a un discriminant nul, donc admet une racine double t_0 . On a alors en posant $\lambda_0 = t_0 u$, $\langle \lambda_0 x + y | \lambda_0 x + y \rangle = 0$, ce qui implique (propriété 3 de la définition) $\lambda_0 x + y = 0$. La réciproque est facile.

THEOREME 11.1.2

Soit E un espace préhilbertien. L'application $x \rightarrow \|x\| = \sqrt{\langle x | x \rangle}$ est une norme sur E .

preuve

La seule propriété non triviale est l'inégalité triangulaire. On a $(\|x\| + \|y\|)^2 - \|x + y\|^2 = 2(\|x\| \|y\| - \operatorname{Re}(\langle x | y \rangle))$ pour tous $x, y \in E$. Or $\operatorname{Re}(\langle x | y \rangle) \leq |\langle x | y \rangle| \leq \|x\| \|y\|$ d'après l'inégalité de Cauchy-Schwarz.

Complément important

L'inégalité $\|x + y\| \leq \|x\| + \|y\|$ est stricte sauf si l'on est dans l'un des deux cas suivants : l'un des vecteurs x ou y est nul ou il existe $k > 0$ tel que $y = kx$.

En effet, supposons qu'on ait l'égalité avec x et y non nuls. D'après la preuve ci dessus, on a les deux égalités $\operatorname{Re}(\langle x | y \rangle) = |\langle x | y \rangle|$ et $|\langle x | y \rangle| = \|x\| \|y\|$. De la deuxième égalité, on déduit qu'il existe $k \in \mathbb{K}$ tel que $y = kx$. En reportant dans la première, il vient $\operatorname{Re}(k \|x\|^2) = |k| \|x\|^2$ ce qui implique $\operatorname{Re}(k) = |k|$ donc $k > 0$ puisque $\|x\| > 0$.

Un espace préhilbertien E sera toujours muni de la norme ci dessus définie, associée au produit scalaire. Toutes les notions topologiques habituelles auront donc un sens dans E . La convergence pour la norme préhilbertienne dans le cas des exemples 2 et 3 du paragraphe précédent s'appelle convergence en moyenne quadratique. Une suite (f_n) d'éléments de E converge donc en moyenne quadratique vers $f \in E$ ssi $\lim_{n \rightarrow \infty} \int_0^{2\pi} |f(t) - f_n(t)|^2 dt = 0$.

THEOREME 11.1.3

Soit E un espace préhilbertien. Le produit scalaire est une application continue de $E \times E$ dans \mathbb{K} .

preuve

C'est un conséquence de l'inégalité de Cauchy-Schwarz. En effet, si $x_0, y_0 \in E$ on a pour tous $x, y \in E$
 $|\langle x | y \rangle - \langle x_0 | y_0 \rangle| \leq |\langle x - x_0 | y \rangle| + |\langle x_0 | y - y_0 \rangle| \leq \|x - x_0\| \|y\| + \|x_0\| \|y - y_0\|$.

11.1.3 Formulaire

Soit E un espace préhilbertien, et $x, y \in E$. On a

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

$$\text{Si } \mathbb{K} = \mathbb{R}, \langle x | y \rangle = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2)$$

$$\text{Si } \mathbb{K} = \mathbb{C}, \langle x | y \rangle = \frac{1}{4}(\|x + y\|^2 - \|x - y\|^2 - i(\|x + iy\|^2 - \|x - iy\|^2)) \quad (\text{identité de polarisation}).$$

11.2 Orthogonalité

11.2.1 Propriétés élémentaires

DEFINITION 11.2.1

Deux vecteurs x et y d'un espace préhilbertien E sont dits orthogonaux si leur produit scalaire est nul. On notera $x \perp y$ la relation x est orthogonal à y .

Noter que $\langle x | y \rangle = 0 \Leftrightarrow \langle y | x \rangle = 0$. La relation "est orthogonal à" est symétrique.

THEOREME 11.2.1 (théorème de Pythagore)

Soient E préhilbertien, et $x, y \in E$. Si $x \perp y$ on a $\|x + y\|^2 = \|x\|^2 + \|y\|^2$.

La réciproque est vraie si $\mathbb{K} = \mathbb{R}$.

Attention! La réciproque est fautive si $\mathbb{K} = \mathbb{C}$.

DEFINITION 11.2.2

Soit A une partie d'un espace préhilbertien E . On appelle orthogonal de A et on notera A^\perp l'ensemble des vecteurs de E orthogonaux à tous les vecteurs de A : $A^\perp = \{x \in E \mid \forall y \in A, \langle y | x \rangle = 0\}$.

Remarques

1. $A^\perp = \text{Vect}(A)^\perp$.
2. $\{0\}^\perp = E$.
3. $E^\perp = \{0\}$ (Le produit scalaire est dit non dégénéré).

PROPOSITION 11.2.1

Pour tout vecteur non nul e d'un espace préhilbertien E , l'orthogonal de e est un hyperplan fermé.

Cela résulte immédiatement du fait que l'application $x \rightarrow \langle e | x \rangle$ est une forme linéaire continue non nulle.

COROLLAIRE 11.2.1

Pour toute partie A d'un espace préhilbertien E , A^\perp est un sous espace vectoriel fermé

En effet, $A^\perp = \bigcap_{a \in A} a^\perp$ est une intersection d'hyperplans fermés.

DEFINITION 11.2.3

Un système $(e_i)_{i \in I}$ de vecteurs d'un espace préhilbertien E est dit orthogonal si $i \neq j \Rightarrow \langle e_i | e_j \rangle = 0$. Il est dit orthonormé si il est orthogonal et si pour tout i , $\|e_i\| = 1$.

THEOREME 11.2.2

Un système orthogonal de vecteurs non nuls (en particulier un système orthonormé) est un système libre.

preuve

Soit $(e_i)_{i \in I}$ un système orthogonal de vecteurs non nuls. Soit $J \subset I$ un sous ensemble fini. Il s'agit de montrer que la famille $(e_j)_{j \in J}$ est libre. Soit $(\lambda_j)_{j \in J}$ des scalaires tels que $\sum_{j \in J} \lambda_j e_j = 0$. Soit $k \in J$. On a $0 = \left\langle \sum_{j \in J} \lambda_j e_j \mid e_k \right\rangle$ ce qui se réduit à $\overline{\lambda_k} \|e_k\|^2 = 0$ donc $\lambda_k = 0$. ■

PROPOSITION 11.2.2

Soit (e_1, \dots, e_n) une base orthonormée d'un espace préhilbertien de dimension finie E . Si $x = \sum_{k=1}^{k=n} x_k e_k$, $y = \sum_{k=1}^{k=n} y_k e_k$ on a

$$x_k = \langle e_k \mid x \rangle \quad \langle x \mid y \rangle = \sum_{k=1}^{k=n} \overline{x_k} y_k \quad \text{et} \quad \|x\|^2 = \sum_{k=1}^{k=n} |x_k|^2$$

11.2.2 Procédé d'orthogonalisation de Gram-Schmidt

THEOREME 11.2.3

Soit $(u_i)_{1 \leq i \leq n}$ (resp $(u_i)_{i \in \mathbb{N}}$) une suite libre finie (resp. une suite libre) de vecteurs de l'espace préhilbertien E . Il existe une unique suite orthonormée $(e_i)_{1 \leq i \leq n}$ (resp. $(e_i)_{i \in \mathbb{N}}$) vérifiant, pour tout entier k compris entre 1 et n (resp. pour tout entier k) : $\text{Vect}(e_1, \dots, e_n) = \text{Vect}(u_1, \dots, u_n)$ et $\langle e_k \mid u_k \rangle > 0$

preuve

Soit $V_k = \text{Vect}(u_1, \dots, u_k)$.

On a $V_1 \subset V_2 \subset \dots \subset V_k$ et $\dim(V_j) = j$; la droite $\mathbb{K}u_j$ est un supplémentaire, dans V_j de V_{j-1} , pour $2 \leq j \leq k$.

1. Existence

On construit la suite $(e_k)_{1 \leq k \leq n}$ par récurrence sur k . On pose $e_1 = u_1 / \|u_1\|$. Supposons les vecteurs e_j construits pour $1 \leq j \leq k$. Ils vérifient donc $\text{Vect}(e_1, \dots, e_j) = \text{Vect}(u_1, \dots, u_j)$ et $\langle e_j \mid u_j \rangle > 0$ pour $1 \leq j \leq k$. En particulier, on a $V_{k+1} = \text{Vect}(e_1, \dots, e_k, u_{k+1})$. On cherche $e'_{k+1} \in V_{k+1}$, orthogonal aux e_j , $1 \leq j \leq k$ tel que $V_{k+1} = \text{Vect}(e_1, \dots, e_k, e'_{k+1})$. Il est clair que cette dernière condition sera remplie si on choisit $e'_{k+1} = u_{k+1} + \sum_{j=1}^k \lambda_j e_j$, quelles que soient les valeurs des scalaires λ_j . On écrit ensuite $\langle e_j \mid e'_{k+1} \rangle = 0$, ce qui donne $\langle e_j \mid u_{k+1} \rangle + \lambda_j = 0$. Le vecteur $e'_{k+1} = u_{k+1} - \sum_{j=1}^k \langle e_j \mid u_{k+1} \rangle e_j$ n'est pas nul, car $u_{k+1} \notin V_k$. On peut donc définir $e_{k+1} = e'_{k+1} / \|e'_{k+1}\|$. La suite $(e_j)_{1 \leq j \leq k+1}$ vérifie les propriétés voulues.

2. Unicité

Supposons que la suite $(\varepsilon_k)_{1 \leq k \leq n}$ réponde aussi à la question. $\text{Vect}(\varepsilon_1) = V_1$ donc ε_1 est colinéaire à e_1 : il existe $\lambda \in \mathbb{K}$ tel que $\varepsilon_1 = \lambda e_1$. On a $\|\varepsilon_1\| = |\lambda| \|e_1\|$ et $\langle \varepsilon_1 \mid u_1 \rangle = \overline{\lambda} \langle e_1 \mid u_1 \rangle$, d'où $|\lambda| = 1$ et $\overline{\lambda} > 0$, soit $\lambda = 1$ et $\varepsilon_1 = e_1$. Supposons prouvé $\varepsilon_j = e_j$ pour $1 \leq j \leq k$ et montrons $\varepsilon_{k+1} = e_{k+1}$. Puisque $\varepsilon_{k+1} \in V_{k+1}$ et que (e_1, \dots, e_{k+1}) est une base orthonormée de V_{k+1} , on peut écrire $\varepsilon_{k+1} = \sum_{j=1}^{k+1} \langle e_j \mid \varepsilon_{k+1} \rangle e_j = \sum_{j=1}^{j=k} \langle e_j \mid \varepsilon_{k+1} \rangle e_j + \langle e_{k+1} \mid \varepsilon_{k+1} \rangle e_{k+1} = \langle e_{k+1} \mid \varepsilon_{k+1} \rangle e_{k+1}$. Les vecteurs ε_{k+1} et e_{k+1} sont donc colinéaires. En raisonnant comme dans le cas de ε_1 , on obtient l'égalité.

En résumé, le système cherché est donné par les formules :

$$e'_1 = u_1, \quad e_1 = \frac{e'_1}{\|e'_1\|}, \quad \text{puis, pour } k \geq 2, \quad e'_k = u_k - \sum_{j=1}^{j=k-1} \langle e_j \mid u_k \rangle e_j, \quad \text{et } e_k = \frac{e'_k}{\|e'_k\|}$$

THEOREME 11.2.4

Tout espace préhilbertien de dimension finie non nulle possède une base orthonormée (bon).

preuve

On choisit une base quelconque de E et on lui applique le procédé d'orthogonalisation pour obtenir une base orthonormée de E . On peut aussi démontrer directement ce théorème, par récurrence sur n . (Voir par exemple la démonstration de l'existence de bases b -orthogonales pour une forme bilinéaire symétrique b).

11.3 Projection orthogonale sur un sev de dimension finie

THEOREME 11.3.1

Soit E un espace préhilbertien et F un sous espace de E non nul et de dimension finie. Alors :

1. $E = F \oplus F^\perp$.
2. La projection p sur F parallèlement à F^\perp est linéaire continue de norme 1.
3. Pour tout x de E on a $\|x - p(x)\| = \inf_{y \in F} \|x - y\| = d(x, F)$ et $p(x)$ est l'unique vecteur v de F vérifiant $\|x - v\| = d(x, F)$. Autrement dit, pour tout $y \in F$, on a $\|x - p(x)\| \leq \|x - y\|$ et cette inégalité est stricte si $y \neq p(x)$
4. Si $(e_i)_{1 \leq i \leq n}$ est une bon de F , on a $p(x) = \sum_{1 \leq i \leq n} \langle e_i | x \rangle e_i$.

L'application p s'appelle la projection orthogonale sur F .

preuve

Tout d'abord, $F \cap F^\perp = \{0\}$ car tout vecteur de cette intersection est orthogonal à lui même donc est nul. Soit $(e_i)_{1 \leq i \leq n}$ une bon de F ; posons $p(x) = \sum_{1 \leq i \leq n} \langle e_i | x \rangle e_i$; p est linéaire et pour tout x on a $p(x) \in F$. Pour $1 \leq k \leq n$ on a $\langle e_k | x - p(x) \rangle = \langle e_k | x \rangle - \sum_{1 \leq i \leq n} \langle e_i | x \rangle \delta_{i,k} = 0$ donc $x - p(x) \in F^\perp$. Comme $x = (x - p(x)) + p(x)$ on a prouvé 1) et 4).

Soit $y \in F$; $x - p(x)$ est orthogonal à $p(x) - y$ donc (Pythagore) $\|x - y\|^2 = \|x - p(x)\|^2 + \|p(x) - y\|^2$ ce qui prouve 3). On a aussi $\|p(x)\|^2 = \|x\|^2 - \|x - p(x)\|^2$ donc $\|p(x)\| \leq \|x\|$; on en déduit que l'application linéaire p est continue et que $\|p\| \leq 1$. Enfin, F étant non nul (c'est uniquement ici qu'intervient cette hypothèse), on peut trouver un vecteur non nul $a \in F$ et on a alors $p(a) = a$ donc $\|p(a)\| = \|a\|$ ce qui prouve que $\|p\| \geq 1$ et donc l'égalité $\|p\| = 1$.

Remarque

Avec les notations de la démonstration du théorème de Gram-Schmidt, si on note, pour $k \geq 2$, $p_k : V_k \rightarrow V_{k-1}$ la projection orthogonale de V_k sur V_{k-1} (c'est à dire la restriction à V_k de la projection orthogonale sur V_{k-1} , la formule $e'_k = u_k - \sum_{j=1}^{k-1} \langle e_j | u_k \rangle e_j$ s'écrit $e'_k = u_k - p_k(u_k)$.

11.4 Suites orthonormales

Si $(\varepsilon_1, \dots, \varepsilon_n)$ est une base orthonormée d'un espace préhilbertien de dimension finie E , on a pour tout $x \in E$ l'égalité $x = \sum_{1 \leq j \leq n} \langle e_j | x \rangle e_j$. On se propose de voir comment ce résultat est modifié si on ne suppose plus E de dimension finie. Dans tout ce paragraphe, $(E, \langle | \rangle)$ est un espace préhilbertien, supposé de dimension infinie. Soit $(e_k)_{k \in \mathbb{N}}$ une suite orthonormée de vecteurs de E . Pour $x \in E$ on pose $x_k = \langle e_k | x \rangle$.

THEOREME 11.4.1 (Inégalité de Bessel)

Soient E un espace préhilbertien, $(e_k)_{k \in \mathbb{N}}$ une suite orthonormée de vecteurs de E . Soient $x \in E$, et pour tout k entier $x_k = \langle e_k | x \rangle$. La série numérique $\sum_{k \geq 0} |x_k|^2$ converge et $\sum_{k \geq 0} |x_k|^2 \leq \|x\|^2$.

preuve

Soit p_n la projection orthogonale sur $F_n = \text{Vect}(e_0, \dots, e_n)$; (e_0, \dots, e_n) étant orthonormé on a $p_n(x) = \sum_{0 \leq k \leq n} x_k e_k$ et

$\|p_n(x)\|^2 = \sum_{0 \leq k \leq n} |x_k|^2$; on a $x - p_n(x) \perp p_n(x)$ donc $\|x\|^2 = \|p_n(x)\|^2 + \|x - p_n(x)\|^2$ soit

$$\left\| x - \sum_{k=0}^{k=n} x_k e_k \right\|^2 = \|x\|^2 - \sum_{k=0}^{k=n} |x_k|^2$$

On en déduit $\sum_{k=0}^{k=n} |x_k|^2 \leq \|x\|^2$; les sommes partielles de la série à termes positifs $\sum |x_k|^2$ sont majorées, ce qui assure sa convergence ; on obtient l'inégalité voulue en faisant tendre n vers l'infini.

THEOREME 11.4.2

Soient E un espace préhilbertien, (e_k) une suite orthonormée de vecteurs de E . Soit enfin $\mathcal{E} = \text{Vect}(e_0, \dots, e_n, \dots)$. Les propriétés suivantes sont équivalentes :

- 1) $\forall x \in E \quad \|x\|^2 = \sum_{k \geq 0} |x_k|^2$ où $x_k = \langle e_k | x \rangle$
- 2) $\forall (x, y) \in E^2$ la série de terme général $\overline{x_k} y_k$ est absolument convergente et $\langle x | y \rangle = \sum_{k=0}^{\infty} \overline{x_k} y_k$.
- 3) $\forall x \in E \quad x = \sum_{k=0}^{\infty} x_k e_k$ i.e. $\lim_{n \rightarrow \infty} \left\| x - \sum_{k=0}^{k=n} x_k e_k \right\| = 0$.
- 4) \mathcal{E} est dense dans E .

preuve

- (1) \Leftrightarrow (2) résulte de l'identité de polarisation et de la majoration $|\overline{x_k} y_k| \leq \frac{1}{2} (|x_k|^2 + |y_k|^2)$.
- (1) \Leftrightarrow (3) résulte de l'égalité $\left\| x - \sum_{k=0}^{k=n} x_k e_k \right\|^2 = \|x\|^2 - \sum_{k=0}^{k=n} |x_k|^2$ obtenue dans la preuve du théorème 1.
- (3) \Rightarrow (4) est clair : pour tout $x \in E$ la suite $\left(\sum_{k=0}^{k=n} x_k e_k \right)_{n \geq 0}$ d'éléments de \mathcal{E} converge vers x .
- (4) \Rightarrow (3) Soient $x \in E$ et $\varepsilon > 0$. Par hypothèse il existe un élément $z \in \mathcal{E}$ tel que $\|x - z\| \leq \varepsilon$. Par définition de \mathcal{E} il existe un entier $N \geq 0$, des scalaires λ_k $0 \leq k \leq n$ tels que $z = \sum_{k=0}^{k=N} \lambda_k e_k$. Soit alors $n \geq N$. Le vecteur $X_n = \sum_{k=0}^{k=n} x_k e_k$ est le projeté orthogonal de x sur $\mathcal{E}_n = \text{Vect}(e_0, \dots, e_n)$ et $z \in \mathcal{E}_n$ donc $\|x - X_n\| \leq \|x - z\| \leq \varepsilon$.
Donc la suite (X_n) converge dans E vers x .

Suites orthonormales indexées par \mathbb{Z}

A. Suites complexes indexées par \mathbb{Z}

Soit $(z_n)_{n \in \mathbb{Z}}$ une famille de nombres complexes indexées par l'ensemble \mathbb{Z} des entiers relatifs. On posera pour m entier naturel $u_m = z_m + z_{-m}$ (on notera que $u_0 = 2z_0$). Soit, pour $n \in \mathbb{N}$, $S_n = \sum_{k=-n}^{k=n} z_k = \frac{u_0}{2} + \sum_{k=1}^{k=n} u_k$. Si la suite (S_n) converge on notera S sa limite. On prendra garde au fait qu'il ne s'agit pas d'une série au sens usuel du terme. En particulier S peut exister alors que les séries $\sum_{m \geq 0} z_m$ et $\sum_{m \geq 0} z_{-m}$ divergent ; pour s'en convaincre il suffit de prendre $z_n = n$ pour $n \in \mathbb{Z}$. Les propriétés ci dessous sont élémentaires.

Propriété 1. La suite (S_m) converge ssi la série de terme général u_m est convergente et alors $S = \frac{u_0}{2} + \sum_{m=1}^{\infty} u_m$.

Propriété 2. Si les deux séries $\sum_{m \geq 1} z_m$ et $\sum_{m \geq 1} z_{-m}$ convergent, la suite (S_m) converge et $S = \sum_{m=0}^{\infty} z_m + \sum_{m=1}^{\infty} z_{-m}$.

Propriété 3. Les trois propriétés suivantes sont équivalentes :

1. Les deux séries $\sum_{m \geq 1} z_m$ et $\sum_{m \geq 1} z_{-m}$ sont absolument convergentes.
2. La série $\sum_{m=1}^{\infty} (|z_m| + |z_{-m}|)$ converge.
3. Il existe un réel K tel que pour toute partie finie F de \mathbb{Z} , $\sum_{k \in F} |z_k| \leq K$.

Dans ce cas, on dit que la série $\sum_{n=-\infty}^{+\infty} z_n$ est absolument convergente et on notera $S = \sum_{n=-\infty}^{+\infty} z_n$.

(Ce cas est celui où la famille $(z_n)_{n \in \mathbb{Z}}$ est sommable).

Propriété 4. Si pour tout m , $z_m \in \mathbb{R}_+$, la suite (S_n) converge ssi les deux séries $\sum_{m \geq 1} z_m$ et $\sum_{m \geq 1} z_{-m}$ convergent .

B. Suite orthonormale indexée par \mathbb{Z}

Considérons de nouveau un espace préhilbertien $(E, \langle \cdot | \cdot \rangle)$ de dimension infinie et soit $(e_n)_{n \in \mathbb{Z}}$ une famille orthonormale d'éléments de E , indexée par \mathbb{Z} . Pour $x \in E$ et $n \in \mathbb{Z}$ on posera encore $x_n = \langle e_n | x \rangle$. Soit enfin $\mathcal{E} = \text{Vect}((e_n)_{n \in \mathbb{Z}})$ le sous-espace de E engendré par les e_n . Le théorème 11.4.2 se généralise comme suit :

THEOREME 11.4.3

Sous les hypothèses précédentes, les propriétés suivantes sont équivalentes :

- 1) Pour tout $x \in E$ la série $\sum_{-\infty}^{+\infty} |x_n|^2$ est convergente et $\|x\|^2 = \sum_{-\infty}^{+\infty} |x_n|^2$.
- 2) $\forall (x, y) \in E^2$ la série $\sum_{-\infty}^{+\infty} \overline{x_k} y_k$ est absolument convergente et $\langle x | y \rangle = \sum_{-\infty}^{+\infty} \overline{x_k} y_k$.
- 3) $\forall x \in E \lim_{n \rightarrow \infty} \left\| x - \sum_{k=-n}^{k=n} x_k e_k \right\| = 0$.
- 4) \mathcal{E} est dense dans E .

La preuve est semblable à celle du théorème 11.4.2

11.5 Polynômes orthogonaux

11.5.1 Introduction

Le programme de l'agrégation comporte "exemples de suites de polynômes orthogonaux". Ce paragraphe n'est donc pas stricto sensu au programme ; il donne le cadre d'ensemble et quelques propriétés générales des suites de polynômes orthogonaux. Ce paragraphe suppose connue la notion de fonction intégrable sur un intervalle quelconque de \mathbb{R} .

On se donne un intervalle I de \mathbb{R} non vide et non réduit à un point et une fonction $p : I \rightarrow \mathbb{R}$ appelée poids vérifiant les trois propriétés suivantes :

- (P1) p est continue sur I .
- (P2) $\forall t \in I p(t) > 0$.
- (P3) Pour tout $n \in \mathbb{N}$, la fonction $t \rightarrow t^n p(t)$ est intégrable sur I .

E désigne l'ensemble des fonctions $f : I \rightarrow \mathbb{R}$ continues telles que $t \rightarrow |f(t)|^2 p(t)$ soit intégrable sur I .

Soient $f, g \in E$. L'inégalité $|f(t)g(t)| p(t) \leq \frac{1}{2} (|f(t)|^2 + |g(t)|^2) p(t)$ montre que $t \rightarrow f(t)g(t)p(t)$ est intégrable sur I . On peut donc poser $\langle f | g \rangle = \int_I f(t)g(t)p(t)dt$.

PROPOSITION 11.5.1

$(E, \langle \cdot | \cdot \rangle)$ est un espace préhilbertien réel.

preuve

E est non vide car il contient la fonction nulle. Si $f \in E$ et $\lambda \in \mathbb{R}$, on a $\lambda f \in E$. Si $f, g \in E$, on a $0 \leq (f+g)^2 p = f^2 p + g^2 p + 2fgp \leq 2f^2 p + 2g^2 p$, donc $(f+g)^2 p$ est intégrable sur I et E est bien un \mathbb{R} -espace vectoriel. $\langle \cdot | \cdot \rangle$ est clairement une forme bilinéaire symétrique positive. Enfin, $\int_I f^2(t)p(t)dt = 0$ implique $\forall t \in I f^2(t)p(t) = 0$ car $f^2 p$ est positive et continue donc, vu (P2), $f = 0$. ■

Notons provisoirement F l'espace des restrictions à I de fonctions polynômes. D'après (P3), E contient les fonctions monômes, donc $F \subset E$. Par ailleurs, on sait, I n'étant pas réduit à un point que F est isomorphe à $\mathbb{R}[X]$. On pourra donc identifier F et $\mathbb{R}[X]$. Dans toute la suite, E et F seront munis du produit scalaire ainsi défini.

Posons pour tout $n \in \mathbb{N}$, $e_n(t) = t^n$ et $V_n = \text{Vect}(e_0, \dots, e_n) \subset F$. V_n est un sous-espace vectoriel de dimension $n+1$ de F , isomorphe à $\mathbb{R}_n[X]$. L'application du procédé d'orthogonalisation de Gram-Schmidt à la suite (e_k) définit une suite orthonormée de polynômes (p_k) .

On appelle suite de polynômes orthogonaux toute suite $(P_n)_{n \geq 0}$ avec $P_n = k_n p_n$ où k_n est une constante réelle non nulle.

Par construction même, on a donc $P_n \in V_n$ et $P_n \notin V_{n-1}$, donc pour tout n , P_n est de degré n .

Attention à la terminologie dans ce paragraphe, on utilisera le terme de polynôme normalisé dans son sens habituel : un polynôme est normalisé si le coefficient du terme de plus haut degré est 1 (et non pas si sa norme vaut 1).

Dans les exemples classiques, la suite $(P_n)_{n \geq 0}$ n'est pas orthonomée, mais seulement une suite orthogonale. Le choix des k_n dépend des exemples. L'une des choix possibles est d'imposer que les P_n soient normalisés.

11.5.2 Exemples classiques

$$\begin{array}{lll} I = [-1, 1] & ; & p(t) = 1 & ; & \text{polynômes de Legendre} \\ I =] - 1, 1 [& ; & p(t) = \frac{1}{\sqrt{1-t^2}} & ; & \text{polynômes de Tchebycheff} \\ I = [0, +\infty [& ; & p(t) = e^{-t} & ; & \text{polynômes de Laguerre} \\ I = \mathbb{R} & ; & p(t) = e^{-t^2/2} & ; & \text{polynômes d'Hermite} \end{array}$$

11.5.3 Propriétés générales

Dans tout ce qui suit, p est fixé, et (P_n) est la suite de polynômes orthogonaux pour le poids p , telle que pour tout n , P_n soit normalisé.

Justifiez l'emploi de l'article défini " la " dans la définition ci dessus.

On a donc en particulier $P_0 = 1$. Vérifier que $P_1 = X + \beta$ avec $\beta = -\frac{\int_I tp(t)dt}{\int_I p(t)dt}$.

Par construction P_n est orthogonal au sous espace engendré par P_0, \dots, P_{n-1} c'est à dire à V_{n-1} identifié avec $\mathbb{R}_{n-1}[X]$. En particulier $P_n - XP_{n-1}$ est de degré strictement inférieur à n donc orthogonal à P_n . On a donc l'égalité

$$\boxed{\forall n \geq 1 \quad \langle XP_{n-1} | P_n \rangle = \langle P_n | P_n \rangle} \quad (11.1)$$

Attention! Cette égalité n'a pas lieu si les polynômes ne sont pas supposés normalisés.

THEOREME 11.5.1

Pour tout n , P_n admet n racines réelles distinctes qui appartiennent toutes à l'intérieur de I .

preuve

Soit $n \geq 1$. On a $0 = \langle 1 | P_n \rangle = \int_I P_n(t)p(t)dt$. On en déduit que P_n admet au moins une racine d'ordre de multiplicité impaire dans l'intérieur de I car sinon, la fonction pP_n serait continue, non nulle et de signe constant, donc l'intégrale serait différente de 0. Soient x_1, \dots, x_r $1 \leq r \leq n$ les racines distinctes de P_n de multiplicité impaire et appartenant à l'intérieur de I . Posons $Q(X) = (X - x_1) \cdots (X - x_r)$. Le polynôme $P_n Q$ est de signe constant sur I donc $\langle P_n | Q \rangle = \int_I P_n(t)Q(t)p(t)dt \neq 0$. Par conséquent, $Q \notin (\mathbb{R}P_n)^\perp = \mathbb{R}_{n-1}[X]$. Comme $\deg(Q_n) \leq n$ on en déduit $\deg(Q) = n$ et $r = n$. Il en résulte que les racines de P_n sont simples.

THEOREME 11.5.2 (Relations de récurrence)

Il existe deux suites de réels $(\lambda_n)_{n \geq 2}$ et $(\mu_n)_{n \geq 2}$ telles que, pour tout $n \geq 2$

$$\boxed{P_n = (X + \lambda_n)P_{n-1} - \mu_n P_{n-2}} \quad (11.2)$$

De plus, on a $\mu_n > 0$ pour tout n .

Attention! si les P_n ne sont pas normalisés la forme de la relation est légèrement différente.

preuve

Soit $n > 1$. $P_n - XP_{n-1}$ appartient à $\mathbb{R}_{n-1}[X]$ dont une base est (P_0, \dots, P_{n-1}) . On peut écrire $P_n - XP_{n-1} = \sum_{j=0}^{n-1} \gamma_j P_j$ où les γ_j sont des réels. La suite (P_j) étant orthogonale, on en déduit $\langle P_n - XP_{n-1} | P_j \rangle = \gamma_j \langle P_j | P_j \rangle$. Or $\langle P_n | P_j \rangle = 0$, donc $\gamma_j \langle P_j | P_j \rangle = -\langle XP_{n-1} | P_j \rangle = -\langle P_{n-1} | XP_j \rangle$, la dernière égalité résultant de l'expression du produit scalaire. Si $j \leq n-3$, XP_j est de degré au plus $n-2$, donc orthogonal à P_{n-1} . Comme $\langle P_j | P_j \rangle > 0$, on en déduit $\gamma_j = 0$. Pour $j = n-2$, il reste $\gamma_{n-2} \langle P_{n-2} | P_{n-2} \rangle = -\langle P_{n-1} | XP_{n-2} \rangle = -\langle P_{n-1} | P_{n-1} \rangle < 0$ d'après (1). Ceci achève la preuve, en prenant $\lambda_n = \gamma_{n-1}$ et $\mu_n = -\gamma_{n-2} = \frac{\|P_{n-1}\|^2}{\|P_{n-2}\|^2} > 0$.

THEOREME 11.5.3

Si l'intervalle I est compact, $F = \mathbb{R}[X]$ est dense dans l'espace $(E, \langle \cdot | \cdot \rangle)$.

preuve

Soit $f \in E$. Il existe une suite de polynômes (A_n) telle que $\|f - A_n\|_\infty = \sup_{t \in I} |f(t) - A_n(t)|$ tende vers 0 (théorème de Weierstrass). On a alors $\|f - A_n\|^2 = \int_I |f(t) - A_n(t)|^2 p(t) dt \leq \|f - A_n\|_\infty^2 \int_I p(t) dt$ quantité qui tend vers 0 quand n tend vers l'infini. ■

11.6 L'espace $l^2_{\mathbb{N}}(\mathbb{C})$

Cette partie présente un exemple d'espace préhilbertien complexe de dimension infinie, particulièrement important en analyse.

DEFINITION 11.6.1

On note $l^2_{\mathbb{N}}(\mathbb{C})$ l'ensemble des suites $(u_n)_{n \in \mathbb{N}}$ de nombres complexes telles que la série $\sum |u_n|^2$ converge.

De même on note $l^2_{\mathbb{Z}}(\mathbb{C})$ l'ensemble des suites $(u_n)_{n \in \mathbb{Z}}$, indexées par \mathbb{Z} , de nombres complexes telles que les séries $\sum_{n \geq 0} |u_n|^2$ et $\sum_{n \geq 1} |u_{-n}|^2$ convergent.

Lemme 11.6.1

Soient (u_n) et (v_n) deux éléments de $l^2_{\mathbb{N}}(\mathbb{C})$ (resp. $l^2_{\mathbb{Z}}(\mathbb{C})$). La série $\sum_{n \geq 0} \overline{u_n} v_n$ (resp. $\sum_{n=-\infty}^{+\infty} \overline{u_n} v_n$) est absolument convergente.

Cela résulte immédiatement de l'inégalité $|\overline{x}y| \leq \frac{1}{2}(|x|^2 + |y|^2)$ valable pour tous nombres complexes x et y .

THEOREME 11.6.1

$l^2_{\mathbb{N}}(\mathbb{C})$ (resp. $l^2_{\mathbb{Z}}(\mathbb{C})$) est un \mathbb{C} -espace vectoriel.

L'application qui au couple $(u = (u_n), v = (v_n)) \in l^2_{\mathbb{N}}(\mathbb{C}) \times l^2_{\mathbb{N}}(\mathbb{C})$ associe le complexe $\langle u | v \rangle = \sum_{n=0}^{\infty} \overline{u_n} v_n$ est un produit scalaire sur $l^2_{\mathbb{N}}(\mathbb{C})$.

De même l'application qui au couple $(u = (u_n), v = (v_n)) \in l^2_{\mathbb{Z}}(\mathbb{C}) \times l^2_{\mathbb{Z}}(\mathbb{C})$ associe le complexe $\langle u | v \rangle = \sum_{n=-\infty}^{\infty} \overline{u_n} v_n$ est un produit scalaire sur $l^2_{\mathbb{Z}}(\mathbb{C})$.

Les vérifications sont faciles et laissées au lecteur.

THEOREME 11.6.2

Les espaces $l^2_{\mathbb{N}}(\mathbb{C})$ et $l^2_{\mathbb{Z}}(\mathbb{C})$ munis de la norme associée au produit scalaire défini ci dessus sont complets.

preuve pour $l^2_{\mathbb{N}}(\mathbb{C})$

Soit $(u^j)_{j \geq 0}$ une suite de Cauchy d'éléments de $l^2_{\mathbb{N}}(\mathbb{C})$. Chaque u^j est donc une suite $u^j = (u_n^j)_{n \geq 0}$ de nombres complexes telle que la série $\sum_{n \geq 0} |u_n^j|^2$ converge. Le fait que la suite (u^j) soit de Cauchy s'exprime par

$$\forall \varepsilon > 0, \exists J(\varepsilon) \in \mathbb{N}, \forall j, k \in \mathbb{N}, j \geq J(\varepsilon) \Rightarrow \|u^j - u^{j+k}\| \leq \varepsilon$$

soit encore

$$(C) \quad \forall \varepsilon > 0, \exists J(\varepsilon) \in \mathbb{N}, \forall j, k \in \mathbb{N}, j \geq J(\varepsilon) \Rightarrow \sum_{n \geq 0} |u_n^j - u_n^{j+k}|^2 \leq \varepsilon^2$$

a) Soit $\varepsilon > 0$ et $n \in \mathbb{N}$. Pour tout $j \geq J(\varepsilon)$ et tout $k \in \mathbb{N}$ on a $|u_n^j - u_n^{j+k}| \leq \|u^j - u^{j+k}\| \leq \varepsilon$ ce qui prouve que pour tout n , la suite $(u_n^j)_{j \in \mathbb{N}}$ est de Cauchy dans \mathbb{C} . Elle est donc convergente. Notons v_n sa limite.

b) Soit N un entier naturel fixé quelconque. Soit $j \geq J(\varepsilon)$ et $k \in \mathbb{N}$. On a d'après (C)

$$\sum_{n=0}^{n=N} |u_n^j - u_n^{j+k}|^2 \leq \sum_{n \geq 0} |u_n^j - u_n^{j+k}|^2 \leq \varepsilon^2$$

En faisant tendre k vers l'infini dans l'inégalité $\sum_{n=0}^{n=N} |u_n^j - u_n^{j+k}|^2 \leq \varepsilon^2$ on obtient

$$(E) \quad \forall j \in \mathbb{N} \quad j \geq J(\varepsilon) \Rightarrow \sum_{n=0}^{n=N} |u_n^j - v_n|^2 \leq \varepsilon^2$$

- En choisissant par exemple $j = J(1)$, on voit que la série à termes positifs $w_n = |u_n^j - v_n|^2$ a ses sommes partielles bornées. Elle est donc convergente, ce qui prouve que $(u^j - v) \in l_{\mathbb{N}}^2(\mathbb{C})$. Comme $u^j \in l_{\mathbb{N}}^2(\mathbb{C})$ on en déduit que $v \in l_{\mathbb{N}}^2(\mathbb{C})$.
- En prenant maintenant $\varepsilon > 0$ quelconque, $j \geq J(\varepsilon)$ et en faisant tendre N vers l'infini dans (E), on obtient

$$\sum_{n=0}^{\infty} |u_n^j - v_n|^2 \leq \varepsilon^2$$

On a donc montré

$$\forall \varepsilon > 0 \quad \exists J \in \mathbb{N} \quad \forall j \in \mathbb{N} \quad j \geq J \Rightarrow \|u^j - v\| \leq \varepsilon$$

ce qui prouve que la suite (u^j) converge, dans $l_{\mathbb{N}}^2(\mathbb{C})$ vers v . Ceci achève la démonstration.

La preuve pour $l_{\mathbb{Z}}^2(\mathbb{C})$ est similaire. On a des résultats analogues pour les espaces préhilbertiens réels $l_{\mathbb{N}}^2(\mathbb{R})$ et $l_{\mathbb{Z}}^2(\mathbb{R})$.

12

Séries de Fourier

12.1 Rappels du cours d'Intégration

12.1.1 Fonctions périodiques

DEFINITION 12.1.1

Une fonction 2π -périodique sera dite continue par morceaux si sa restriction à l'intervalle $[0, 2\pi]$ l'est. Il en résulte que la restriction de f à tout intervalle compact est continue par morceaux. On notera \mathcal{P} la \mathbb{C} algèbre des fonctions 2π -périodiques et continues par morceaux.

Une fonction 2π -périodique f de \mathbb{R} dans \mathbb{C} sera dite C^1 par morceaux si il existe une subdivision $(x_0 = 0, x_1, \dots, x_n = 2\pi)$ de l'intervalle $[0, 2\pi]$ et pour $i \in \{0, 1, \dots, n-1\}$ des fonctions de classe C^1 , $g_i : [x_i, x_{i+1}] \rightarrow \mathbb{C}$ telles que

$$\forall x \in]x_i, x_{i+1}[\quad f(x) = g_i(x)$$

On notera qu'une fonction C^1 par morceaux n'est pas nécessairement continue. Elle est par contre continue par morceaux.

Le lemme suivant facile est d'un usage courant.

Lemme 12.1.1

Soit f de \mathbb{R} dans \mathbb{C} continue par morceaux ; on a pour tout $a \in \mathbb{R}$

$$\int_0^{2\pi} f(t) dt = \int_a^{a+2\pi} f(t) dt.$$

Si f est continue par morceaux, elle admet en tout point x des limites à gauche et à droite. On notera un peu abusivement

$$f(x^+) = \lim_{\substack{y \rightarrow x \\ y > x}} f(y) \quad \text{et} \quad f(x^-) = \lim_{\substack{y \rightarrow x \\ y < x}} f(y)$$

DEFINITION 12.1.2

On notera E l'espace vectoriel des fonctions 2π -périodiques, continues par morceaux vérifiant

$$\forall x \in \mathbb{R} \quad f(x) = \frac{1}{2} (f(x^+) + f(x^-)).$$

Pour f et g dans E on pose

$$\langle f, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} \overline{f(t)} g(t) dt.$$

THEOREME 12.1.1

\langle, \rangle est un produit scalaire sur E et donc (E, \langle, \rangle) est un espace préhilbertien complexe.

preuve

Soit $f \in E$ telle que $\langle f | f \rangle = 0$. Soit $b \in \mathbb{R}$ tel que b et $b + 2\pi$ soient des points de continuités de f . La fonction $|f|^2$ est positive et vérifie $\int_b^{b+2\pi} |f(t)|^2 dt = 0$. D'après le théorème 10.3.2, elle est nulle en tout point de continuité de f sur $[b, b + 2\pi]$, c'est à dire partout sur cet intervalle sauf au plus en un nombre fini de points. Soit x un point de discontinuité de f . Il existe un intervalle $I =]x - a, x + a[\subset [b, b + 2\pi]$ tel que f est continue sur $I \setminus \{x\}$ donc nulle sur cet ensemble. Les limites à droite et à gauche de f en x sont nulles, donc compte tenu de la définition de E , on a $f(x) = 0$ et finalement $f(x) = 0$ pour $x \in [b, b + 2\pi]$. Par périodicité, $f = 0$. La vérification des autres propriétés d'un produit scalaire est immédiate.

Remarque

On aura à considérer deux normes distinctes sur E . D'une part, la norme définie par le produit scalaire ; on la notera $\| \cdot \|_2$. D'autre part, la norme de la convergence uniforme que l'on notera $\| \cdot \|_\infty$. On aura donc, pour $f \in E$

$$\|f\|_2 = \sqrt{\frac{1}{2\pi} \int_0^{2\pi} |f(t)|^2 dt} \quad \text{et} \quad \|f\|_\infty = \sup_{t \in [0, 2\pi]} |f(t)|$$

On a pour toute $f \in E$, $\|f\|_2 \leq \|f\|_\infty$.

Exercice : Vérifier que ces deux normes ne sont pas équivalentes.

Lemme 12.1.2

Soit $f \in \mathcal{P}$; il existe une fonction $f_0 \in E$ et une seule dont la restriction à $[0, 2\pi]$ coïncide avec celle de f sauf au plus en un nombre fini de points. On a $\int_0^{2\pi} f_0(t) dt = \int_0^{2\pi} f(t) dt$. On dira que f_0 est la fonction de E associée à f .

preuve

Existence : On pose, pour tout $x \in \mathbb{R}$ $f_0(x) = \frac{1}{2} (f(x^+) + f(x^-))$. On a $f_0(x) = f(x)$ en tout point x où f est continue. Soit $x \in \mathbb{R}$. Il existe un $\eta > 0$ telle que f soit continue sur $A =]x - \eta, x[\cup]x, x + \eta[$. Pour $y \in A$ on a donc $f_0(y) = f(y)$ et par conséquent $f_0(x^+) = f(x^+)$ et $f_0(x^-) = f(x^-)$ donc $f_0(x) = \frac{1}{2} (f_0(x^+) + f_0(x^-))$ ce qui prouve que $f_0 \in E$.

Unicité : Soient $f_0, f_1 \in E$ répondant à la question. Soient $x \in \mathbb{R}$ et η, A comme ci dessus. On a $f_0(y) = f_1(y) = f(y)$ pour tout $y \in A$ et donc $f_0(x^+) = f_1(x^+)$ et $f_0(x^-) = f_1(x^-)$ donc $f_0(x) = f_1(x)$.

L'égalité des intégrales est évidente.

Pour $f, g \in \mathcal{P}$ on notera encore $\langle f | g \rangle = \frac{1}{2\pi} \int_0^{2\pi} \overline{f(t)} g(t) dt$. $\langle | \cdot \rangle$ est, sur \mathcal{P} , une forme sesquilinéaire à symétrie hermitienne positive, mais qui n'est plus définie.

12.1.2 Lemme de Riemann-Lebesgue

Lemme 12.1.3

Soit f de $[a, b]$ dans \mathbb{C} continue par morceaux ; les intégrales $\int_a^b \sin(\lambda t) f(t) dt$ et $\int_a^b \cos(\lambda t) f(t) dt$ tendent vers 0 quand λ réel tend vers l'infini.

Ce théorème est prouvé dans les compléments d'analyse (chapitre intégrales dépendant d'un paramètre, exemple de la transformation de Fourier).

12.1.3 Produit de convolution

DEFINITION 12.1.3

Soient f et g deux fonctions de \mathbb{R} dans \mathbb{C} , 2π -périodiques et continues par morceaux. On appelle produit de convolution de f et g et on note $f * g$ la fonction définie par

$$f * g(x) = \frac{1}{2\pi} \int_0^{2\pi} f(x-t) g(t) dt.$$

On résume dans le lemme ci-dessous les principales propriétés du produit de convolution :

Lemme 12.1.4

Soient $f, g \in \mathcal{P}$.

1) $f * g = g * f$.

- 2) $f * g$ est 2π -périodique.
- 3) L'application $(f, g) \rightarrow f * g$ est bilinéaire.
- 4) $f * g$ est continue sur \mathbb{R} .

preuve

Les points 1, 2 et 3 se vérifient facilement. Montrons la continuité de $f * g$. Fixons f et distinguons plusieurs cas pour g .

1. Si g est continue, pour toute suite (x_n) tendant vers x , la suite de fonctions $h_n(t) := f(t)g(x_n - t)$ converge simplement vers la fonction $h(t) = f(t)g(x - t)$ et la convergence est dominée par la fonction intégrable $\|f\|_\infty \|g\|_\infty \chi_{[0, 2\pi]}$ donc d'après le théorème de convergence dominée $f * g(x_n)$ converge vers $f * g(x)$, d'où la continuité de $f * g$ dans ce cas.
2. Si $g|_{[0, 2\pi]}$ est la fonction caractéristique d'un intervalle (α, β) on a $2\pi f * g(x) = \int_\alpha^\beta f(x - t) dt = \int_{x-\beta}^{x-\alpha} f(u) du$ et la continuité résulte des théorèmes élémentaires sur les intégrales fonctions de leurs bornes.

3. Dans le cas général, la restriction de g à $[0, 2\pi[$ peut s'écrire sous la forme $g_0 + \sum_{k=1}^{k=p} \lambda_k \chi_k$ où g_0 est continue, les λ_k sont des nombres complexes et les χ_k sont des fonctions caractéristiques d'intervalles contenus dans $[0, 2\pi]$. Alors $f * g = f * g_0 + \sum_{k=1}^{k=p} \lambda_k f * \chi_k$ est continue comme somme de fonctions continues.

12.2 Polynômes et séries trigonométriques

Notations

1. Pour $k \in \mathbb{Z}$ on note e_k la fonction de \mathbb{R} dans \mathbb{C} définie par $e_k(t) = e^{ikt}$. e_k est évidemment un élément de E .
2. Pour $k \in \mathbb{N}$, on note $C_k(t) = \cos(kt)$ et $S_k(t) = \sin(kt)$. C_k et S_k sont des éléments de E .
3. On note \mathcal{E} le sous espace vectoriel de E engendré par les fonctions (e_k) , $k \in \mathbb{Z}$ et $\mathcal{E}_n = \text{Vect}(e_{-n}, e_{-(n-1)}, \dots, e_0, \dots, e_{(n-1)}, e_n)$. La suite croissante de sous espaces vectoriels \mathcal{E}_n a pour réunion \mathcal{E} . Par les formules d'Euler, on voit que $\mathcal{E}_n = \text{Vect}(C_0, C_1, \dots, C_n, S_1, \dots, S_n)$. \mathcal{E} est donc aussi le sous espace vectoriel de E engendré par les fonctions C_k , $k \in \mathbb{N}$ et S_k , $k \in \mathbb{N}^*$.

DEFINITION 12.2.1 (Polynômes trigonométriques)

On appelle polynôme trigonométrique toute fonction de \mathcal{E} .

THEOREME 12.2.1

1. $(e_k)_{k \in \mathbb{Z}}$ est un système orthonormé de (E, \langle, \rangle) .
2. De même, $(C_0, \sqrt{2}C_1, \sqrt{2}S_1, \dots, \sqrt{2}C_n, \sqrt{2}S_n, \dots)$ est un système orthonormé dans (E, \langle, \rangle)

La preuve est élémentaire. Il en résulte que ces deux systèmes sont libres, donc que $\dim \mathcal{E}_n = 2n + 1$.

Soit P un polynôme trigonométrique. Si P est constant il appartient à \mathcal{E}_0 ; sinon, il existe un unique entier n tel que $P \in \mathcal{E}_n$ et $P \notin \mathcal{E}_{n-1}$. On peut donc l'écrire

$$P = \sum_{k=-n}^{k=n} c_k e_k = \frac{a_0}{2} + \sum_{k=1}^{k=n} (a_k C_k + b_k S_k) \text{ soit encore } P(t) = \sum_{k=-n}^{k=n} c_k e^{ikt} = \frac{a_0}{2} + \sum_{k=1}^{k=n} (a_k \cos(kt) + b_k \sin(kt))$$

avec

$$(\mathcal{R}) \quad \begin{cases} c_0 &= a_0/2 \\ c_k &= 1/2(a_k - ib_k), \quad k \geq 1 \\ c_{-k} &= 1/2(a_k + ib_k), \quad k \geq 1 \end{cases} \Leftrightarrow \begin{cases} a_k &= c_k + c_{-k}, \quad k \geq 0 \\ b_k &= i(c_k - c_{-k}), \quad k \geq 1 \end{cases}$$

Puisque $(e_j)_{-n \leq j \leq n}$ est une base orthonormée de \mathcal{E}_n on a

$$\forall k \in \mathbb{Z} \quad c_k = \langle e_k | P \rangle = \frac{1}{2\pi} \int_0^{2\pi} e^{-ikt} P(t) dt$$

et de même,

$$\forall k \in \mathbb{N} \quad a_k = \frac{1}{\pi} \int_0^{2\pi} P(t) \cos(kt) dt \quad \text{et} \quad \forall k \in \mathbb{N}^* \quad b_k = \frac{1}{\pi} \int_0^{2\pi} P(t) \sin(kt) dt$$

Attention! a_n et b_n sont en général complexes; ce ne sont donc pas en général les parties réelles et imaginaires de $2c_n$.

DEFINITION 12.2.2 (Séries trigonométriques)

On appellera série trigonométrique une série de fonctions de \mathbb{R} dans \mathbb{C} dont le terme général u_n est, pour $n \geq 1$, de la forme

$$u_n = a_n \cos(nt) + b_n \sin(nt)$$

le terme d'indice 0 étant constant et écrit $u_0(t) = a_0/2$.

Remarque

La somme partielle d'ordre n d'une telle série s'écrit

$$S_n(t) = \sum_{k=0}^{k=n} u_n(t) = \sum_{k=-n}^{k=n} c_k e^{ikt}$$

où les c_k sont définis par les relations (\mathcal{R}).

La convergence de la série trigonométrique signifie l'existence d'une limite à la suite $S_n(t)$ et non bien entendu la convergence séparée des séries $\sum_{k \geq 0} c_k e^{ikt}$ et $\sum_{k < 0} c_k e^{ikt}$ (c.f. 11.4)

12.3 Coefficients et série de Fourier

DEFINITION 12.3.1

Soit $f \in \mathcal{P}$, on appelle coefficient de Fourier (exponentiel) de f , d'indice $n \in \mathbb{Z}$ le nombre complexe

$$c_n(f) = \langle e_n, f \rangle = \frac{1}{2\pi} \int_0^{2\pi} f(t) e^{-int} dt.$$

On appelle coefficients de Fourier (trigonométriques) de f , d'indice $n \in \mathbb{N}$ les nombres

$$a_n(f) = c_n(f) + c_{-n}(f) = \frac{1}{\pi} \int_0^{2\pi} f(t) \cos(nt) dt$$

et pour $n \geq 1$

$$b_n(f) = i(c_n(f) - c_{-n}(f)) = \frac{1}{\pi} \int_0^{2\pi} f(t) \sin(nt) dt$$

On appelle série de Fourier de f la série trigonométrique associée à la suite ci dessus : les sommes partielles de cette série sont donc

$$S_n(f)(t) = \sum_{k=-n}^{k=n} c_k(f) e^{ikt} = \frac{a_0(f)}{2} + \sum_{k=1}^{k=n} (a_k(f) \cos(kt) + b_k(f) \sin(kt))$$

Remarques

(1) On a pour tout $n \in \mathbb{N}$

$$a_n(f) = 2 \langle C_n, f \rangle \text{ et } b_n(f) = 2 \langle S_n, f \rangle.$$

(2) Si la fonction f est paire on a pour tout $n \in \mathbb{N}$ $b_n(f) = 0$ et $a_n(f) = \frac{2}{\pi} \int_0^\pi f(t) \cos(nt) dt$.

(3) Si la fonction f est impaire on a pour tout $n \in \mathbb{N}$ $a_n(f) = 0$ et $b_n(f) = \frac{2}{\pi} \int_0^\pi f(t) \sin(nt) dt$.

(4) Si la fonction f est à valeurs réelles, les coefficients trigonométriques sont réels et on a

$$c_{-n}(f) = \overline{c_n(f)} \quad a_n(f) = 2\operatorname{Re}(c_n(f)) \text{ et } b_n(f) = -2\operatorname{Im}(c_n(f))$$

(5) Soit $f \in \mathcal{P}$ et $f_0 \in E$ sa fonction associée (lemme 12.1.2). Ces deux fonctions ont les mêmes coefficients de Fourier et donc la même série de Fourier.

(6) Il résulte du lemme de Riemann-Lebesgue que pour toute $f \in \mathcal{P}$, on a $\lim_{|n| \rightarrow \infty} c_n(f) = 0$.

EXEMPLE 12.3.1

Soit $f : \mathbb{R} \rightarrow \mathbb{R}$ la fonction 2π périodique telle que $f(x) = \frac{\pi-x}{2}$ pour $0 < x < 2\pi$ et $f(0) = 0$. Il est immédiat que $f \in E$.

D'autre part, f est impaire car si $x \in]-2\pi, 0[$, on a $f(x) = f(x+2\pi) = \frac{\pi-(x+2\pi)}{2} = \frac{-x-\pi}{2} = -f(-x)$. Donc $a_n(f) = 0$ pour tout $n \in \mathbb{N}$.

Ensuite pour $n \in \mathbb{N}^*$, $b_n(f) = \frac{1}{\pi} \int_0^\pi (\pi-x) \sin(nx) dx = -\frac{\pi-x}{n\pi} \cos(nx) \Big|_0^\pi - \frac{1}{n\pi} \int_0^\pi \cos(nx) dx = \frac{1}{n}$. La série de Fourier

de f est donc la série $\sum_{n=1}^{\infty} \frac{\sin(nx)}{n}$.

12.4 Densité des polynômes trigonométriques

12.4.1 Théorème de densité

Le but de cette section est de prouver le théorème suivant

THEOREME 12.4.1

Soit E_0 le sous espace vectoriel de E formé des fonctions continues. L'espace \mathcal{E} des polynômes trigonométriques est dense dans l'espace E_0 muni de la norme de la convergence uniforme.

Ce théorème est une conséquence immédiate des deux propositions suivantes :

PROPOSITION 12.4.1

Soit (P_n) une suite de polynômes trigonométriques vérifiant les propriétés suivantes :

$$\begin{aligned} \text{(P1)} \quad & \forall n \in \mathbb{N}, \forall t \in \mathbb{R}, P_n(t) \geq 0 \\ \text{(P2)} \quad & \forall n \in \mathbb{N}, \int_0^{2\pi} P_n(t) dt = 1 \\ \text{(P3)} \quad & \forall \delta \in]0, \pi[, \lim_{n \rightarrow \infty} \int_{\delta}^{\pi} P_n(t) dt = \lim_{n \rightarrow \infty} \int_{-\pi}^{-\delta} P_n(t) dt = 0 \end{aligned}$$

Pour toute fonction continue et 2π -périodique $f : \mathbb{R} \rightarrow \mathbb{C}$ la suite $f * P_n$ est une suite de polynômes trigonométriques convergeant uniformément vers f .

PROPOSITION 12.4.2

Il existe des suites de polynômes trigonométriques vérifiant les trois propriétés (P1), (P2) et (P3).

Preuve de la proposition 12.4.1

Soit $f : \mathbb{R} \rightarrow \mathbb{C}$ continue et 2π -périodique.

1. Soit P un polynôme trigonométrique. La fonction $f * P$ est un polynôme trigonométrique.

En effet, si $P(t) = \sum_{k=-n}^{k=n} c_k e^{ikt}$ on a $f * P(x) = \frac{1}{2\pi} \int_0^{2\pi} f(t) P(x-t) dt = \frac{1}{2\pi} \sum_{k=-n}^{k=n} \left(c_k \int_0^{2\pi} f(t) e^{-ikt} dt \right) e^{ikx}$ Donc, les $f * P_n$ sont des polynômes trigonométriques.

2. La suite $f * P_n$ converge uniformément vers f .

Soit $\varepsilon > 0$. La fonction f est continue sur \mathbb{R} et 2π périodique. Elle est donc bornée sur \mathbb{R} et uniformément continue sur \mathbb{R} . Il existe donc un $\delta > 0$ que l'on peut choisir strictement inférieur à π tel que

$$\forall u, v \in \mathbb{R} \quad |v - u| \leq \delta \Rightarrow |f(v) - f(u)| \leq \frac{\varepsilon}{2}$$

Fixons un tel δ . Compte tenu de la propriété (P2), on peut écrire

$$f(x) - f * P_n(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} (f(x) - f(x-t)) P_n(t) dt = I_1(n) + I_2(n) + I_3(n)$$

en posant

$$I_1(n) = \frac{1}{2\pi} \int_{-\delta}^{\delta} (f(x) - f(x-t)) P_n(t) dt$$

$$I_2(n) = \frac{1}{2\pi} \int_{\delta}^{\pi} (f(x) - f(x-t)) P_n(t) dt \quad \text{et} \quad I_3(n) = \frac{1}{2\pi} \int_{-\pi}^{-\delta} (f(x) - f(x-t)) P_n(t) dt$$

- Pour $|t| \leq \delta$ on a $|f(x) - f(x-t)| \leq \varepsilon/2$ donc $|I_1(n)| \leq \frac{1}{2\pi} \frac{\varepsilon}{2} \int_{-\delta}^{\delta} P_n(t) dt$. Or P_n est une fonction positive, donc $\int_{-\delta}^{\delta} P_n(t) dt \leq \int_{-\pi}^{\pi} P_n(t) dt = 1$. Donc, pour tout n , $|I_1(n)| \leq \varepsilon/2$.
- D'autre part, on a $|I_2(n)| \leq \frac{1}{2\pi} 2 \|f\|_{\infty} \int_{\delta}^{\pi} P_n(t) dt$ et de même $|I_3(n)| \leq \frac{1}{2\pi} 2 \|f\|_{\infty} \int_{-\pi}^{-\delta} P_n(t) dt$. D'après la propriété (P3), il existe un entier N tel que pour tout $n \geq N$ on ait $\int_{\delta}^{\pi} P_n(t) dt \leq \frac{\varepsilon\pi}{4 \|f\|_{\infty}}$ et $\int_{-\pi}^{-\delta} P_n(t) dt \leq \frac{\varepsilon\pi}{4 \|f\|_{\infty}}$. Cet entier N est indépendant de x et pour tout $n \geq N$ on a $|I_2(n)| + |I_3(n)| \leq \varepsilon/2$.

- Alors, pour tout $n \geq N$ et tout $x \in \mathbb{R}$, on a $|f(x) - f * P_n(x)| \leq \varepsilon$.

Ce qui achève la preuve.

Preuve de la proposition 12.4.2

Posons pour tout entier naturel n , $h_n(x) = \cos^{2n}\left(\frac{x}{2}\right)$, $a_n = \int_{-\pi}^{\pi} h_n(x) dx$ et $P_n(x) = \frac{2\pi}{a_n} h_n(x)$. La suite (P_n) répond à la question.

En effet, par construction P_n vérifie $\frac{1}{2\pi} \int_{-\pi}^{\pi} P_n(x) dx = 1$ et $P_n(x) \geq 0$ pour tout x . D'autre part, $\cos^2(x/2) = \frac{2 + e^{ix} + e^{-ix}}{4}$ donc P_n est un polynôme trigonométrique.

Il reste à vérifier la propriété (P3).

Pour $\delta \leq |t| \leq \pi$ on a $|P_n(x)| \leq b_n$ où $b_n = \frac{2\pi}{a_n} \cos^{2n}(\delta/2)$. Le changement de variable $x = 2t$ donne $a_n = 2 \int_{-\pi}^{\pi} \cos^{2n}(t) dt = 16 \int_0^{\pi/2} \cos^{2n} t dt$. Cette dernière intégrale est une intégrale de Wallis. Un exercice classique montre qu'elle est équivalente à $\sqrt{\frac{\pi}{2n}}$ donc $b_n \sim A\sqrt{n} \cos^{2n}(\delta/2)$ où $A > 0$ est une constante. Donc $\lim_{n \rightarrow \infty} b_n = 0$ d'où l'on déduit (P3).

12.4.2 Théorème de Fejer

Nous allons donner une autre preuve de la proposition 12.4.2 qui fournira un résultat plus précis que le théorème 12.4.1

DEFINITION 12.4.1

Soit $f \in \mathcal{P}$. On appelle sommes de Féjer la suite des moyennes arithmétiques des sommes partielles de la série de Fourier de f , c'est à dire la suite de fonctions définie par

$$\Sigma_n(f)(x) = \frac{1}{n+1} (S_0(f)(x) + S_1(f)(x) + \dots + S_n(f)(x))$$

où $S_n(x) = \sum_{k=-n}^{k=n} c_k(f) e^{ikx}$ est la somme partielle d'indice n de la série de Fourier de f .

THEOREME 12.4.2 (Théorème de Féjer)

Soit $f : \mathbb{R} \rightarrow \mathbb{C}$ une fonction 2π périodique et **continue**. La suite des sommes de Féjer de f converge uniformément vers f sur \mathbb{R} .

Noyau de Dirichlet et Noyau de Fejer

DEFINITION 12.4.2

1) On appelle noyau de Dirichlet la suite $(D_n)_{n \geq 0}$ de polynômes trigonométriques définie par

$$D_n(t) = \sum_{k=-n}^{k=n} e^{ikt}$$

2) On appelle noyau de Fejer la suite des moyennes arithmétiques des noyaux de Dirichlet, c'est à dire la suite de polynômes trigonométriques $(K_n)_{n \geq 0}$ définie par

$$K_n(t) = \frac{1}{n+1} (D_0(t) + \dots + D_n(t))$$

L'introduction de ces suites de polynômes trigonométriques est justifiée par le

Lemme 12.4.1

Soit $f \in \mathcal{E}$. On a pour tout $n \in \mathbb{N}$, $S_n(f) = D_n * f$ et $\Sigma_n(f) = K_n * f$.

preuve

$$S_n(f)(x) = \sum_{k=-n}^{k=n} c_k(f) e^{ikx} = \frac{1}{2\pi} \sum_{k=-n}^{k=n} \int_0^{2\pi} f(t) e^{ik(x-t)} dt = \frac{1}{2\pi} \int_0^{2\pi} f(t) D_n(x-t) dt = f * D_n(x)$$

La deuxième égalité résulte de la linéarité du produit de convolution.

Calcul des noyaux

PROPOSITION 12.4.3

On a

$$D_n(t) = 1 + 2 \sum_{k=1}^{k=n} \cos(kt) = \begin{cases} \frac{\sin\left(\left(n + \frac{1}{2}\right)t\right)}{\sin\left(\frac{t}{2}\right)} & \text{si } t \notin 2\pi\mathbb{Z} \\ 2n + 1 & \text{si } t \in 2\pi\mathbb{Z} \end{cases}$$

preuve

La première relation définissant $D_n(t)$ résulte des formules d'Euler. Supposons $t \not\equiv 0 \pmod{2\pi}$; alors $e^{it} \neq 1$ donc

$$D_n(t) = \sum_{k=-n}^{k=n} e^{ikt} = \frac{e^{-nit} - e^{(n+1)it}}{1 - e^{it}}$$

d'où en mettant en facteur $e^{it/2}$ en haut et en bas

$$D_n(t) = \frac{e^{-(n+1/2)it} - e^{(n+1/2)it}}{e^{-it/2} - e^{it/2}} = \frac{\sin\left(\left(n + \frac{1}{2}\right)t\right)}{\sin\left(\frac{t}{2}\right)}$$

Le cas $t \equiv 0 \pmod{2\pi}$ est immédiat. On peut aussi retrouver le résultat en prenant la limite quand t tend vers 0 dans la formule précédente, puisque D_n est une fonction continue.

PROPOSITION 12.4.4

$$K_n(t) = \begin{cases} \frac{1}{n+1} \frac{\sin^2\left(\frac{n+1}{2}t\right)}{\sin^2\left(\frac{t}{2}\right)} & \text{si } t \notin 2\pi\mathbb{Z} \\ n+1 & \text{si } t \in 2\pi\mathbb{Z} \end{cases}$$

preuve

Supposons $t \not\equiv 0 \pmod{2\pi}$; on remplace $D_k(t)$ par $\frac{e^{-kit} - e^{(k+1)it}}{1 - e^{it}}$. Il vient

$$(n+1)(1 - e^{it})K_n(t) = \sum_{k=0}^{k=n} e^{-ikt} - \sum_{k=0}^{k=n} e^{i(k+1)t}$$

donc

$$(n+1)(1 - e^{it})K_n(t) = \frac{1 - e^{-(n+1)it}}{1 - e^{-it}} - \frac{e^{it} - e^{(n+2)it}}{1 - e^{it}}$$

On multiplie le numérateur et le dénominateur de la dernière fraction par e^{-it} . Il vient

$$K_n(t) = \frac{1}{n+1} \frac{1}{(1 - e^{it})(1 - e^{-it})} \left(2 - e^{(n+1)it} - e^{-(n+1)it}\right) = \frac{1}{n+1} \frac{2(1 - \cos((n+1)t))}{2(1 - \cos(t))} = \frac{1}{n+1} \frac{\sin^2\left(\frac{n+1}{2}t\right)}{\sin^2\left(\frac{t}{2}\right)}$$

Le cas $t \equiv 0 \pmod{2\pi}$ s'obtient par exemple en prenant la limite quand t tend vers 0 dans la formule précédente.

PROPOSITION 12.4.5

La suite (K_n) possède les propriétés (P1), (P2) et (P3).

preuve

(P1) La formule précédente montre que $K_n(t) \geq 0$ pour tout n et tout $t \in \mathbb{R}$.

(P2) Comme $\int_0^{2\pi} e^{ikt} dt = 0$ si $k \neq 0$ on a $\int_0^{2\pi} D_n(t) dt = 1$ pour tout n , donc aussi $\int_0^{2\pi} K_n(t) dt = 1$.

(P3) Si $\delta \leq |t| \leq \pi$ on a $\sin^2(t/2) \geq \sin^2(\delta/2)$, donc $0 \leq K_n(t) \leq \frac{1}{(n+1)\sin^2(\delta/2)}$ d'où le résultat.

Le théorème de Féjer est alors une conséquence de la proposition 12.4.1

12.5 Convergence en moyenne quadratique

12.5.1 Meilleure approximation quadratique

THEOREME 12.5.1

Soit $f \in E$. Pour tout polynôme trigonométrique $P \in \mathcal{E}_n$ on a $\|f - S_n(f)\|_2 \leq \|f - P\|_2$ avec égalité si et seulement si $P = S_n(f)$.

Lemme 12.5.1 (Inégalité de Bessel)

Soit $f \in \mathcal{P}$ on a

$$\sum_{k=-n}^{k=n} |c_k(f)|^2 \leq \|f\|_2^2 = \frac{1}{2\pi} \int_0^{2\pi} |f(t)|^2 dt.$$

Ces deux résultats sont des conséquences immédiates du théorème 11.4.1 et de sa preuve. Rappelons en brièvement les démonstrations :

$S_n(f)$ est la projection orthogonale de f sur \mathcal{E}_n donc $\|f - P\|_2^2 = \|f - S_n(f)\|_2^2 + \|S_n(f) - P\|_2^2$ car $f - S_n(f)$ est orthogonal à $S_n(f) - P$ d'où le théorème; pour l'inégalité de Bessel, on peut supposer que $f \in E$ car les deux membres de l'inégalité ne changent pas de valeur si on remplace f par sa fonction associée f_0 . On a $\|S_n(f)\|_2^2 = \|f\|_2^2 - \|f - S_n(f)\|_2^2 \leq \|f\|_2^2$. La suite croissante $\left(\sum_{k=-n}^{k=n} |c_k(f)|^2 \right)_{n \geq 0}$ est majorée, donc elle converge. D'où la conclusion.

Remarque

La convergence de cette série permet de retrouver, sans utiliser le lemme de Riemann Lebesgue, le fait que $c_n(f)$ tend vers 0 quand n tend vers $\pm\infty$.

12.5.2 Un théorème de densité

THEOREME 12.5.2

Le sous espace \mathcal{E} des polynômes trigonométriques est dense dans l'espace préhilbertien $(E, \langle \cdot | \cdot \rangle)$

Soit $f \in E$. Il s'agit de montrer que pour tout $\varepsilon > 0$ il existe un polynôme trigonométrique $P \in \mathcal{E}$ tel que $\|f - P\|_2 \leq \varepsilon$.

preuve

Commençons par montrer que pour $f \in E$ et tout $\varepsilon > 0$ il existe $\Psi \in E$ continue telle que $\|f - \Psi\|_2 \leq \varepsilon/2$.

- Soit d'abord $0 \leq a < b \leq 2\pi$ et $\varphi : [a, b] \rightarrow \mathbb{C}$ continue. Montrons que pour tout $\varepsilon > 0$ il existe $\psi : [a, b] \rightarrow \mathbb{C}$ continue telle que $\psi(a) = \psi(b) = 0$ et $\frac{1}{2\pi} \int_a^b |\varphi(x) - \psi(x)|^2 dx \leq \varepsilon^2$. Soit $h > 0$ vérifiant $2h < b - a$. Soit $\psi : [a, b] \rightarrow \mathbb{C}$ continue telle que $\psi(a) = \psi(b) = 0$, égale à φ sur l'intervalle $[a + h, b - h]$ et affine sur chacun des deux intervalles $[a, a + h]$ et $[b - h, b]$. Soit $x \in [a, a + h]$. On a $\psi(x) = \frac{\varphi(a+h)}{h}(x - a)$ donc $|\psi(x)| \leq \|\varphi\|_\infty$ et par conséquent $|\varphi(x) - \psi(x)| \leq 2\|\varphi\|_\infty$. La même majoration est valable sur $[b - h, b]$. On en déduit

$$\frac{1}{2\pi} \int_a^b |\varphi(x) - \psi(x)|^2 dx = \frac{1}{2\pi} \int_a^{a+h} |\varphi(x) - \psi(x)|^2 dx + \frac{1}{2\pi} \int_{b-h}^b |\varphi(x) - \psi(x)|^2 dx \leq \frac{4}{\pi} \|\varphi\|_\infty^2 h$$

qui peut être rendu inférieur à ε^2 par choix de h .

- Soit $f \in E$; il existe une subdivision $x_0 = 0 < x_1 < \dots < x_p = 2\pi$ et pour $0 \leq k \leq p-1$ des fonctions continues $f_k : [x_k, x_{k+1}] \rightarrow \mathbb{C}$ telles que $f(x) = f_k(x)$ pour $x_k < x < x_{k+1}$. Soit $\varepsilon > 0$; on dispose pour chaque k d'une fonction $\psi_k : [x_k, x_{k+1}] \rightarrow \mathbb{C}$ continue, vérifiant $\psi_k(x_k) = \psi_k(x_{k+1}) = 0$ et $\frac{1}{2\pi} \int_{x_k}^{x_{k+1}} |f_k(x) - \psi_k(x)|^2 dx \leq \frac{\varepsilon^2}{4p}$.

Notons Ψ_k la fonction $[0, 2\pi] \rightarrow \mathbb{C}$ égale à ψ_k sur $[x_k, x_{k+1}]$ et à 0 ailleurs. Soit $\Psi = \sum_{k=0}^{p-1} \Psi_k$. C'est une fonction continue sur $[0, 2\pi]$. Comme $\Psi(0) = \Psi(2\pi) = 0$, Ψ se prolonge en une fonction 2π -périodique et continue, qui vérifie

$$\|f - \Psi\|_2^2 = \frac{1}{2\pi} \sum_{k=0}^{p-1} \int_{x_k}^{x_{k+1}} |f_k(x) - \psi_k(x)|^2 dx \leq \frac{\varepsilon^2}{4}.$$

La fonction Ψ étant continue, il existe un polynôme trigonométrique P vérifiant $\|\Psi - P\|_\infty \leq \varepsilon/2$ et donc $\|\Psi - P\|_2 \leq \varepsilon/2$. On a donc $\|f - P\|_2 \leq \varepsilon$. ■

12.5.3 Théorème de Parseval

Les résultats qui suivent sont des conséquences du théorème 11.4.3 vu lors de l'étude des espaces préhilbertiens et de la densité de \mathcal{E} dans l'espace préhilbertien E .

THEOREME 12.5.3

Soit $f \in E$. La série de Fourier de f converge en moyenne quadratique vers f . Autrement dit $\lim_{n \rightarrow \infty} \|f - S_n(f)\|_2 = 0$.

Rappelons la preuve :

Soit $\varepsilon > 0$. D'après le théorème précédent, il existe un entier naturel N et un polynôme trigonométrique $P \in \mathcal{E}_N$ tel que $\|f - P\|_2 \leq \varepsilon$; pour tout $n \geq N$ on a $P \in \mathcal{E}_n$, donc d'après le théorème de meilleure approximation quadratique, $\|f - S_n(f)\|_2 \leq \|f - P\|_2 \leq \varepsilon$. ■

THEOREME 12.5.4 (Formule de Parseval)

Soit $f \in \mathcal{P}$. On a

$$\sum_{-\infty}^{+\infty} |c_n(f)|^2 = \frac{1}{2\pi} \int_0^{2\pi} |f(t)|^2 dt$$

et

$$\frac{1}{2} |a_0(f)|^2 + \sum_{n=1}^{\infty} (|a_n(f)|^2 + |b_n(f)|^2) = \frac{1}{\pi} \int_0^{2\pi} |f(t)|^2 dt$$

preuve

Les deux membres des égalités ci dessus ne changent pas de valeurs si on remplace f par sa fonction associée $f_0 \in E$. On peut donc supposer $f \in E$. Dans ce cas, on peut appliquer le théorème 11.4.3. Rappelons la preuve :

On a $\sum_{k=-n}^{k=n} |c_k(f)|^2 = \|S_n(f)\|_2^2 = \|f\|_2^2 - \|f - S_n(f)\|_2^2$. Le résultat découle du théorème précédent.

Exemple

L'application de ce théorème à la fonction considérée à l'exemple 12.3.1 donne

$$\sum_{n \geq 1} \frac{1}{n^2} = \frac{1}{\pi} \int_0^{2\pi} \frac{(\pi - x)^2}{4} dx = \frac{\pi^2}{6}$$

COROLLAIRE 12.5.1

Pour $f, g \in E$, la série de terme général $\overline{c_n(f)} c_n(g)$ est absolument convergente et

$$\sum_{-\infty}^{+\infty} \overline{c_n(f)} c_n(g) = \frac{1}{2\pi} \int_0^{2\pi} \overline{f(t)} g(t) dt$$

Résulte du théorème 11.4.3

COROLLAIRE 12.5.2

Soient $f, g \in E$ telles que $\forall n \in \mathbb{Z} \ c_n(f) = c_n(g)$. Alors $f = g$.

preuve

La fonction $f - g$ est dans E et a tous ses coefficients de Fourier nuls. D'après l'égalité de Parseval sa norme est nulle, donc $f = g$.

Remarque

Si f, g sont dans \mathcal{P} et si $c_n(f) = c_n(g)$ pour tout n , on peut seulement conclure que $f(x) = g(x)$ en tout point $x \in [0, 2\pi]$ sauf en un nombre fini (éventuellement nul) de points.

12.6 Etude de la convergence uniforme

12.6.1 Convergence uniforme des séries trigonométriques

THEOREME 12.6.1

Soit $(\gamma_n)_{n \in \mathbb{Z}}$ une suite complexe telle que la suite $S_n = \sum_{k=-n}^{k=n} \gamma_k e^{ikx}$ soit uniformément convergente. Alors la limite de cette suite est une fonction continue et 2π périodique f et on a pour tout $n \in \mathbb{Z} \ c_n(f) = \gamma_n$.

En termes de coefficients trigonométriques, l'énoncé devient :

Soit $\alpha_0/2 + \sum_{k=1}^{\infty} (\alpha_k \cos(kx) + \beta_k \sin(kx))$ une série trigonométrique uniformément convergente. Sa somme est une fonction continue f dont les coefficients de Fourier trigonométriques sont les α_k et les β_k .

Remarquons que les hypothèses du théorème sont vérifiées si la série indexée par \mathbb{Z} , $\sum_{-\infty}^{+\infty} |\gamma_n|$ est convergente (resp. si la série

$\sum_{n=1}^{\infty} (|\alpha_n| + |\beta_n|)$ est convergente).

preuve

S_n est continue et converge uniformément vers f , donc f est continue. Soit $k \in \mathbb{Z}$. On a $|f(t)e^{-ikt} - S_n(t)e^{-ikt}| = |f(t) - S_n(t)| \leq \|f - S_n\|_{\infty}$ donc la suite de fonction $(t \rightarrow S_n(t)e^{-ikt})_{n \geq 0}$ converge uniformément sur $[0, 2\pi]$ vers la fonction $t \rightarrow f(t)e^{-ikt}$. Il en résulte que $c_k(f) = \lim_{n \rightarrow \infty} \frac{1}{2\pi} \int_0^{2\pi} S_n(t)e^{-ikt} dt$. Or dès que $n \geq |k|$, cette dernière intégrale est égale à γ_k . Donc $c_k(f) = \gamma_k$.

COROLLAIRE 12.6.1

Soit $f \in E$. Si la série de Fourier de f est uniformément convergente alors f est continue et égale en tout point à la somme de sa série de Fourier.

preuve

En effet, soit F la somme de la série de Fourier de f . La convergence uniforme implique que F est continue et que la série de Fourier de F est égale à celle de f d'après le théorème précédent. Il résulte du corollaire 12.5.2 que $F = f$.

12.6.2 Estimations de coefficients de Fourier

DEFINITION 12.6.1

Soit f une fonction 2π -périodique, continue et C^1 par morceaux. Par définition, f admet en tout point une dérivée à droite f'_d et une dérivée à gauche f'_g et en tout point $x \in [0, 2\pi]$ sauf au plus un nombre fini, f est dérivable et on a $f'_d(x) = f'_g(x)$. On notera dans la suite f' la fonction égale en tout point x à $\frac{1}{2} (f'_g(x) + f'_d(x))$. f' coïncide donc avec la dérivée de f là où elle existe et est une fonction appartenant à E .

Lemme 12.6.1 (Intégration par parties)

Soient $f, g : [a, b] \rightarrow \mathbb{C}$ deux fonctions avec f continue et de classe C^1 par morceaux et g de classe C^1 . On a

$$\int_a^b f'(t)g(t)dt = f(b)g(b) - f(a)g(a) - \int_a^b f(t)g'(t)dt$$

preuve

Il existe une subdivision $(x_0 = a, x_1, \dots, x_n = b)$ de l'intervalle $[a, b]$ et pour $i \in \{0, 1, \dots, n-1\}$ des fonctions de classe C^1 , $f_i : [x_i, x_{i+1}] \rightarrow \mathbb{C}$ telles que $\forall x \in]x_i, x_{i+1}[\quad f(x) = f_i(x)$. On a

$$\int_a^b f'(t)g(t)dt = \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} f'_k(t)g(t)dt = \sum_{k=0}^{n-1} \left(f_k(x_{k+1})g(x_{k+1}) - f_k(x_k)g(x_k) - \int_{x_k}^{x_{k+1}} f_k(t)g'(t)dt \right)$$

Or f est continue, donc $f_k(x_k) = f(x_k)$, $f_k(x_{k+1}) = f(x_{k+1})$ et $f_k(t) = f(t)$ d'où

$$\int_a^b f'(t)g(t)dt = \sum_{k=0}^{n-1} (f(x_{k+1})g(x_{k+1}) - f(x_k)g(x_k)) - \sum_{k=0}^{n-1} \int_{x_k}^{x_{k+1}} f(t)g'(t)dt = f(b)g(b) - f(a)g(a) - \int_a^b f(t)g'(t)dt$$

PROPOSITION 12.6.1

Soit $f : \mathbb{R} \rightarrow \mathbb{C}$ 2π -périodique, continue et C^1 par morceaux. On a

$$\forall n \in \mathbb{Z} \quad c_n(f') = nic_n(f)$$

et

$$\forall n \in \mathbb{N} \quad a_n(f') = nb_n(f) \quad b_n(f') = -na_n(f)$$

preuve

C'est une application du lemme précédent avec $g(t) = e^{-int}$. Il vient :

$$c_n(f') = \frac{1}{2\pi} \int_0^{2\pi} f'(t)e^{-nit}dt = \frac{1}{2\pi} \left(f(2\pi)e^{-2ni\pi} - f(0) + ni \int_0^{2\pi} f(t)e^{-nit}dt \right) = nic_n(f)$$

COROLLAIRE 12.6.2

Soit $f : \mathbb{R} \rightarrow \mathbb{C}$ de classe C^p . On a pour tout k entre 1 et p et tout $n \in \mathbb{Z}$,

$$c_n(f^{(k)}) = (ni)^k c_n(f)$$

Ce résultat subsiste si f est C^{n-1} et si $f^{(n-1)}$ est C^1 par morceaux.

COROLLAIRE 12.6.3

Soit $f : \mathbb{R} \rightarrow \mathbb{C}$ de classe C^p . On a $c_n(f) = o_{n \rightarrow \pm\infty} \left(\frac{1}{|n|^p} \right)$. En particulier, si f est C^∞ , on a $c_n(f) = o \left(\frac{1}{|n|^k} \right)$ pour tout entier naturel k .

12.6.3 Application

THEOREME 12.6.2

Soit $f : \mathbb{R} \rightarrow \mathbb{C}$ 2π -périodique, continue et C^1 par morceaux. La série de Fourier de f converge normalement, donc uniformément sur \mathbb{R} et sa somme est égale à f en tout point.

preuve

Pour $n \neq 0$ on a $|c_n(f)| = \frac{1}{|n|} |c_n(f')| \leq \frac{1}{2} \left(\frac{1}{n^2} + |c_n(f')|^2 \right)$. D'après le théorème de Parseval, la série $\sum_{n \in \mathbb{Z}} |c_n(f')|^2$ converge. On en déduit la convergence de la série $\sum_{n \in \mathbb{Z}} |c_n(f)|$, donc la convergence normale de la série de Fourier de f . On sait dans ce cas que sa somme est f (corollaire 12.6.1). ■

12.7 Étude de la convergence simple

THEOREME 12.7.1 (Théorème de Dirichlet)

Soit f de \mathbb{R} dans \mathbb{C} 2π -périodique et de classe C^1 par morceaux. Alors pour tout $x \in \mathbb{R}$ la série de Fourier de f converge simplement et on a

$$\frac{1}{2}(f(x^+) + f(x^-)) = \lim_{n \rightarrow \infty} \sum_{-n}^n c_k(f)e^{ikx} = \frac{a_0(f)}{2} + \sum_{k=1}^{\infty} (a_k(f) \cos(kx) + b_k(f) \sin(kx)).$$

Si f est dans E on obtient

$$f(x) = \lim_{n \rightarrow \infty} \sum_{-n}^n c_k(f) e^{ikt} = \frac{a_0(f)}{2} + \sum_{k=1}^{\infty} (a_k(f) \cos(kt) + b_k(f) \sin(kt)).$$

preuve

- Rappelons que le noyau de Dirichlet a été défini comme la suite de fonctions $(D_n)_{n \in \mathbb{N}}$ telles que $D_n(t) = \sum_{k=-n}^{k=n} e^{ikt}$ et que $S_n(f)(x) = f * D_n(x)$. Le noyau de Dirichlet vérifie les propriétés suivantes :

$$(D1) \quad \frac{1}{2\pi} \int_0^\pi D_n(t) dt = \frac{1}{2\pi} \int_{-\pi}^0 D_n(t) dt = \frac{1}{2} \quad \text{car } D_n \text{ est paire et } \frac{1}{2\pi} \int_{-\pi}^\pi D_n(t) dt = 1$$

$$(D2) \quad \forall t \in \mathbb{R} \setminus 2\pi\mathbb{Z} \quad D_n(t) = \frac{\sin\left(\left(n + \frac{1}{2}\right)t\right)}{\sin\left(\frac{t}{2}\right)}$$

- On a

$$S_n(f)(x) = f * D_n(x) = \frac{1}{2\pi} \int_{-\pi}^\pi f(x-t) D_n(t) dt = \frac{1}{2\pi} \int_{-\pi}^0 f(x-t) D_n(t) dt + \frac{1}{2\pi} \int_0^\pi f(x-t) D_n(t) dt$$

donc, en utilisant (D1)

$$S_n(f)(x) - \frac{1}{2} (f(x^+) + f(x^-)) = I^+(n, x) + I^-(n, x)$$

en posant

$$I^+(n, x) = \frac{1}{2\pi} \int_{-\pi}^0 (f(x-t) - f(x^+)) D_n(t) dt = \frac{1}{2\pi} \int_{-\pi}^0 h^+(t) \sin((n+1/2)t) dt \quad \text{avec } h^+(t) = \frac{f(x-t) - f(x^+)}{\sin(t/2)}$$

et

$$I^-(n, x) = \frac{1}{2\pi} \int_0^\pi (f(x-t) - f(x^-)) D_n(t) dt = \frac{1}{2\pi} \int_0^\pi h^-(t) \sin((n+1/2)t) dt \quad \text{avec } h^-(t) = \frac{f(x-t) - f(x^-)}{\sin(t/2)}$$

- La fonction h^+ est continue sur $[-\pi, 0[$. Au voisinage de 0, on a $h^+(t) \sim 2 \frac{f(x-t) - f(x^+)}{t}$. Par hypothèse, f est C^1 par morceaux, donc admet en x une dérivée à droite, donc $\lim_{\substack{t \rightarrow 0 \\ t < 0}} h^+(t) = -2f'_d(x)$ et par conséquent h^+ se prolonge en une fonction continue sur $[0, \pi]$. Il résulte alors du lemme de Riemann Lebesgue que $\lim_{n \rightarrow \infty} I^+(n, x) = 0$. On voit de même que $\lim_{n \rightarrow \infty} I^-(n, x) = 0$ ce qui achève la preuve.

Exemple

Reprenons l'exemple 12.3.1. La fonction considérée est clairement C^1 par morceaux et appartient à E . Le théorème de Dirichlet s'applique. On a donc

$$\forall x \in \mathbb{R} \quad f(x) = \sum_{n=1}^{\infty} \frac{\sin(nx)}{n}$$

et en particulier

$$\forall x \in]0, 2\pi[\quad \frac{\pi - x}{2} = \sum_{n=1}^{\infty} \frac{\sin(nx)}{n}$$

Remarque

La fonction f n'est pas continue en 0. Il en résulte que la série de terme général $\frac{\sin(nx)}{n}$ n'est pas uniformément convergente sur $[0, 2\pi]$. Par contre une application classique de la transformation d'Abel permet de montrer que cette série converge uniformément sur tout intervalle de la forme $[\delta, 2\pi - \delta]$ où δ vérifie $0 < \delta < \pi$.

12.8 Cas des fonctions T -périodiques

On indique ci dessous les formules relatives aux séries de Fourier de fonctions T périodiques où T est un réel strictement positif. On notera E_T le \mathbb{C} -espace vectoriel des fonctions T -périodiques, continues par morceaux, vérifiant $f(x) = \frac{1}{2}(f(x^+) + f(x^-))$. On munit E_T du produit scalaire défini par

$$\langle f, g \rangle = \frac{1}{T} \int_0^T \overline{f(t)}g(t)dt.$$

Muni de ce produit scalaire, E_T est un espace préhilbertien complexe. On pose $\omega = 2\pi/T$ (en physique ω s'appelle la pulsation). Dans cet espace le produit de convolution est défini par

$$f * g(x) = \frac{1}{T} \int_0^T f(x-t)g(t)dt$$

On définit enfin les fonctions e_k, C_k, S_k par $e_k(t) = e^{ik\omega t}$, $C_k(t) = \cos(k\omega t)$ et $S_k(t) = \sin(k\omega t)$. Les systèmes $(e_k)_{k \in \mathbb{Z}}$ et $(C_0, \sqrt{2}C_1, \sqrt{2}S_1, \dots, \sqrt{2}C_n, \sqrt{2}S_n, \dots)$ sont orthonormés. Les polynômes trigonométriques sont définis de la même manière. Pour $f \in E_T$, les coefficients de Fourier de f sont définis par

$$c_n(f) = \langle e_n | f \rangle = \frac{1}{T} \int_0^T f(t)e^{-in\omega t}dt, \quad a_n(f) = \frac{2}{T} \int_0^T f(t)\cos(n\omega t)dt, \quad b_n(f) = \frac{2}{T} \int_0^T f(t)\sin(n\omega t)dt$$

On a les mêmes relations (\mathcal{R}) entre ces coefficients. La série de Fourier de f est la série dont les sommes partielles sont

$$S_n(f)(t) = \sum_{k=-n}^{k=n} c_k(f)e^{ik\omega t} = \frac{a_0(f)}{2} + \sum_{k=1}^{k=n} (a_k(f)\cos(k\omega t) + b_k(f)\sin(k\omega t))$$

Si f est C^1 par morceaux, on définit f' de la même manière et on a pour tout $n \in \mathbb{Z}$

$$c_n(f') = ni\omega c_n(f), \quad a_n(f') = n\omega b_n(f), \quad b_n(f') = -n\omega a_n(f)$$

et si f est C^k , on a $c_n(f^{(k)}) = (ni\omega)^k c_n(f)$.

Tous les théorèmes s'étendent. On peut le prouver à partir de ceux démontrés dans le cas $T = 2\pi$ par un changement de variables. Si $f \in E_T$, la fonction f^\sharp définie par $f^\sharp(u) = f(u/\omega)$ est 2π -périodique et a les mêmes propriétés de régularité que f . Les coefficients de Fourier usuels $c_n(f^\sharp)$ de f^\sharp vérifient $c_n(f^\sharp) = c_n(f)$. L'application à f^\sharp des théorèmes des paragraphes précédents donne les résultats correspondants pour f . En particulier la formule de Parseval s'écrit

$$\sum_{-\infty}^{\infty} |c_n(f)|^2 = \frac{1}{T} \int_0^T |f(t)|^2 dt \quad \text{ou} \quad \frac{1}{2}|a_0(f)|^2 + \sum_{n=1}^{\infty} (|a_n(f)|^2 + |b_n(f)|^2) = \frac{2}{T} \int_0^T |f(t)|^2 dt$$

13

Calcul différentiel

13.1 Préliminaires : notations

Dans tout ce chapitre, on utilisera les notations suivantes :

1. E et F sont des \mathbb{R} -espaces vectoriels normés de dimension finie, p et n respectivement. Ω est un ouvert non vide de E et $f : \Omega \rightarrow F$ une application. On résumera tout ceci par la notation $E \supset \Omega \xrightarrow{f} F$.
2. Pour $a \in \Omega$, on désignera par r_a un réel strictement positif tel que $B_E(a, r_a) := \{x \in E \mid d(a, x) = \|x - a\| < r_a\} \subset \Omega$. Si il n'y a pas d'ambiguïté, on notera $B(a, r_a)$ pour $B_E(a, r_a)$.
3. $\mathcal{B}_E = (e_1, \dots, e_p)$ est une base de E , $\mathcal{B}_F = (u_1, \dots, u_n)$ une base de F . Ces bases définissent des isomorphismes $\Phi_{\mathcal{B}_E} : \mathbb{R}^p \rightarrow E$ et $\Phi_{\mathcal{B}_F} : \mathbb{R}^n \rightarrow F$ par $\Phi_{\mathcal{B}_E}(x_1, \dots, x_p) = x_1 e_1 + \dots + x_p e_p$ et $\Phi_{\mathcal{B}_F}(y_1, \dots, y_n) = y_1 u_1 + \dots + y_n u_n$.
4. L'application f est donnée par $f(x) = \sum_{i=1}^n f_i(x) u_i$ où $f_i : \Omega \rightarrow \mathbb{R}$. Les f_i dépendent évidemment de la base \mathcal{B}_F . On les appellera composantes de f .
5. Soit $\tilde{\Omega} = \Phi_{\mathcal{B}_E}^{-1}(\Omega) = \{[x_1, \dots, x_p] \in \mathbb{R}^p \mid x = \sum_{j=1}^p x_j e_j \in \Omega\}$. $\tilde{\Omega}$ est un ouvert de \mathbb{R}^p . On définit $\tilde{f}_i : \tilde{\Omega} \rightarrow \mathbb{R}$ par

$$\tilde{f}_i(x_1, \dots, x_p) = f_i \left(\sum_{j=1}^p x_j e_j \right)$$

Enfin, on notera $\tilde{f} : \tilde{\Omega} \rightarrow \mathbb{R}^n$ l'application

$$(x_1, \dots, x_p) \rightarrow (\tilde{f}_1(x_1, \dots, x_p), \dots, \tilde{f}_n(x_1, \dots, x_p))$$

\tilde{f} est l'expression de f en coordonnées, les bases \mathcal{B}_E et \mathcal{B}_F étant fixées et on a

$$\tilde{f} = \Phi_{\mathcal{B}_F}^{-1} \circ f \circ \Phi_{\mathcal{B}_E} \quad f = \Phi_{\mathcal{B}_F} \circ \tilde{f} \circ \Phi_{\mathcal{B}_E}^{-1}$$

Enfin, si $E = \mathbb{R}^p$ et $F = \mathbb{R}^n$ on choisira toujours, sauf mention explicite du contraire pour \mathcal{B}_E et \mathcal{B}_F les bases canoniques, auquel cas \tilde{f} et f sont identiques.

13.2 Différentiabilité

13.2.1 Dérivée suivant un vecteur

Soient $E \supset \Omega \xrightarrow{f} F$, $e \in E$ et $a \in \Omega$. Si $e \neq 0$, la fonction d'une variable réelle $t \rightarrow f(a + te)$ est définie pour $|t| < r_a / \|e\|$. Sinon elle est définie pour tout t .

DEFINITION 13.2.1

On dit que la fonction f admet au point a une dérivée suivant le vecteur e si la fonction $t \rightarrow f(a + te)$ est dérivable en 0. Dans ce cas, on notera cette dérivée $\partial_e f(a)$.

On a donc, sous réserve d'existence

$$\partial_e f(a) = \lim_{t \rightarrow 0} \frac{f(a + te) - f(a)}{t}$$

Cette dérivée est donc un élément de F .

On notera que toute fonction admet une dérivée suivant le vecteur nul et que cette dérivée est le vecteur nul de F .

\mathcal{B}_F étant fixée, on a $\frac{f(a+te)-f(a)}{t} = \sum_{i=1}^{i=n} \frac{f_i(a+te)-f_i(a)}{t} u_i$ de sorte que f admet une dérivée suivant e ssi chacune des applications coordonnées f_i en admet une, auquel cas on aura

$$\partial_e f(a) = \sum_{i=1}^{i=n} \partial_e f_i(a) u_i$$

Si $a = \sum_{j=1}^{j=p} a_j e_j$ et $e = \sum_{j=1}^{j=p} \alpha_j e_j$ on a

$$\frac{f_i(a + te) - f_i(a)}{t} = \frac{\tilde{f}_i(a_1 + t\alpha_1, \dots, a_p + t\alpha_p) - \tilde{f}_i(a_1, \dots, a_p)}{t}$$

donc f_i admet en a une dérivée selon e ssi \tilde{f}_i admet en (a_1, \dots, a_p) une dérivée selon $(\alpha_1, \dots, \alpha_p)$ et dans ce cas ces dérivées sont égales. On peut donc toujours étudier le problème de l'existence des dérivées de f en coordonnées.

Attention! f peut avoir des dérivées en un point a selon **tous** les vecteurs E sans être continue en a .

Considérons par exemple l'application $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par $f(0, 0) = 0$ et $f(x, y) = \frac{x^2 y}{x^4 + y^2}$ si $(x, y) \neq (0, 0)$.

Si $e = (a, b) \neq (0, 0)$ on a $\delta(t) := \frac{f(te) - f(0)}{t} = \frac{a^2 b}{t^2 a^4 + b^2}$. Si $b = 0$, $\delta(t) = 0$ pour tout t , donc tend vers 0. Si $b \neq 0$, $\lim_{t \rightarrow 0} \delta(t) = a^2/b$. Donc f admet des dérivées selon tout vecteur de \mathbb{R}^2 . Mais f n'est pas continue en $(0, 0)$ car $\lim_{n \rightarrow \infty} f(\frac{1}{n}, \frac{1}{n^2}) = \frac{1}{2} \neq 0 = f(0, 0)$.

Dérivées partielles

La base $\mathcal{B}_E = (e_1, \dots, e_p)$ de base de E étant fixée, soit j un entier, $1 \leq j \leq p$. On appelle dérivée partielle de f suivant la j -ième variable la dérivée, si elle existe, de f selon le j -ième vecteur de la base \mathcal{B}_E . On note cette dérivée partielle $\partial_j f(a)$ ou $\frac{\partial f}{\partial x_j}(a)$ ou encore $f'_{x_j}(a)$.

On prendra garde au fait que cette notion dépend de la base choisie. Dans le cas particulier $E = \mathbb{R}^p$, on prend pour \mathcal{B}_E , la base canonique. On a alors, sous réserve d'existence

$$\frac{\partial f}{\partial x_j}(a) = \partial_{e_j} f(a) = \lim_{t \rightarrow 0} \frac{f(a_1, \dots, a_{j-1}, a_j + t, a_{j+1}, \dots, a_p) - f(a_1, \dots, a_p)}{t}$$

13.2.2 Différentiabilité

Soit $h \in E$ tel que $\|h\| < r_a$. On a alors $a + h \in \Omega$. Ceci permet de poser la définition suivante :

DEFINITION 13.2.2

On dit que f est différentiable en $a \in \Omega$ si il existe une application linéaire $L \in L(E, F)$ telle que

$$\lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{f(a + h) - f(a) - L(h)}{\|h\|} = 0$$

Ceci revient à dire qu'on a un développement limité d'ordre 1

$$f(a+h) = f(a) + L(h) + \|h\|\varepsilon(h) \quad \text{avec} \quad \lim_{h \rightarrow 0} \varepsilon(h) = 0$$

où ε est une fonction définie sur la boule $B_E(0, r_a)$.

PROPOSITION 13.2.1

Si f est différentiable en a , f admet en a une dérivée selon tout vecteur e de E , dérivée donnée par

$$\partial_e f(a) = L(e)$$

preuve

C'est vrai si $e = 0$. Sinon, pour $t \neq 0$

$$\frac{f(a+te) - f(a)}{t} = \pm \frac{f(a+te) - f(a) - L(te)}{\|te\|} \|e\| + L(e)$$

le signe \pm étant $+$ si $t > 0$ et $-$ sinon. Par hypothèse, le premier terme du second membre tend vers 0 quand t tend vers 0. ■

COROLLAIRE 13.2.1

Si f est différentiable en a , l'application linéaire L figurant dans la définition est unique.

En effet, $\mathcal{B}_E = (e_1, \dots, e_p)$ désignant toujours une base de E , on a $L(e_j) = \partial_{e_j} f(a)$.

L s'appelle la différentielle de f en a . Il y a plusieurs notations pour L : $d_a f$, $df(a)$ ou encore $f'(a)$. Dans la suite, nous utiliserons le plus souvent $df(a)$.

DEFINITION 13.2.3

On dit que f est différentiable sur Ω si elle est différentiable en tout point de Ω . Dans ce cas, on dispose d'une application $df : \Omega \rightarrow L(E, F)$ qui à $x \in \Omega$ associe la différentielle de f en x , $df(x)$.

THEOREME 13.2.1

Soit Ω un ouvert de E et $f : \Omega \rightarrow F$. Si f est différentiable en $a \in \Omega$, alors f est continue en a .

preuve

E étant de dimension finie, l'application linéaire $df(a) : E \rightarrow F$ est continue, donc $f(a+h) = f(a) + df(a)(h) + \|h\|\varepsilon(h)$ tend vers $f(a)$ quand h tend vers 0.

PROPOSITION 13.2.2

1. Soient $f, g : \Omega \rightarrow F$ différentiables en $a \in \Omega$ et $\lambda \in \mathbb{R}$. La fonction $f + \lambda g$ est différentiable en a et $d(f + \lambda g)(a) = df(a) + \lambda dg(a)$.

2. L'ensemble des fonctions $\Omega \rightarrow F$ différentiables en a est un sous espace vectoriel de l'espace vectoriel $\mathcal{F}(\Omega, F)$ des fonctions définies sur Ω à valeurs dans F et l'application $f \rightarrow df(a)$ est une application linéaire de cet espace dans $L(E, F)$.

3. L'ensemble des fonctions $\Omega \rightarrow F$ différentiables en tout point de Ω est un sous espace vectoriel de $\mathcal{F}(\Omega, F)$ et l'application $f \rightarrow f' = df$ est une application linéaire de cet espace dans $\mathcal{F}(\Omega, L(E, F))$.

PROPOSITION 13.2.3

Soient $E \supset \Omega \xrightarrow{f} F$ et $a \in \Omega$. f est différentiable en a ssi chacune des applications composantes f_i est différentiable en a et alors, pour tout $h \in E$, on a

$$df(a)(h) = \sum_{i=1}^{i=n} df_i(a)(h)$$

Les vérifications sont faciles.

PROPOSITION 13.2.4

Soit $f : \Omega \rightarrow F$ et $a \in \Omega$.

1. Soient G un espace vectoriel de dimension finie et $\Phi \in L(F, G)$ une application linéaire. Si f est différentiable en a , $\Phi \circ f$ est différentiable en a et $d(\Phi \circ f)(a) = \Phi \circ df(a)$. Si Φ est un isomorphisme, f est différentiable en a ssi $\Phi \circ f$ l'est.

2. Soient H un espace vectoriel de dimension finie et $\psi : H \rightarrow E$ un isomorphisme. L'ensemble $\Omega' = H^{-1}(\Omega)$ est un ouvert de H et f est différentiable en a ssi $f \circ \Psi$ est différentiable en $c = \Psi^{-1}(a)$. On a alors $d(f \circ \Psi)(c) = df(a) \circ \Psi$.

preuve

1. Soit $h \in B_E(0, r_a)$. On a $\Phi \circ f(a+h) = \Phi(f(a) + df(a)(h) + \|h\|\varepsilon(h)) = \Phi \circ f(a) + \Phi \circ df(a)(h) + \|h\|\Phi(\varepsilon(h))$. Or E étant de dimension finie, Φ est continue, donc $\Phi(\varepsilon(h))$ tend vers 0 avec h . Par conséquent $\Phi \circ f$ est différentiable en a et $d(\Phi \circ f)(a) = \Phi \circ df(a)$.

Si Φ est un isomorphisme et si $\Phi \circ f$ est différentiable en a , on applique ce qui précède en remplaçant f par $\Phi \circ f$ et Φ par Φ^{-1} pour en conclure que f est différentiable en a .

2. Ψ est continue. Ω' est ouvert comme image réciproque d'un ouvert par une application continue. Il existe un $r > 0$ tel que $B_H(c, r) \subset \Omega'$ et $f(B_H(c, r)) \subset B(a, r_a)$.

Supposons f différentiable en a . Soit $k \in B_H(0, r)$. On a $f \circ \Psi(c+k) = f(a + \Psi(k)) = f(a) + df(a)(\Psi(k)) + \|\Psi(k)\|\varepsilon(\Psi(k))$.

La continuité de l'application linéaire Ψ fournit une constante C telle que $\|\Psi(k)\| \leq C\|k\|$. Quand k tend vers 0, $\Psi(k)$ aussi, donc $\varepsilon(\Psi(k))$ aussi et $\frac{\|\Psi(k)\|\varepsilon(\Psi(k))}{\|k\|}$ tend vers 0 avec k ce qui prouve que $f \circ \Psi$ est différentiable en c . On montre la réciproque comme dans le 1.

Application

D'après la proposition précédente, l'application $f : \Omega \rightarrow F$ est différentiable sur Ω (resp. en a) ssi l'application $\tilde{f} = \Phi_{B_F}^{-1} \circ f \circ \Phi_{B_E}$ est différentiable sur $\tilde{\Omega}$ (resp. en $\tilde{a} = \Phi_{B_E}^{-1}(a)$) et pour tout $x \in \Omega$ on a, en posant $\tilde{x} = \Phi_{B_E}^{-1}(x)$,

$$df(x) = \Phi_{B_F} \circ d\tilde{f}(\tilde{x}) \circ \Phi_{B_E}^{-1}$$

En clair, une fois choisie les bases B_E et B_F , si \tilde{f} est l'expression en coordonnées de f , f est différentiable en a ssi \tilde{f} l'est en $\tilde{a} = (a_1, \dots, a_p)$ et l'écriture en coordonnées de $df(a) : E \rightarrow F$ est $d\tilde{f}(\tilde{a})$.

Ce qui précède est une justification précise que l'on peut faire tous les calculs dans des bases et se ramener ainsi à des applications $\mathbb{R}^p \rightarrow \mathbb{R}^n$.

Remarque importante

Les notions de différentiabilité comme de dérivée selon un vecteur sont définies en termes de limites. Comme toutes les normes sur E et sur F sont équivalentes, il est facile de voir que ces notions sont indépendantes des normes choisies. Autrement dit, si f est différentiable en a pour les normes N_E et N_F sur E et F , elle l'est encore, et avec la même différentielle si on munit E et F d'autres normes N'_E et N'_F .

13.2.3 Exemples

EXEMPLE 13.2.1 (Application constante)

Si $f : \Omega \rightarrow F$ est constante, f est différentiable en tout $a \in \Omega$ et $df(a) = 0$.

EXEMPLE 13.2.2 (Application affine)

Soit $\varphi \in L(E, F)$ et $f : \Omega \rightarrow F$ une application affine d'application linéaire associée φ . f est différentiable en tout point $a \in \Omega$ et $df(a) = \varphi$.

En effet, $f(a+h) - f(a) = \varphi(h)$ par définition d'une application affine.

Notons que dans ce cas, l'application $df : \Omega \rightarrow L(E, F)$ est constante.

EXEMPLE 13.2.3 (Cas des applications $\mathbb{R} \rightarrow F$)

Soient Ω un ouvert de \mathbb{R} , $f : \Omega \rightarrow F$ et $a \in \Omega$. f est différentiable en $a \Leftrightarrow f$ est dérivable en a . Alors, si $f'(a) \in F$ est la dérivée usuelle de f , on a, pour $h \in \mathbb{R}$, $df(a)(h) = hf'(a)$.

L'application $L(\mathbb{R}, F) \rightarrow F$ qui à $L \in L(\mathbb{R}, F)$ associe $L(1) \in F$ est un isomorphisme et alors $L(h) = hL(1)$ pour tout

$h \in \mathbb{R}$. Ceci étant, on a

$$\begin{aligned} f \text{ différentiable en } a &\Leftrightarrow \\ \exists L \in L(\mathbb{R}, F), f(a+h) = f(a) + L(h) + o(|h|) &\Leftrightarrow \\ \exists \lambda \in F, f(a+h) = f(a) + h\lambda + o(h) &\Leftrightarrow \\ f \text{ est dérivable en } a \text{ et } f'(a) = \lambda & \end{aligned}$$

L'équivalence annoncée en résulte. On voit que l'identification entre $L(\mathbb{R}, F)$ et F identifie $df(a) \in L(\mathbb{R}, F)$ et $f'(a) \in F$. Ceci justifie que l'on note souvent $f'(a)$ pour la différentielle de f en a dans le cas général.

EXEMPLE 13.2.4 (Application bilinéaire)

Soient E_1, E_2 et F trois espaces de dimensions finie et $B : E_1 \times E_2 \rightarrow F$ une application bilinéaire. B est différentiable en tout point (a_1, a_2) de $E_1 \times E_2$ et

$$\forall h = (h_1, h_2) \in E_1 \times E_2 \quad dB(a_1, a_2)(h_1, h_2) = B(a_1, h_2) + B(h_1, a_2)$$

En effet, l'application $h = (h_1, h_2) \rightarrow L(h) = B(a_1, h_2) + B(h_1, a_2)$ est linéaire et $B(a_1 + h_1, a_2 + h_2) - B(a) - L(h) = B(h_1, h_2)$. Munissons $E = E_1 \times E_2$ de la norme définie par $\|h\| = \|h_1\| + \|h_2\|$. L'application bilinéaire B est continue, donc il existe $C > 0$ telle que $\|B(h_1, h_2)\| \leq C\|h_1\| \cdot \|h_2\| \leq C\|h\|^2$. On en déduit

$$\frac{\|B(a_1 + h_1, a_2 + h_2) - B(a) - L(h)\|}{\|h\|} \leq C\|h\|. \blacksquare$$

13.2.4 Expression de la différentielle

Soient $E \supset \Omega \xrightarrow{f} F$ et \mathcal{B}_E une base de E . Soit $h = \sum_{j=1}^{j=p} h_j e_j \in E$. Supposons f différentiable en a . On a

$$df(a)(h) = df(a) \left(\sum_{j=1}^{j=p} h_j e_j \right) = \sum_{j=1}^{j=p} h_j df(a)(e_j) = \sum_{j=1}^{j=p} h_j \frac{\partial f}{\partial x_j}(a)$$

THEOREME 13.2.2

Soient $E \supset \Omega \xrightarrow{f} F$, $a \in \Omega$ et \mathcal{B}_E une base fixée de E . f est différentiable en a si et seulement si les deux conditions suivantes sont vérifiées :

(1) f admet des dérivées partielles en a dans toutes les variables x_1, \dots, x_p .

$$(2) \quad \lim_{h \rightarrow 0} \frac{f(a+h) - f(a) - \sum_{j=1}^{j=p} h_j \frac{\partial f}{\partial x_j}(a)}{\|h\|} = 0 \quad \text{où } h = \sum_{j=1}^{j=p} h_j e_j$$

preuve

On sait que si f est différentiable en a elle admet en a des dérivées partielles. Vu l'expression ci dessus de la différentielle, la condition (2) exprime simplement la différentiabilité de f . La condition est donc nécessaire.

Supposons inversement que f vérifie (1) et (2). L'application $L : h \rightarrow \sum_{j=1}^{j=p} h_j \frac{\partial f}{\partial x_j}(a)$ est une application linéaire de E dans F

et la condition (2) se traduit par $\lim_{h \rightarrow 0} \frac{f(a+h) - f(a) - L(h)}{\|h\|} = 0$ ce qui prouve que f est différentiable en a .

Exemple

On se propose d'étudier la différentiabilité en $(0,0)$ de l'application $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par $f(0,0) = 0$ et $f(x,y) = \frac{x^3 y}{x^2 + y^2}$ si $(x,y) \neq (0,0)$. Les applications $x \rightarrow f(x,0)$ et $y \rightarrow f(0,y)$ sont nulles donc dérivables en 0. Par conséquent, f admet des dérivées partielles en $(0,0)$ toutes deux nulles. Si elle est différentiable en $(0,0)$ sa différentielle L ne peut être que l'application nulle.

On munit \mathbb{R}^2 de la norme euclidienne. On a alors $f(x,y) - f(0,0) - L(x,y) = f(x,y)$ et

$$\left| \frac{f(x,y)}{\|(x,y)\|} \right| = \frac{|x^3 y|}{(x^2 + y^2)^{3/2}} \leq (x^2 + y^2)^{1/2} = \|(x,y)\|$$

Donc $\lim_{(x,y) \rightarrow (0,0)} \frac{f(x,y)}{\|(x,y)\|} = 0$ ce qui prouve que f est différentiable en $(0,0)$ et que sa différentielle en ce point est nulle.

13.2.5 Matrice jacobienne

Soient $\mathbb{R}^p \supset \Omega \xrightarrow{f} \mathbb{R}^n$, $f = (f_1, \dots, f_n)$. Supposons f différentiable en a . On appelle matrice jacobienne de f en a la matrice, dans les bases canoniques de l'application linéaire $df(a)$. Compte tenu de l'expression de $df(a)$ établie ci dessus cette matrice est

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \cdots & \frac{\partial f_1}{\partial x_p}(a) \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1}(a) & \cdots & \frac{\partial f_n}{\partial x_p}(a) \end{pmatrix}$$

Dans le cas $E \supset \Omega \xrightarrow{f} F$, la matrice ci dessus est la matrice de $d\tilde{f}(\tilde{a})$ et c'est aussi la matrice dans les bases $(\mathcal{B}_E, \mathcal{B}_F)$ de $df(a)$.

13.2.6 Opérations

Dans les preuves qui suivent, $\varepsilon, \varepsilon_1, \varepsilon_2$ sont des fonctions définies sur $B(0, r_a)$ qui tendent vers 0 quand h tend vers 0.

THEOREME 13.2.3

Soit $f : \Omega \rightarrow \mathbb{C}^*$ différentiable en $a \in \Omega$. Alors $g = \frac{1}{f}$ est différentiable en a et

$$dg(a) = -\frac{1}{f(a)^2} df(a)$$

preuve

Ici \mathbb{C}^* est considéré comme un ouvert de \mathbb{C} identifié en tant qu'espace vectoriel à \mathbb{R}^2 .

$$g(a+h) - g(a) = -\frac{1}{f(a)f(a+h)}(f(a+h) - f(a)) = -\frac{1}{f(a)f(a+h)}(df(a)(h) + \|h\|\varepsilon(h))$$

donc

$$g(a+h) - g(a) + \frac{1}{f(a)^2} df(a)(h) = \left(\frac{1}{f(a)^2} - \frac{1}{f(a)f(a+h)} \right) df(a)(h) - \|h\| \frac{\varepsilon(h)}{f(a)f(a+h)}$$

E étant de dimension finie, $df(a)$ est une application linéaire continue, donc il existe une constante $C > 0$ telle que $\|df(a)(h)\| \leq C\|h\|$. D'autre part, f différentiable en a est continue en a , donc $\lim_{h \rightarrow 0} f(a+h) = f(a)$. On obtient

$$\left| g(a+h) - g(a) + \frac{1}{f(a)^2} df(a)(h) \right| \leq \|h\| \eta(h)$$

où $\eta(h) = \left| \frac{1}{f(a)^2} - \frac{1}{f(a)f(a+h)} \right| C + \frac{|\varepsilon(h)|}{|f(a)f(a+h)|}$ tend vers 0 quand h tend vers 0. ■

THEOREME 13.2.4

Soient E, F_1, F_2, G quatre espaces vectoriels réels de dimension finie, Ω un ouvert non vide de E , $f_1 : \Omega \rightarrow F_1, f_2 : \Omega \rightarrow F_2$ des applications différentiables en a et $B : F_1 \times F_2 \rightarrow G$ une application bilinéaire. L'application $g : \Omega \rightarrow G$ définie par $g = B(f_1, f_2)$, c'est à dire $g(x) = B(f_1(x), f_2(x))$ est différentiable en a et pour tout $h \in E$

$$dg(a)(h) = B(df_1(a)(h), f_2(a)) + B(f_1(a), df_2(a)(h))$$

preuve

Notons L l'application $h \rightarrow B(df_1(a)(h), f_2(a)) + B(f_1(a), df_2(a)(h))$. C'est une application linéaire de E dans G . Posons $\Delta(h) = \|g(a+h) - g(a) - L(h)\|$. Il s'agit de montrer que $\Delta(h)/\|h\|$ tend vers 0 quand h tend vers 0.

- L'application bilinéaire B est continue puisque F_1 et F_2 sont de dimension finie. Il existe une constante $K > 0$ telle que $\|B(y_1, y_2)\| \leq K \|y_1\| \|y_2\|$ pour tout $(y_1, y_2) \in F_1 \times F_2$.
- Il existe des fonctions $\varepsilon_1, \varepsilon_2$ de limite nulle en 0 telles que $u_k(h) := f_k(a+h) - f_k(a) = df_k(a)(h) + \|h\| \varepsilon_k(h)$ pour $k = 1, 2$. Par continuité des applications linéaires $df_k(a)$, il existe des constantes C_k telles que $\|df_k(a)(h)\| \leq C_k \|h\|$ ce qui implique $\|u_k(h)\| \leq \|h\| [C_k + |\varepsilon_k(h)|]$.
- En utilisant la bilinéarité de B , il vient

$$\begin{aligned} g(a+h) - g(a) &= B(f_1(a) + u_1(h), f_2(a) + u_2(h)) - B(f_1(a), f_2(a)) \\ &= B(u_1(h), f_2(a)) + B(f_1(a), u_2(h)) + B(u_1(h), u_2(h)) \\ &= B(df_1(a)(h) + \|h\| \varepsilon_1(h), f_2(a)) + B(f_1(a), df_2(a)(h) + \|h\| \varepsilon_2(h)) + B(u_1(h), u_2(h)) \end{aligned}$$

donc

$$g(a+h) - g(a) - L(h) = \|h\| [B(\varepsilon_1(h), f_2(a)) + B(f_1(a), \varepsilon_2(h))] + B(u_1(h), u_2(h))$$

Par continuité de B , le terme en facteur de $\|h\|$ tend vers 0 avec h . Pour le dernier terme on a

$$\|B(u_1(h), u_2(h))\| \leq K \|u_1(h)\| \|u_2(h)\| \leq K \|h\|^2 (C_1 + |\varepsilon_1(h)|)(C_2 + |\varepsilon_2(h)|)$$

donc

$$\lim_{h \rightarrow 0} \frac{\|\Delta(h)\|}{\|h\|} = 0$$

Ces deux théorèmes généralisent les formules bien connues $\left(\frac{1}{u}\right)' = -\frac{u'}{u^2}$ et $(uv)' = u'v + uv'$.

13.2.7 Composition

THEOREME 13.2.5

Soient E, F, G trois espaces vectoriels réels de dimension finie. Soient Ω un ouvert non vide de E et U un ouvert non vide de F . Soient $f : \Omega \rightarrow F$ et $g : U \rightarrow G$ des applications. Soit enfin $a \in \Omega$. On suppose vérifiées les propriétés suivantes :

- 1) $f(\Omega) \subset U$.
- 2) f est différentiable en a .
- 3) g est différentiable en $b = f(a)$.

Alors $g \circ f : \Omega \rightarrow G$ est différentiable en a et

$$d(g \circ f)(a) = dg(f(a)) \circ df(a)$$

COROLLAIRE 13.2.2

Avec les mêmes notations que dans le théorème, si f est différentiable sur Ω et g sur U , $g \circ f$ est différentiable sur Ω .

preuve

La preuve est essentiellement celle de la composition de deux développements limités d'ordre 1.

- Il existe $r_b > 0$ telle que $B_F(b, r_b) \subset U$. Puisque f est différentiable en a , elle est continue en a , donc il existe $\rho > 0$, $\rho < r_a$ tel que $B_E(a, \rho) \subset \Omega$ et $f(B_E(a, \rho)) \subset B_F(b, r_b)$.
- Il existe des fonctions $\varepsilon_1 : B_E(0, \rho) \rightarrow F$ et $\varepsilon_2 : B_F(0, r_b) \rightarrow G$ vérifiant

$$f(a+h) = f(a) + df(a)(h) + \|h\| \varepsilon_1(h), \quad \lim_{h \rightarrow 0} \varepsilon_1(h) = 0$$

$$g(b+k) = g(b) + dg(b)(k) + \|k\| \varepsilon_2(k), \quad \lim_{k \rightarrow 0} \varepsilon_2(k) = 0$$

- Par continuité des applications linéaires $df(a)$ et $dg(b)$ on dispose de constantes $C > 0$ et $C' > 0$ telles que

$$\|df(a)(h)\| \leq C \|h\| \quad \text{et} \quad \|dg(b)(k)\| \leq C' \|k\|$$

- Puisque $b = f(a)$, pour $\|h\| < \rho$ on a $\|f(a+h) - f(a)\| < r_b$. Posons $k = f(a+h) - f(a) = df(a)(h) + \|h\| \varepsilon_1(h)$. On a

$$\|k\| \leq (C + \|\varepsilon_1(h)\|) \|h\|$$

On peut maintenant terminer la démonstration :

$$g \circ f(a+h) - g \circ f(a) - dg(b) \circ df(a)(h) = dg(b)(\|h\| \varepsilon_1(h)) + \|k\| \varepsilon_2(k)$$

donc

$$\|g \circ f(a+h) - g \circ f(a) - dg(b) \circ df(a)(h)\| \leq \|h\| \eta(h)$$

avec

$$\eta(h) = C' \|\varepsilon_1(h)\| + (C + \|\varepsilon_1(h)\|) \|\varepsilon_2(k)\|$$

Quand h tend vers 0, k tend vers 0, donc $\eta(h)$ tend aussi vers 0 ce qui achève la preuve.

Ce théorème, et les deux précédents, permettent dans la plupart des cas de prouver la différentiabilité d'une application. Détaillons sur un exemple.

Exemple Etudions la différentiabilité de l'application $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par

$$f(0,0) = 0 \quad \text{et} \quad f(x,y) = \frac{(\cos(xy) - 1) \sin \sqrt{x^2 + y^2}}{x^2 + y^2} \quad \text{si} \quad (x,y) \neq (0,0)$$

Posons $\Omega = \mathbb{R}^2 \setminus \{(0,0)\}$

- Les applications $(x,y) \rightarrow x$ et $(x,y) \rightarrow y$ sont linéaires donc différentiables en tout point. En utilisant le théorème sur les produits, on voit que les applications $v : (x,y) \rightarrow xy$, $(x,y) \rightarrow x^2$ et $(x,y) \rightarrow y^2$ sont aussi différentiables. Il en est donc de même de $D : (x,y) \rightarrow x^2 + y^2$.
- La fonction $t \rightarrow \sqrt{t}$ est dérivable sur \mathbb{R}^{+*} , donc différentiable sur ce domaine (exemple 13.2.3). Par conséquent, d'après le théorème de composition, l'application $u_1 : (x,y) \rightarrow \sqrt{D(x,y)} = \sqrt{x^2 + y^2}$ est différentiable sur Ω . Comme l'application $\sin : \mathbb{R} \rightarrow \mathbb{R}$ est dérivable, elle est différentiable, donc l'application $u_2 : (x,y) \rightarrow \sin \sqrt{x^2 + y^2}$ est différentiable sur Ω .
- De même, de la dérivabilité de la fonction \cos , on déduit par composition que la fonction $\cos \circ v$ est différentiable sur \mathbb{R}^2 . Il en est donc de même de la fonction $u_3 : (x,y) \rightarrow \cos(xy) - 1$.
- Le théorème 13.2.4 assure que $u_3 \cdot u_2$ est différentiable sur Ω . Le théorème sur les quotients assure que $1/D$ est différentiable sur Ω . Enfin, le théorème 13.2.4 montre que $f = (u_3 \cdot u_2) \cdot \frac{1}{D}$ est différentiable sur Ω .

- Reste à étudier la différentiabilité en 0. Posons $\rho = \sqrt{x^2 + y^2}$. ρ est la norme euclidienne de (x,y) . D'après l'inégalité des accroissements finis, on a pour tout $t \in \mathbb{R}$ $|\cos(t) - 1| \leq t^2/2$ et $|\sin t| \leq |t|$. On en déduit, pour $(x,y) \neq (0,0)$ $|f(x,y)| \leq \frac{x^2 y^2}{2\sqrt{x^2 + y^2}} \leq \frac{1}{2} \rho^3$. Comme $f(0,0) = 0$, on voit que $f(x,y) - f(0,0) = o(\|(x,y)\|)$ ce qui montre que f est différentiable en $(0,0)$ et que $df(0,0)$ est l'application linéaire nulle.

THEOREME 13.2.6

Soient Ω un ouvert non vide de \mathbb{R}^p et U un ouvert non vide de \mathbb{R}^n . Soient $f : \Omega \rightarrow \mathbb{R}^n$ et $g : U \rightarrow \mathbb{R}^q$ des applications. Soit enfin $a \in \Omega$. On suppose vérifiée les propriétés suivantes :

- 1) $f(\Omega) \subset U$.
- 2) f est différentiable en a .
- 3) g est différentiable en $b = f(a)$.

Notons $x = (x_1, \dots, x_p)$ un élément générique de Ω , $y = (y_1, \dots, y_n)$ un élément générique de U . Posons $f(x) = (f_1(x), \dots, f_n(x))$. Soit enfin $h = g \circ f : \Omega \rightarrow \mathbb{R}^q$. On sait que h est différentiable en a . On a, pour $1 \leq j \leq p$

$$\frac{\partial h}{\partial x_j}(a) = \sum_{i=1}^{i=n} \frac{\partial f_i}{\partial x_j}(a) \frac{\partial g}{\partial y_i}(b)$$

preuve

On a $dh(a) = dg(b) \circ df(a)$ donc $\text{Mat}(dh(a)) = \text{Mat}(dg(b))\text{Mat}(df(a))$. $\frac{\partial h}{\partial x_j}(a)$ est la j -ième colonne de $\text{Mat}(dh(a))$. L'égalité à prouver n'est autre qu'une forme condensée de la formule donnant le produit de deux matrices.

Cette formule est souvent appelée "règle de chaînes" (*chain rule* dans les ouvrages en anglais).

Notons qu'il s'agit d'une écriture condensée puisque $b = f(a)$. En particulier si f est différentiable sur Ω et g sur U on a en fait l'égalité de fonctions

$$\frac{\partial h}{\partial x_j} = \sum_{i=1}^{i=n} \frac{\partial f_i}{\partial x_j} \cdot \left(\frac{\partial g}{\partial y_i} \circ f \right)$$

Bien entendu, on a des formules identiques dans le cas $E \supset \Omega \xrightarrow{f} F$ si les espaces E et F sont munis de bases.

13.2.8 Notation différentielle

Définition

Supposons que Ω soit un ouvert de \mathbb{R}^p et soit $f : \Omega \rightarrow F$ différentiable sur Ω . Soit (e_1, \dots, e_p) la base canonique de \mathbb{R}^p et $(\varepsilon_1, \dots, \varepsilon_p)$ la base duale. Soit $x = (x_1, \dots, x_p)$ un élément quelconque de Ω . Pour $1 \leq j \leq p$ désignons encore par x_j l'application qui à $x \in \Omega$ associe sa j -ième coordonnée x_j . C'est bien sur un abus de notation, puisqu'on désigne de la même manière une fonction et la valeur qu'elle prend en un point quelconque. L'application x_j est la restriction à Ω de l'application linéaire ε_j . Elle est donc différentiable en tout point de Ω et $dx_j(a) = \varepsilon_j$. Par conséquent dx_j est l'application constante de Ω dans $L(\mathbb{R}^p, \mathbb{R})$ qui à $x \in \Omega$ associe ε_j .

Soit $h = (h_1, \dots, h_p) \in \mathbb{R}^p$. On a

$$df(x)(h) = \sum_{j=1}^{j=p} h_j \frac{\partial f}{\partial x_j}(x) = \sum_{j=1}^{j=p} [dx_j(x)(h)] \frac{\partial f}{\partial x_j}(x)$$

Soit encore

$$\forall x \in \Omega \quad df(x) = \sum_{j=1}^{j=p} [dx_j(x)] \cdot \frac{\partial f}{\partial x_j}(x)$$

Ce que l'on écrit sous la forme dite notation différentielle

$$(1) \quad df = \sum_{j=1}^{j=p} \frac{\partial f}{\partial x_j} dx_j$$

Remarque sur les notations

Soient E et F deux \mathbb{K} -espaces vectoriels, $\varphi \in E^*$ (dual de E) et $u \in F$. On note usuellement $u \otimes \varphi$ l'application linéaire de E dans F qui à $t \in E$ associe $\varphi(t)u$.

C'est cette écriture que l'on utilise ici, en supprimant le \otimes . $\frac{\partial f}{\partial x_j}(x) dx_j$ est l'application linéaire de E dans F qui à $h \in E$ associe $dx_j(h) \frac{\partial f}{\partial x_j}(x)$.

Si la fonction f est à valeurs réelles, cette écriture ne pose aucun problème : les $\frac{\partial f}{\partial x_j}(x)$ sont des nombres et les dx_j des formes linéaires.

Composition

Revenons maintenant à la situation du théorème 13.2.6. On suppose $F = \mathbb{R}^n$, U est un ouvert de \mathbb{R}^n et f est telle que $f(\Omega) \subset U$. On se donne aussi $g : U \rightarrow G$ différentiable en tout point de U . Notons $y = (y_1, \dots, y_n)$ un élément générique de U . On a donc

$$(2) \quad dg = \sum_{i=1}^{i=n} \frac{\partial g}{\partial y_i} dy_i$$

Ecrivons $f = (f_1, \dots, f_p)$ et soit $h = g \circ f$. Passer de g à h revient à remplacer dans l'expression $g(y_1, \dots, y_n)$ y_i par $f_i(x_1, \dots, x_p)$. La relation (1) appliquée à f_i s'écrit

$$df_i = \sum_{j=1}^{j=p} \frac{\partial f_i}{\partial x_j} dx_j$$

La règle de chaîne donne

$$\frac{\partial h}{\partial x_j}(x) = \sum_{i=1}^{i=n} \frac{\partial g}{\partial y_i}(f(x)) \frac{\partial f_i}{\partial x_j}(x)$$

donc

$$\begin{aligned}
 dh(x) &= \sum_{j=1}^{j=p} \frac{\partial h}{\partial x_j}(x) dx_j = \sum_{j=1}^{j=p} \left(\sum_{i=1}^{i=n} \frac{\partial g}{\partial y_i}(f(x)) \frac{\partial f_i}{\partial x_j}(x) \right) dx_j \\
 &= \sum_{i=1}^{i=n} \frac{\partial g}{\partial y_i}(f(x)) \left(\sum_{j=1}^{j=p} \frac{\partial f_i}{\partial x_j}(x) dx_j \right) \\
 &= \sum_{i=1}^{i=n} \frac{\partial g}{\partial y_i}(f(x)) df_i(x)
 \end{aligned}$$

Autrement dit, pour obtenir la différentielle de $g \circ f$, on remplace dans l'expression (2) de la différentielle de g les $\frac{\partial g}{\partial y_i}$ par $\frac{\partial g}{\partial y_i} \circ f$ et les dy_i par les df_i , puis on regroupe les termes en dx_j :

$$\text{si } dg = \sum_{i=1}^{i=n} \frac{\partial g}{\partial y_i} dy_i \text{ alors } d(g \circ f) = \sum_{i=1}^{i=n} \left(\frac{\partial g}{\partial y_i} \circ f \right) df_i$$

C'est cette propriété qui fait l'intérêt de la notation différentielle.

Exemple : passage en polaires

Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ différentiable. Soit $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ définie par

$$\varphi(\rho, \theta) = (\rho \cos \theta, \rho \sin \theta)$$

et $F = f \circ \varphi$. L'application φ s'écrit $\varphi = (\varphi_1, \varphi_2)$ avec

$$\begin{aligned}
 \varphi_1(\rho, \theta) &= \rho \cos \theta \\
 \varphi_2(\rho, \theta) &= \rho \sin \theta
 \end{aligned}$$

d'où

$$\begin{aligned}
 d\varphi_1 &= \cos \theta d\rho - \rho \sin \theta d\theta \\
 d\varphi_2 &= \sin \theta d\rho + \rho \cos \theta d\theta
 \end{aligned}$$

On a

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy$$

Donc

$$\begin{aligned}
 dF(\rho, \theta) &= \frac{\partial f}{\partial x}(\cdot, \cdot)(\cos \theta d\rho - \rho \sin \theta d\theta) + \frac{\partial f}{\partial y}(\cdot, \cdot)(\sin \theta d\rho + \rho \cos \theta d\theta) \\
 &= \left(\cos \theta \frac{\partial f}{\partial x}(\cdot, \cdot) + \sin \theta \frac{\partial f}{\partial y}(\cdot, \cdot) \right) d\rho + \left(-\sin \theta \frac{\partial f}{\partial x}(\cdot, \cdot) + \cos \theta \frac{\partial f}{\partial y}(\cdot, \cdot) \right) \rho d\theta
 \end{aligned}$$

où, pour abrégier $(\cdot, \cdot) = (\rho \cos \theta, \rho \sin \theta)$

On en déduit

$$\begin{aligned}
 \frac{\partial F}{\partial \rho}(\rho, \theta) &= \cos \theta \frac{\partial f}{\partial x}(\rho \cos \theta, \rho \sin \theta) + \sin \theta \frac{\partial f}{\partial y}(\rho \cos \theta, \rho \sin \theta) \\
 \frac{\partial F}{\partial \theta}(\rho, \theta) &= -\rho \sin \theta \frac{\partial f}{\partial x}(\rho \cos \theta, \rho \sin \theta) + \rho \cos \theta \frac{\partial f}{\partial y}(\rho \cos \theta, \rho \sin \theta)
 \end{aligned}$$

13.2.9 Gradient

Soit Ω un ouvert d'un espace euclidien $(E, \langle \cdot | \cdot \rangle)$ et $f : \Omega \rightarrow \mathbb{R}$ différentiable sur Ω . La structure euclidienne de E permet d'identifier E et son dual E^* par l'isomorphisme canonique qui à tout vecteur e de E associe la forme linéaire $v \rightarrow \langle e | v \rangle$. Or la différentielle de f en un point x de Ω est une forme linéaire $df(x)$ sur E , c'est à dire un élément de E^* . Il existe donc un unique vecteur $V(x)$ de E tel que pour tout $h \in E$, $df(x)(h) = \langle V(x) | h \rangle$. Ce vecteur $V(x)$ s'appelle le gradient de f en x et se note $\text{grad} f(x)$ ou plus souvent $\nabla f(x)$.

Si $\mathcal{B} = (e_1, \dots, e_p)$ est une base **orthonormée** de E et si (h_1, \dots, h_p) désignent les coordonnées dans cette base d'un vecteur arbitraire h on a $df(x)(h) = \sum_{j=1}^{j=p} \frac{\partial f}{\partial x_j}(x) h_j$ de sorte que $\nabla f(x)$ est le vecteur

$$\nabla f(x) = \sum_{j=1}^{j=p} \frac{\partial f}{\partial x_j}(x) e_j$$

où, rappelons le, $\frac{\partial f}{\partial x_j}(x)$ est la dérivée de f en x selon le vecteur e_j .

13.3 Fonctions de classe C^1

E et F étant deux espaces vectoriels normés de dimension finie, l'espace $L(E, F)$ des applications linéaires de E dans F est de dimension finie et peut être muni d'une norme. On choisira en général la norme subordonnée à celle de E et de F qui est donnée, pour $\varphi \in L(E, F)$ par

$$\|\varphi\| = \sup_{x \in E, \|x\| \leq 1} \|\varphi(x)\|$$

DEFINITION 13.3.1

Soient Ω un ouvert de E et $f : \Omega \rightarrow F$ une application. On dit que f est de classe C^1 sur Ω si elle est différentiable et si l'application différentielle $f' = df : \Omega \rightarrow L(E, F)$ est continue sur Ω .

Si Ω est un ouvert de \mathbb{R} , l'identification de $L(\mathbb{R}, F)$ avec F identifie la différentielle $df(x)$ avec la dérivée usuelle $f'(x)$. Donc f est C^1 au sens de la nouvelle définition si elle l'est au sens habituel. Cette définition est donc une généralisation de la définition des applications de classe C^1 d'un ouvert de \mathbb{R} dans E .

THEOREME 13.3.1

1. Soit $E \supset \Omega \xrightarrow{f} F$ et (f_1, \dots, f_n) les composantes de f dans une base \mathcal{B}_F . f est C^1 ssi chacune des f_i est C^1 .
2. Soient $g, h : \Omega \rightarrow F$ de classe C^1 et $\lambda \in \mathbb{R}$. $\lambda g + h$ est de classe C^1 . L'ensemble $C^1(\Omega, F)$ des fonctions de classe C^1 de Ω dans F est un sous espace vectoriel de $\mathcal{F}(\Omega, F)$ et l'application $f \rightarrow df$ de $C^1(\Omega, F)$ dans l'espace $C^0(\Omega, L(E, F))$ des applications continues de Ω dans $L(E, F)$ est linéaire.
3. Soit $f : \Omega \rightarrow \mathbb{C}^*$ une application de classe C^1 . Alors $1/f$ est aussi de classe C^1 .
4. Soient E, F_1, F_2, G quatre espaces vectoriels réels de dimension finie, Ω un ouvert non vide de E , $f_1 : \Omega \rightarrow F_1, f_2 : \Omega \rightarrow F_2$ des applications de classe C^1 et $B : F_1 \times F_2 \rightarrow G$ une application bilinéaire. L'application $g : \Omega \rightarrow G$ définie par $g = B(f_1, f_2)$, c'est à dire $g(x) = B(f_1(x), f_2(x))$ est de classe C^1 .
5. Soient E, F, G trois espaces vectoriels réels de dimension finie. Soient Ω un ouvert non vide de E et U un ouvert non vide de F . Soient $f : \Omega \rightarrow F$ et $g : U \rightarrow G$ des applications de classe C^1 telles que $f(\Omega) \subset U$. Alors $g \circ f : \Omega \rightarrow G$ est de classe C^1 .

preuve

Prouvons le point 5. D'après le théorème 13.2.6, $g \circ f$ est différentiable en tout point de Ω et $d(g \circ f)(x) = dg(f(x)) \circ df(x)$.

Montrons que $d(g \circ f) : \Omega \rightarrow L(E, G)$ est continue.

- Par hypothèse, l'application $\Omega \rightarrow L(E, F), x \rightarrow df(x)$ est continue.
- De même, l'application $U \rightarrow L(F, G), y \rightarrow dg(y)$ est continue. f étant différentiable est continue, donc par composition l'application $dg \circ f : \Omega \rightarrow L(F, G)$ est continue.
- Il en résulte que l'application $\Theta : \Omega \rightarrow L(F, G) \times L(E, F)$ définie par $\Theta(x) = (dg(f(x)), df(x))$ est continue.
- L'application $B : L(F, G) \times L(E, F) \rightarrow L(E, G)$ définie par $B(\varphi, \psi) = \varphi \circ \psi$ est continue car B est bilinéaire et $L(E, F), L(F, G)$ sont de dimension finie. On en déduit que la composée $B \circ \Theta = d(g \circ f)$ est continue. ■

La preuve des autres points est laissée au lecteur. Pour 3 (resp. 4) on utilise la formule donnée par le théorème 13.2.3 (resp. 13.2.4).

THEOREME 13.3.2

Soient $E \supset \Omega \xrightarrow{f} F$ et $\mathcal{B}_E = (e_1, \dots, e_p)$ une base de E . f est de classe C^1 ssi f admet des dérivées partielles dans toutes les variables x_1, \dots, x_p et si ces dérivées partielles sont continues.

Si f est C^1 elle est différentiable donc admet des dérivées partielles. La dérivée partielle par rapport à la j -ième variable est donnée par $x \rightarrow \partial_{e_j} f(x) = df(x)(e_j)$. Si $x \rightarrow df(x)$ est continue, il en est de même de $x \rightarrow df(x)(e_j)$.

Réciproque : on sait qu'une fonction peut avoir des dérivées partielles sans être continue, donc à fortiori sans être différentiable. La difficulté de la preuve consiste à montrer que si f admet des dérivées partielles continues sur Ω , elle est différentiable.

Ensuite, on voit facilement en utilisant $df(x) = \sum_{j=1}^{j=p} \frac{\partial f}{\partial x_j}(x) dx_j$ que l'application df est continue.

Ce résultat est admis par le programme. On en trouvera une preuve dans l'appendice.

Remarque

Il en résulte que si f admet des dérivées partielles continues relativement à une base \mathcal{B}_E de E , il en est de même relativement à une autre base \mathcal{B}'_E , ce qui peut évidemment se voir aussi avec la règle de chaîne.

13.4 Exemples et applications

EXEMPLE 13.4.1 (Courbes)

Soit $E \supset \Omega \xrightarrow{f} F$ de classe C^1 . Soit d'autre part I un intervalle ouvert de \mathbb{R} non vide et $\gamma : I \rightarrow \Omega$ une courbe paramétrée de classe C^1 . Soit enfin $t_0 \in I$. L'application $\Gamma = g \circ \gamma : I \rightarrow F$ est une courbe paramétrée de classe C^1 . Posons $m_0 = \gamma(t_0)$. On a

$$\Gamma'(t_0) = df(m_0)(\gamma'(t_0))$$

Autrement dit, le vecteur vitesse au temps t_0 de Γ est l'image par $df(\gamma(t_0))$ du vecteur vitesse de γ au temps t_0 .

En effet, le théorème de composition donne $d\Gamma(t_0) = df(m_0) \circ d\gamma(t_0)$. Mais $d\gamma(t_0)$ est l'application linéaire $\mathbb{R} \rightarrow E$ définie par $h \rightarrow h\gamma'(t_0)$, donc $d\Gamma(t_0)(h) = df(m_0)(h\gamma'(t_0)) = h df(m_0)(\gamma'(t_0))$ d'où la conclusion.

EXEMPLE 13.4.2 (Plan tangent)

Soit Ω un ouvert de \mathbb{R}^2 et $f : \Omega \rightarrow \mathbb{R}$ une fonction de classe C^1 . Soit Σ le graphe de f :

$$\Sigma = \{(x, y, z) \in \mathbb{R}^3 \mid (x, y) \in \Omega \text{ et } z = f(x, y)\}$$

On dit que Σ est la surface d'équation $z = f(x, y)$. Soit $(x_0, y_0) \in \Omega$. La différentiabilité de f se traduit, pour $(x, y) \in \Omega$, par

$$f(x, y) = f(x_0, y_0) + (x - x_0) \frac{\partial f}{\partial x}(x_0, y_0) + (y - y_0) \frac{\partial f}{\partial y}(x_0, y_0) + \|(x - x_0), (y - y_0)\| \varepsilon(x - x_0, y - y_0)$$

où ε tend vers 0 en 0. Considérons la fonction affine définie sur \mathbb{R}^2 par

$$A(X, Y) = f(x_0, y_0) + (X - x_0) \frac{\partial f}{\partial x}(x_0, y_0) + (Y - y_0) \frac{\partial f}{\partial y}(x_0, y_0)$$

Son graphe est le plan affine Π d'équation $Z = A(X, Y)$. Ce plan s'appelle le plan tangent à la surface au point (x_0, y_0) . Le plan vectoriel directeur $\vec{\Pi}$ de Π a pour équation $Z = X \frac{\partial f}{\partial x}(x_0, y_0) + Y \frac{\partial f}{\partial y}(x_0, y_0)$.

Soient I un intervalle ouvert de \mathbb{R} contenant 0 et $\gamma : I \rightarrow \mathbb{R}^3$ une courbe de classe C^1 dont le support est inclus dans Σ et telle que $\gamma(0) = (x_0, y_0)$. En écrivant $\gamma(t) = (x(t), y(t), z(t))$ on voit que le support de γ est inclus dans Σ ssi $\forall t \in I \quad z(t) = f(x(t), y(t))$. En utilisant la règle de chaîne, on obtient

$$z'(t) = x'(t) \frac{\partial f}{\partial x}(x(t), y(t)) + y'(t) \frac{\partial f}{\partial y}(x(t), y(t))$$

et pour $t = 0$

$$z'(0) = x'(0) \frac{\partial f}{\partial x}(x_0, y_0) + y'(0) \frac{\partial f}{\partial y}(x_0, y_0)$$

ce qui montre que le vecteur vitesse $\gamma'(0) = (x'(0), y'(0), z'(0))$ appartient au plan vectoriel $\overrightarrow{\Pi}$. La tangente au point de paramètre 0 de γ est donc une droite affine contenue dans le plan Π .

Réciproquement, soit D une droite affine contenue dans Π et passant par le point $M_0 = (x_0, y_0, z_0)$. Soit $e = (a, b, c)$ un vecteur directeur de D . On a $e \in \overrightarrow{\Pi}$ donc $c = a \frac{\partial f}{\partial x}(x_0, y_0) + b \frac{\partial f}{\partial y}(x_0, y_0)$. Considérons la courbe paramétrée $\gamma(t) = (x_0 + at, y_0 + bt, f(x_0 + at, y_0 + bt))$. γ est définie et C^1 dans un intervalle ouvert J de \mathbb{R} contenant 0, le support de γ est contenu dans Σ , et $\gamma'(0) = (a, b, a \frac{\partial f}{\partial x}(x_0, y_0) + b \frac{\partial f}{\partial y}(x_0, y_0)) = (a, b, c) = e$. Comme $\gamma(0) = M_0$, la tangente à γ en le paramètre 0 est la droite D .

Π est donc la réunion des droites affines passant par M_0 qui sont les tangentes d'arcs paramétrés de classe C^1 , tracés sur Σ , et passant par ce point.

EXEMPLE 13.4.3 (Norme euclidienne)

Soit $(E, \langle \cdot | \cdot \rangle)$ un espace euclidien et $\| \cdot \|$ la norme associée. C'est une application de E dans \mathbb{R} . Posons $\varphi(x) = \|x\|$. L'application φ est de classe C^1 sur $E \setminus \{0\}$ et sa différentielle en un point $a \neq 0$ est l'application linéaire

$$h \rightarrow d\varphi(a)(h) = \frac{\langle a | h \rangle}{\|a\|}$$

Autrement dit

$$\nabla\varphi(a) = \frac{a}{\|a\|}$$

preuve

Soit $\Omega = E \setminus \{0\}$. On a $\varphi = u_2 \circ u_1$ avec $u_1 : \Omega \rightarrow \mathbb{R}$, $x \rightarrow \langle x | x \rangle$ et $u_2 : \mathbb{R}_+^* \rightarrow \mathbb{R}$, $u_2(t) = \sqrt{t}$. On a bien $u_1(\Omega) \subset \mathbb{R}_+^*$.

- On a $u_1(a+h) = u_1(a) + 2\langle a | h \rangle + \|h\|^2$. $h \rightarrow \langle a | h \rangle$ est linéaire et $\|h\|^2 = o(\|h\|)$. Donc u_1 est différentiable en a et $du_1(a)(h) = 2\langle a | h \rangle$.

- u_2 est C^1 sur \mathbb{R}_+^* et pour $t_0 \in \mathbb{R}_+^*$, $du_2(t_0)$ est la multiplication par $u_2'(t_0) = \frac{1}{2\sqrt{t_0}}$. Par conséquent $\varphi = u_2 \circ u_1$ est différentiable en a et $d\varphi(a)(h) = du_2(\|a\|^2) \circ du_1(a)(h) = \frac{\langle a | h \rangle}{\|a\|}$. Il est clair que $d\varphi(a)$ dépend continuellement de a . ■

Remarque 1 : Si F est un espace vectoriel non réduit à $\{0\}$ et si N est une norme sur F , N n'est pas différentiable en 0. En effet, si e est un élément non nul de F , l'application de \mathbb{R} dans \mathbb{R} $t \rightarrow N(te) = |t|N(e)$ n'est pas dérivable en 0 puisque $N(e) \neq 0$. Or si N était différentiable en 0, cette application serait dérivable en 0.

Remarque 2 : Le calcul fait pour u_1 est en fait valable pour toute forme quadratique q sur un espace vectoriel réel E de dimension finie : si b est la forme bilinéaire symétrique associée à q , on a pour tout $a, h \in E$ $q(a+h) = q(a) + 2b(a, h) + q(h)$. L'application $h \rightarrow 2b(a, h)$ est linéaire et $q(h) = o(\|h\|)$. (Le vérifier !) donc $dq(a)(h) = 2b(a, h)$.

EXEMPLE 13.4.4 (Déterminant)

Considérons l'application $D : M_n(\mathbb{R}) \rightarrow \mathbb{R}$ définie par $D(M) = \det(M)$.

- D est C^1 . En effet, si $X = (x_{i,j})$ on a $\det(X) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) x_{1,\sigma(1)} \cdots x_{n,\sigma(n)}$. L'application $X \rightarrow x_{i,j}$ est linéaire donc C^1 et D s'écrit donc comme somme de $n!$ fonctions qui sont C^1 chacune comme produits de n fonctions de classe C^1 .

- On a $\det(X) = \sum_{i=1}^{i=n} x_{i,j} A_{i,j}$ où $A_{i,j}$ est le cofacteur de $x_{i,j}$. La variable $x_{i,j}$ n'apparaît pas dans le développement de $A_{i,j}$ de sorte que

$$\frac{\partial D}{\partial x_{i,j}}(X) = A_{i,j}$$

Soit $H = (h_{i,j}) \in M_n(\mathbb{R})$. On a donc $dD(X)(H) = \sum_{i,j} h_{i,j} A_{i,j}$. Soit $\text{Cof}(X)$ la matrice des cofacteurs de X . On voit que

$$\sum_{i,j} h_{i,j} A_{i,j} = \sum_{i=1}^{i=n} \left(\sum_{j=1}^{j=n} h_{i,j} [{}^t\text{Cof}(X)]_{j,i} \right) = \text{tr}(H {}^t\text{Cof}(X)) \text{ donc}$$

$$dD(X)(H) = \text{tr}(H {}^t\text{Cof}(X))$$

En particulier, si on choisit $X = I_n$, la matrice des cofacteurs est I_n et la différentielle de l'application déterminant au point I_n n'est autre que l'application trace.

Application

Soit $A \in M_n(\mathbb{R})$ et $f : \mathbb{R} \rightarrow M_n(\mathbb{R})$ l'application $t \rightarrow \det(\exp(tA))$. On sait que $t \rightarrow \exp(tA)$ est dérivable en 0 et que sa dérivée est A . Par composition on a donc

$$f'(0) = \text{tr}(A)$$

EXEMPLE 13.4.5 (Inversion des matrices)

Soit $\Phi : GL(n, \mathbb{R}) \rightarrow M_n(\mathbb{R})$ l'application définie par $\Phi(M) = M^{-1}$. Cette application est de classe C^1 et pour $A \in GL(n, \mathbb{R})$ on a

$$\forall H \in M_n(\mathbb{R}) \quad d\Phi(A)(H) = -A^{-1}HA^{-1}$$

• Φ est C^1 . Désignons par $T_{i,j}$ l'application qui à une matrice M associe le cofacteur de l'élément d'indice (i, j) . $T_{i,j}$ est une fonction polynomiale en les coefficients de M donc $M \rightarrow T_{i,j}(M)$ est C^1 . Ensuite $\Phi(M)_{i,j} = \frac{T_{j,i}(M)}{\det(M)}$ est C^1 sur $GL(n, \mathbb{R})$ comme quotient de deux fonctions C^1 celle figurant au dénominateur ne s'annulant pas. Donc Φ est C^1 .

• Soit $A \in GL(n, \mathbb{R})$. Considérons l'application $M \rightarrow M\Phi(M)$. Cette application est constante, donc sa différentielle est nulle. Mais d'autre part, c'est la composée de l'application bilinéaire $(M, M') \rightarrow MM'$ et de l'application $M \rightarrow (M, \Phi(M))$. La différentielle de la seconde est $H \rightarrow (H, d\Phi(A)(H))$ donc la différentielle en A de $M \rightarrow M\Phi(M)$ est $H \rightarrow H\Phi(A) + Ad\Phi(A)(H)$. Par conséquent, on a $\forall H \in M_n(\mathbb{R}) \quad H\Phi(A) + Ad\Phi(A)(H) = 0$ donc $d\Phi(A)(H) = -A^{-1}H\Phi(A) = -A^{-1}HA^{-1}$.

Remarque : Le résultat établi ci dessus généralise donc la formule donnant la dérivée de $x \rightarrow 1/x$. En effet, pour $n = 1$, $GL(1, \mathbb{R})$ s'identifie à \mathbb{R}^* et Φ est l'application $x \rightarrow 1/x$. Sa dérivée en $a \neq 0$ est $-1/a^2$, donc sa différentielle en a est $h \rightarrow -h/a^2$. Les résultats coïncident (ouf!) car la multiplication dans $M_1(\mathbb{R})$ est commutative!

13.5 Inégalités des accroissements finis

Si a et b sont deux points d'un espace vectoriel réel E , on notera $[a, b]$ le segment d'extrémités a et b , c'est à dire l'ensemble $\{(1-t)a + tb; 0 \leq t \leq 1\}$.

THEOREME 13.5.1

Soient Ω un ouvert non vide d'un espace vectoriel normé de dimension finie E , F un autre espace vectoriel normé de dimension finie, $f : \Omega \rightarrow F$ une application de classe C^1 et $L(E, F)$ l'espace des applications linéaires de E dans F muni de la norme subordonnée aux normes de E et F .

Soient $a, b \in \Omega$ tels que $[a, b] \subset \Omega$ et $K \geq 0$ tel que

$$\forall x \in [a, b] \quad \|df(x)\| \leq K$$

Alors

$$\|f(b) - f(a)\| \leq K\|b - a\|$$

preuve

Soit $\varphi : [0, 1] \rightarrow F$ définie par $\varphi(t) = f(a + t(b-a))$. Par composition, φ est de classe C^1 . On a $\varphi'(t) = df(a + t(b-a))(b-a)$ donc $\|\varphi'(t)\| \leq K\|b - a\|$. L'inégalité des accroissements finis pour une fonction à valeurs dans un espace vectoriel normé donne $\|\varphi(1) - \varphi(0)\| \leq K\|b - a\|$ ce qui est le résultat énoncé.

COROLLAIRE 13.5.1

Soit Ω un ouvert connexe de E et $df : \Omega \rightarrow F$ de classe C^1 . Si la différentielle de f est identiquement nulle sur Ω , f est constante.

preuve

Soient $a \in \Omega$ et $D = \{x \in \Omega \mid f(x) = f(a)\}$.

- f étant continue, D est un fermé de Ω comme image réciproque par f du fermé $\{f(a)\}$ de F .
- Soit $m \in D$. Soit $r > 0$ tel que $B(m, r) \subset \Omega$. Toute boule est convexe, donc si $z \in B(m, r)$ le segment $[m, z]$ est contenu dans $B(m, r)$ donc dans Ω . En tout point u de ce segment on a $df(u) = 0$, donc d'après l'inégalité des accroissements finis appliquée avec $K = 0$, on a $f(z) = f(m) = f(a)$. Ceci montre que $B(m, r) \subset D$ donc que D est un ouvert de E donc de Ω .
- D est un fermé ouvert non vide ($a \in D$) de Ω connexe, donc $D = \Omega$ et $\forall x \in \Omega \quad f(x) = f(a)$ donc f est constante.

Enfin, pour les fonctions à valeurs réelles, on dispose d'une généralisation de l'égalité des accroissements finis :

PROPOSITION 13.5.1

Soient Ω un ouvert d'un espace vectoriel de dimension finie E , $f : \Omega \rightarrow \mathbb{R}$ de classe C^1 et $a, b \in \Omega$ tels que $[a, b] \subset \Omega$. Alors

il existe $\theta \in]0, 1[$ tel que

$$f(b) - f(a) = df(a + \theta(b - a))(b - a) = \sum_{j=1}^{j=p} \frac{\partial f}{\partial x_j} ((1 - \theta)a + \theta b)(b_j - a_j)$$

Pour le montrer il suffit d'appliquer le théorème des accroissements finis usuels à la fonction $\varphi : [0, 1] \rightarrow \mathbb{R}$ introduite dans la démonstration du théorème ci dessus.

13.6 Dérivées d'ordre supérieur

13.6.1 Dérivées partielles d'ordre supérieur

DEFINITION 13.6.1

Soient $E \supset \Omega \xrightarrow{f} F$ et \mathcal{B}_E une base de E . Soient j, i entre 1 et p . Sous réserve d'existence, on note $\frac{\partial^2 f}{\partial x_i \partial x_j}$ la dérivée partielle par rapport à la i -ième variable de la dérivée partielle $\frac{\partial f}{\partial x_j}$ de f par rapport à la j -ième variable :

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial}{\partial x_i} \left(\frac{\partial}{\partial x_j} \right)$$

Plus généralement, on définit par récurrence les dérivées partielles de tout ordre, sous réserve d'existence :

$$\frac{\partial^m f}{\partial x_{i_m} \cdots \partial x_{i_1}} = \frac{\partial f}{\partial x_{i_m}} \left(\frac{\partial^{m-1} f}{\partial x_{i_{m-1}} \cdots \partial x_{i_1}} \right)$$

Si $\dim E = p$, il y a p^m dérivées partielles d'ordre m .

DEFINITION 13.6.2

On dira que f est k fois dérivable sur Ω si toutes les dérivées partielles d'ordre k de f sont définies sur Ω .

Si f est k fois dérivable, toutes les dérivées partielles d'ordre $m \leq k$ existent donc dans Ω .

Soit $\mathcal{B}_F = (u_1, \dots, u_n)$ une base de F . Si on écrit $f(x) = f_1(x)u_1 + \dots + f_n(x)u_n$ avec $f_i : \Omega \rightarrow \mathbb{R}$ on voit facilement que f est k fois dérivable ssi chacune des f_i est k fois dérivable.

13.6.2 Fonctions de classe C^k

Si l'application $f : \Omega \rightarrow F$ est différentiable, on dispose de l'application $f' = df : \Omega \rightarrow L(E, F)$ et $L(E, F)$ est un espace vectoriel réel de dimension finie.

DEFINITION 13.6.3

Soient E, F deux espaces vectoriels normés de dimension finie, Ω un ouvert de E et $f : \Omega \rightarrow F$. Soit k un entier, $k \geq 2$. On dit que f est de classe C^k sur Ω si f est de classe C^1 et si l'application $f' = df : \Omega \rightarrow L(E, F)$ est de classe C^{k-1} .

On dit que f est de classe C^∞ si elle est de classe C^k pour tout k .

Il s'agit donc d'une définition par induction sur k . On définit ensuite la différentielle k -ième de f comme la différentielle de la différentielle $(k - 1)$ -ième de f . On la note en général $d^k f$ ou $f^{(k)}$.

Posons $F_0 = F$, $F_1 = L(E, F)$ et pour $k \geq 2$, $F_k = L(E, F_{k-1})$. Une définition équivalente est la suivante : f est de classe C^k si il existe une suite d'applications $f^{(j)} : \Omega \rightarrow F_j$ de classe C^1 telles que $f^{(0)} = f$ et $f^{(j)} = df^{(j-1)}$ pour $1 \leq j \leq k$.

13.6.3 Étude de la différentielle seconde

Si f est C^2 , $d^2 f$ est une application de Ω dans $L(E, L(E, F))$.

Un peu d'algèbre multilinéaire

Soient E et F deux \mathbb{K} -espaces vectoriels, $B(E, F)$ l'espace des applications bilinéaires de $E \times E$ dans F . Soit $u \in L(E, L(E, F))$. Pour $x \in E$, $u(x) \in L(E, F)$ donc pour $y \in E$, on peut calculer $u(x)(y)$ qui est un élément de F . Définissons $b_u : E \times E \rightarrow F$ par

$$b_u(x, y) = (u(x))(y)$$

On vérifie facilement que b_u est bilinéaire. On dispose donc d'une application $b : L(E, L(E, F)) \rightarrow B(E, F)$ $u \rightarrow b_u$.

- Cette application est linéaire. (immédiat)
- Elle est injective : si $b_u = 0$, on a pour tout x fixé dans E $u(x)(y) = 0$ pour tout y , donc $u(x)$ est l'application linéaire nulle. Ceci ayant lieu pour tout x , $u = 0$ et $\ker b = \{0\}$.
- Elle est surjective. Soit en effet $\beta : E \times E \rightarrow F$ bilinéaire. Pour $x \in E$, l'application $y \rightarrow \beta(x, y)$ est linéaire de E dans F . Notons la $u(x)$. Il est immédiat que $x \rightarrow u(x)$ est une application linéaire de E dans $L(E, F)$ et que $b_u = \beta$.

Exercice

Montrer de même qu'il existe un isomorphisme naturel entre F_k et l'espace des applications k -linéaires de E^k dans F .

Application à $d^2 f$

Par conséquent, si f est C^2 , $d^2 f$ peut être considérée comme une application de Ω dans l'espace $B(E, F)$ des applications bilinéaires de E dans F .

Fixons une base $\mathcal{B}_E = (e_1, \dots, e_p)$ de E et notons $(\varepsilon_1, \dots, \varepsilon_p)$ sa base duale. L'application df est alors donnée par

$$df = \sum_{j=1}^{j=p} \frac{\partial f}{\partial x_j} \varepsilon_j$$

Alors $f \in C^2 \Leftrightarrow df \in C^1 \Leftrightarrow \frac{\partial f}{\partial x_j} \in C^1$ pour tout j entre 1 et p . Dans ce cas on a, pour tout $h = \sum_{j=1}^{j=p} h_j e_j \in E$

$$d^2 f(x)(h) = \sum_{j=1}^{j=p} d \left(\frac{\partial f}{\partial x_j} \right) (x)(h) \varepsilon_j$$

Mais

$$d \left(\frac{\partial f}{\partial x_j} \right) (x) = \sum_{i=1}^{i=p} \frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_j} \right) (x) \varepsilon_i = \sum_{i=1}^{i=p} \frac{\partial^2 f}{\partial x_i \partial x_j} (x) \varepsilon_i$$

donc

$$d^2 f(x)(h) = \sum_{j=1}^{j=p} \left(\sum_{i=1}^{i=p} \frac{\partial^2 f}{\partial x_i \partial x_j} (x) h_i \right) \varepsilon_j$$

Si maintenant $y = \sum_{j=1}^{j=p} y_j e_j \in E$ on a

$$d^2 f(x)(h)(y) = \sum_{j=1}^{j=p} \left(\sum_{i=1}^{i=p} \frac{\partial^2 f}{\partial x_i \partial x_j} (x) h_i \right) y_j = \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j} (x) h_i y_j$$

On voit que $d^2 f(x)$ s'identifie à la forme bilinéaire de matrice

$$H_f(x) = (d^2 f(x)(e_i, e_j)) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j} (x) \right)$$

Dans le cas général où f est à valeurs dans F , on a encore, en utilisant le calcul ci-dessus pour chaque composante de f

$$d^2 f(x)(h)(y) = \sum_{j=1}^{j=p} \left(\sum_{i=1}^{i=p} \frac{\partial^2 f}{\partial x_i \partial x_j} (x) h_i \right) y_j = \sum_{i,j} h_i y_j \frac{\partial^2 f}{\partial x_i \partial x_j} (x)$$

où cette fois les dérivées partielles secondes évaluées en x sont des vecteurs de F .

13.6.4 Caractérisation des fonctions de classe C^k

On voit facilement qu'une fonction $E \supset \Omega \xrightarrow{f} F$ est de classe C^k ssi chacune de ses composantes dans une base de F l'est.

THEOREME 13.6.1

Soient $E \supset \Omega \xrightarrow{f} F$ et \mathcal{B}_E une base de E . Pour que f soit de classe C^k sur Ω il faut et il suffit que f soit k fois dérivable sur Ω et que toutes les dérivées partielles d'ordre k (relativement à la base \mathcal{B}_E) soient continues sur Ω

Ce résultat est déjà connu pour $k = 1$. Montrons le pour $k = 2$. En écrivant f dans une base de F , on peut se ramener au cas où f est à valeurs réelles.

• Supposons f de classe C^2 . On a, (ε_j) désignant la base duale de \mathcal{B}_E

$$df(x) = \sum_{j=1}^{j=p} \frac{\partial f}{\partial x_j}(x) \varepsilon_j$$

et les $\frac{\partial f}{\partial x_j}$ sont les composantes de l'application df dans la base (ε_j) . Puisque df est C^1 , ces composantes sont elles mêmes C^1 , donc admettent des dérivées partielles d'ordre 1 continues. Ce qui prouve que toutes les dérivées partielles d'ordre 2 de f existent et sont continues.

• Supposons que toutes les dérivées partielles d'ordre 2 existent et sont continues. Chacune des fonctions $\frac{\partial f}{\partial x_j}$ a donc des dérivées partielles continues. Ces fonctions sont donc C^1 . En particulier elles sont continues, donc f est différentiable.

Comme $df = \sum_{j=1}^{j=p} \frac{\partial f}{\partial x_j} \varepsilon_j$, chacune des composantes de df est C^1 , donc df est C^1 et f est C^2 .

La démonstration dans le cas général se fait par récurrence sur k . Elle est seulement un peu plus lourde à écrire.

EXEMPLE 13.6.1

Les applications considérés dans les exmples 13.4.3, 13.4.4 et 13.4.5 sont de classe C^∞ .

13.6.5 Théorème de Schwarz

THEOREME 13.6.2

Soient $E \supset \Omega \xrightarrow{f} F$ et \mathcal{B}_E une base de E . Soient $a \in \Omega$ et i, j deux entiers distincts compris entre 1 et p . On suppose que les dérivées partielles secondes $\frac{\partial^2 f}{\partial x_i \partial x_j}$ et $\frac{\partial^2 f}{\partial x_j \partial x_i}$ sont définies dans un voisinage de a et continues en a . Alors

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a)$$

Ce théorème est admis par le programme.

COROLLAIRE 13.6.1

Soit $E \supset \Omega \xrightarrow{f} F$ de classe C^2 . Pour tout $x \in \Omega$, l'application bilinéaire $d^2 f(x) : E \times E \rightarrow F$ est symétrique.

preuve

Choisissons une base \mathcal{B}_E de E . En prenant les dérivées partielles par rapport à cette base, $d^2 f(x)$ est donnée par la formule

$$d^2 f(x)(h)(y) = \sum_{i,j} h_i y_j \frac{\partial^2 f}{\partial x_i \partial x_j}(x)$$

La symétrie résulte de l'égalité $\frac{\partial^2 f}{\partial x_i \partial x_j}(x) = \frac{\partial^2 f}{\partial x_j \partial x_i}(x)$ valable en tout x .

On peut prouver par récurrence sur k que pour une fonction f de classe C^k , dans le calcul de $\frac{\partial^k f}{\partial x_{i_1} \cdots \partial x_{i_k}}$ on peut sans changer la valeur du résultat permuter arbitrairement l'ordre des dérivations. Par conséquent, si dans la liste (i_1, \dots, i_p) , il y a

m_1 fois l'indice j_1, \dots, m_r fois l'indice j_r avec $m_1 + \dots + m_r = k$ on a

$$\frac{\partial^k f}{\partial x_{i_1} \dots \partial x_{i_k}} = \frac{\partial^k f}{\underbrace{\partial x_{j_r} \dots \partial x_{j_r}}_{m_r \text{ fois}} \dots \underbrace{\partial x_{j_1} \dots \partial x_{j_1}}_{m_1 \text{ fois}}} =: \frac{\partial^k f}{\partial x_{j_r}^{m_r} \dots \partial x_{j_1}^{m_1}}$$

la dernière expression étant une notation condensée. Ainsi pour une fonction de deux variables x et y de classe C^3 , il y a en tout quatre dérivées partielles en général distinctes à calculer :

$$\frac{\partial^3 f}{\partial x^3}, \frac{\partial^3 f}{\partial y \partial x^2}, \frac{\partial^3 f}{\partial y^2 \partial x}, \frac{\partial^3 f}{\partial y^3}$$

où, par exemple

$$\frac{\partial^3 f}{\partial y \partial x^2} = \frac{\partial^3 f}{\partial y \partial x \partial x} = \frac{\partial^3 f}{\partial x \partial y \partial x} = \frac{\partial^3 f}{\partial x \partial x \partial y}$$

13.7 Formule de Taylor Young à l'ordre 2

THEOREME 13.7.1

Soient $E \supset \Omega \xrightarrow{f} F$ de classe C^2 , $a \in \Omega$ et $r_a > 0$ tel que $B(a, r_a) \subset \Omega$. On a, pour $h \in B(0, r_a)$

$$f(a+h) = f(a) + df(a)(h) + \frac{1}{2}d^2f(a)(h, h) + \|h\|^2\varepsilon(h) \quad \text{avec} \quad \lim_{h \rightarrow 0} \varepsilon(h) = 0$$

ou, en coordonnées, avec $h = (h_1, \dots, h_p)$

$$f(a+h) = f(a) + \sum_{j=1}^{j=p} h_j \frac{\partial f}{\partial x_j}(a) + \frac{1}{2} \sum_{1 \leq i, j \leq p} h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(a) + \|h\|^2\varepsilon(h)$$

Ce théorème est admis. Une preuve est donnée en appendice.

13.8 Problèmes d'extremums

13.8.1 Définitions

DEFINITION 13.8.1 (Extremum local)

Soient D un sous ensemble de E , $x_0 \in D$ et $f : D \rightarrow \mathbb{R}$.

1. On dit que f présente un maximum local en x_0 si il existe un voisinage U de x_0 dans E tel que

$$x \in U \cap D \Rightarrow f(x) \leq f(x_0)$$

2. On dit que f présente un maximum local strict en x_0 si il existe un voisinage U de x_0 dans E tel que

$$(x \in U \text{ et } x \neq x_0) \Rightarrow f(x) < f(x_0)$$

Définitions analogues pour un minimum.

f présente un extremum local (resp. strict) en x_0 si f a en x_0 un maximum ou un minimum local (resp. strict).

Dans cette définition, D est un sous ensemble quelconque de E , non nécessairement ouvert.

DEFINITION 13.8.2

Soient D un sous ensemble de E , $x_0 \in D$ et $f : D \rightarrow \mathbb{R}$.

On dit que f a un maximum (resp. un maximum strict) en x_0 si $\forall x \in D$ $f(x) \leq f(x_0)$ (resp. si $\forall x \in D$, $x \neq x_0 \Rightarrow f(x) < f(x_0)$).

Définition analogues pour un minimum.

Un maximum est évidemment un maximum local. Si f est continue sur D et si D est compact, on sait que f a toujours au moins un minimum et un maximum sur D .

DEFINITION 13.8.3

Soient Ω un ouvert de E , $a \in \Omega$ et $f : \Omega \rightarrow \mathbb{R}$ différentiable sur Ω . On dit que a est un point critique de f si $df(a) = 0$. Si E est muni d'une structure euclidienne, a est un point critique de f si $\nabla f(a) = 0$.

13.8.2 Extremums : conditions nécessaires

THEOREME 13.8.1

Soient Ω un ouvert de E , $a \in \Omega$ et $f : \Omega \rightarrow \mathbb{R}$ différentiable sur Ω . Si f présente un extremum local en a , alors a est un point critique de f .

preuve

Soit $e \in E$. Puisque Ω est ouvert, la fonction d'une variable réelle $t \rightarrow f(a + te)$ est définie dans un intervalle ouvert contenant 0 et présente en 0 un extremum local. Comme elle est dérivable en 0, sa dérivée en 0 est nulle. Mais cette dérivée vaut $df(a)(e)$. Donc $\forall e \in E$ $df(a)(e) = 0$ ce qui prouve que $df(a) = 0$.

Conséquence :

Soit D un sous ensemble quelconque de E et $f : D \rightarrow \mathbb{R}$ différentiable sur l'intérieur de D . Si f présente un extrémum local en $x_0 \in D$, il y a deux possibilités :

$$\begin{aligned} \text{ou } x_0 \in \overset{\circ}{D} \text{ et } df(x_0) = 0 \\ \text{ou } x_0 \in D \setminus \overset{\circ}{D} \end{aligned}$$

Soient Ω un ouvert de E et $f : \Omega \rightarrow \mathbb{R}$ de classe C^2 . Soit $a \in \Omega$. La formule de Taylor en a s'écrit

$$f(a + h) = f(a) + df(a)(h) + \frac{1}{2}d^2f(a)(h, h) + \|h\|^2\varepsilon(h)$$

D'après le théorème de Schwarz, l'application $(h, k) \rightarrow d^2f(a)(h, k)$ est une forme bilinéaire symétrique et $h \rightarrow d^2f(a)(h, h)$ est la forme quadratique associée. On la notera $H_f(a)(h)$. $H_f(a)$ s'appelle la Hessienne de f en a . En coordonnées

$$H_f(a)(h) = \sum_{1 \leq i, j \leq p} \frac{\partial^2 f}{\partial x_i \partial x_j}(a) h_i h_j$$

THEOREME 13.8.2

Soient Ω un ouvert de E , $a \in \Omega$ et $f : \Omega \rightarrow \mathbb{R}$ de classe C^2 .

Si f présente un maximum local en a la forme quadratique $H_f(a)$ est négative.

Si f présente un minimum local en a , la forme quadratique $H_f(a)$ est positive.

preuve

Supposons que f présente un extremum local en a . Alors a est un point critique de f . Soit $h \in E$, $h \neq 0$. La formule de Taylor appliquée au vecteur th , $t \in \mathbb{R}$ s'écrit

$$f(a + th) = f(a) + \frac{1}{2}t^2 H_f(a)(h) + t^2 \|h\|^2 \varepsilon(th)$$

donc

$$H_f(a)(h) = \lim_{t \rightarrow 0} \frac{2}{t^2} (f(a + th) - f(a))$$

Si f présente un maximum (resp. minimum) local en a , le second membre est négatif (resp. positif) ou nul pour $|t|$ assez petit, d'où la conclusion.

13.8.3 Extremums : condition suffisante

DEFINITION 13.8.4

Soient Ω un ouvert de E , $f : \Omega \rightarrow \mathbb{R}$ de classe C^2 et $a \in \Omega$ un point critique de f . Le point critique est dit non dégénéré si la forme quadratique $H_f(a)$ est non dégénérée (i.e. de rang $p = \dim E$).

THEOREME 13.8.3

Soient Ω un ouvert de E , $f : \Omega \rightarrow \mathbb{R}$ de classe C^2 et $a \in \Omega$ un point critique de f .

1. Si la forme quadratique $H_f(a)$ est définie négative, f présente en a un maximum local strict.
2. Si la forme quadratique $H_f(a)$ est définie positive, f présente en a un minimum local strict.
3. Si la forme quadratique $H_f(a)$ n'est pas de signe constant, f ne présente pas d'extremum local en a .

preuve

Le point 3. résulte de la condition nécessaire d'extremum local.

On peut faire la démonstration en coordonnées, ce qui revient à supposer $E = \mathbb{R}^p$. Munissons \mathbb{R}^p de sa structure euclidienne usuelle. Supposons $H_f(a)$ définie positive. La matrice $\left(\frac{\partial^2 f}{\partial x_i \partial x_j}(a)\right)$ est diagonalisable et ses valeurs propres sont strictement positives. Soit λ la plus petite de ces valeurs propres. On a alors $H_f(a)(h) \geq \lambda \|h\|^2$. La formule de Taylor donne alors

$$f(a+h) - f(a) \geq \left(\frac{\lambda}{2} - \varepsilon(h)\right) \|h\|^2$$

Puisque $\varepsilon(h)$ tend vers 0 avec h , il existe $\rho > 0$ tel que $B(a, \rho) \subset \Omega$ et tel que $\|h\| < \rho \Rightarrow |\varepsilon(h)| < \frac{\lambda}{4}$. Alors, pour $x \in B(a, \rho)$ on a $f(x) - f(a) \geq \frac{\lambda}{4} \|x - a\|^2$ ce qui prouve que f admet en a un minimum local strict.

Si $H_f(a)$ est définie négative, on applique ce qui précède à $-f$ pour conclure.

Conclusion : Si a est un point critique de f de classe C^2 , le seul cas d'ambiguïté est le cas où la forme quadratique $H_f(a)$ est de signe constant et dégénérée.

Si $H_f(a)$ est non dégénérée de signe non constant, on dit que le point critique a est un point col ou selle.

Cas de la dimension 2

Soit Ω un ouvert de \mathbb{R}^2 et $f : \Omega \rightarrow \mathbb{R}$. On notera (x, y) les coordonnées dans \mathbb{R}^2 . Soit $(a, b) \in \mathbb{R}^2$.

Notations de Monge : on note

$$p = \frac{\partial f}{\partial x}(a, b) \quad q = \frac{\partial f}{\partial y}(a, b) \quad r = \frac{\partial^2 f}{\partial x^2}(a, b) \quad s = \frac{\partial^2 f}{\partial x \partial y}(a, b) \quad t = \frac{\partial^2 f}{\partial y^2}(a, b)$$

La formule de Taylor s'écrit

$$f(a+h, b+k) = f(a, b) + (ph + qk) + \frac{1}{2}(rh^2 + 2shk + tk^2) + (h^2 + k^2)\varepsilon(h, k) \quad \lim_{(h,k) \rightarrow (0,0)} \varepsilon(h, k) = 0$$

La matrice de la forme quadratique $H_f(a, b)$ s'écrit $\begin{pmatrix} r & s \\ s & t \end{pmatrix}$. Cette forme quadratique est non dégénérée ssi $rt - s^2 \neq 0$. Si $rt - s^2 > 0$ les deux valeurs propres sont de même signe. Comme $r = H_f(a)(1, 0)$ si $r > 0$ la forme est définie positive et si $r < 0$ elle est définie négative. D'où les critères en dimension 2 :

- (a, b) point critique de $f \Leftrightarrow p = q = 0$
- Si (a, b) est un point critique, alors
 - si $rt - s^2 > 0$ et $r > 0$, (a, b) est un minimum local strict.
 - si $rt - s^2 > 0$ et $r < 0$, (a, b) est un maximum local strict.
 - si $rt - s^2 < 0$ le point critique est de type selle et f n'a pas d'extremum en (a, b) .

Dans ce qui précède on peut évidemment remplacer r par s . Notons que si $rt - s^2 = 0$ la forme quadratique $H_f(a, b)$ est de rang 0 ou 1, donc est de signe constant. dans ce cas on ne peut pas conclure en utilisant les théorèmes ci dessus et il faut faire une étude plus détaillée.

13.8.4 Exemples

Exemple 1 Etudier les extremums de $f(x, y) = xy \ln(x + y)$.

f est C^∞ sur $D = \{(x, y) \in \mathbb{R}^2 \mid x + y > 0\}$.

$$\frac{\partial f}{\partial x}(x, y) = y \ln(x + y) + \frac{xy}{x + y} \quad \frac{\partial f}{\partial y}(x, y) = x \ln(x + y) + \frac{xy}{x + y}$$

$$\frac{\partial^2 f}{\partial x^2} = \frac{2y}{x+y} - \frac{xy}{(x+y)^2} \quad \frac{\partial^2 f}{\partial x \partial y} = \ln(x+y) + 1 - \frac{xy}{(x+y)^2} \quad \frac{\partial^2 f}{\partial y^2} = \frac{2x}{x+y} - \frac{xy}{(x+y)^2}$$

Les points critiques sont solutions du système

$$\begin{cases} (x-y) \ln(x+y) = 0 \\ (x+y) \ln(x+y) + 2 \frac{xy}{x+y} = 0 \end{cases}$$

On a donc deux cas à envisager :

- $x = y$ et $2x \ln(2x) + x = 0$ soit $x_0 = y_0 = \frac{1}{2}e^{-1/2}$. En ce point, $r = t = 3/4$, $s = 1/4$; donc $rt - s^2 > 0$ avec $r > 0$ et le point (x_0, y_0) est un minimum local strict.
- $x + y = 1$ et $xy = 0$ ce qui donne les deux points $A = (1, 0)$ et $B = (0, 1)$. En A on a $r = 0$, $s = 1$, $t = 2$, en B , on a $r = 2$, $s = 1$, $t = 0$ et dans les deux cas $rt - s^2 < 0$, donc A et B sont des points selles. Il est clair que A et B sont de même nature car la fonction f vérifie $f(x, y) = f(y, x)$ pour tout $(x, y) \in D$.

Exemple 2

Pour établir qu'un extremum local est un point critique, on a utilisé le fait que si f avait en $a \in \Omega$ par exemple un maximum local, alors la fonction d'une variable réelle $t \rightarrow f(a + th)$ avait en $t = 0$ un maximum local.

L'exemple qui suit montre que la réciproque de cette propriété n'est pas vraie : soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par $f(x, y) = (y - x^2)(y - 3x^2) = y^2 - 4yx^2 + 3x^4$. L'écriture de f montre que le DL₂ en $(0, 0)$ est $f(x, y) = y^2 + o(x^2 + y^2)$. Par conséquent $(0, 0)$ est un point critique et la hessienne en 0 est la forme quadratique $(x, y) \rightarrow y^2$ qui est positive dégénérée.

- Soit $(a, b) \neq (0, 0)$. Soit $\varphi(t) = f(ta, tb) = t^2(b^2 - 4tba^2 + 3t^2a^4)$. Si $b \neq 0$, $\varphi(t) \underset{t \rightarrow 0}{\sim} b^2 t^2$ donc $\varphi(t) > 0$ pour t non nul assez petit. Si $b = 0$, alors $a \neq 0$ et $\varphi(t) = 3a^4 t^4$ donc là encore $\varphi(t) > 0$ pour t non nul assez petit. Il s'ensuit que la restriction de f à la droite D de vecteur directeur (a, b) passant par $(0, 0)$ présente un minimum local strict à l'origine.
- Pourtant, l'origine n'est pas un minimum local. Soit V un voisinage de $(0, 0)$. Pour $|x|$ assez petit, $(x, 2x^2) \in V$ et $f(x, 2x^2) = -x^4 < 0 = f(0, 0)$ pour $x \neq 0$.

13.8.5 Complément : étude des points selle

Ce petit complément n'est pas explicitement au programme, mais est facile et très instructif géométriquement.

Soit Ω un ouvert de \mathbb{R}^n , $f : \Omega \rightarrow \mathbb{R}$ de classe C^2 . Considérons un point critique non dégénéré de type selle de f . On peut toujours, quitte à faire un changement d'origine dans \mathbb{R}^n supposer que ce point critique est l'origine. On peut aussi supposer pour simplifier que $f(0) = 0$. On munit \mathbb{R}^n de la norme euclidienne usuelle. La formule de Taylor en 0 s'écrit alors

$$f(x) = \frac{1}{2} \sum_{i,j} \frac{\partial^2 f}{\partial x_i \partial x_j}(0) x_i x_j + \|x\|^2 \varepsilon(x)$$

où (x_1, \dots, x_n) sont les coordonnées de x dans la base canonique.

Soit $H = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(0) \right)$. H est une matrice symétrique réelle. La forme quadratique de matrice H est non dégénérée. Soit (p, q) sa signature. On a $p + q = n$. Soient $\lambda_1, \dots, \lambda_p$ les valeurs propres strictement positives de H et $-\lambda_{p+1}, \dots, -\lambda_n$ les valeurs propres strictement négatives. Soit enfin $\mathcal{B} = (e_1, \dots, e_n)$ une base orthonormée de \mathbb{R}^n formée de vecteurs propres de H et telle que

$$H(X_1 e_1 + \dots + X_n e_n) = \lambda_1 X_1^2 + \dots + \lambda_p X_p^2 - \lambda_{p+1} X_{p+1}^2 - \dots - \lambda_n X_n^2$$

Soit alors (X_1, \dots, X_n) les coordonnées de x dans la base \mathcal{B} . La formule de Taylor devient

$$f(x) = \frac{1}{2} (\lambda_1 X_1^2 + \dots + \lambda_p X_p^2 - \lambda_{p+1} X_{p+1}^2 - \dots - \lambda_n X_n^2) + (X_1^2 + \dots + X_n^2) \varepsilon(x)$$

Soit $F = \text{Vect}(e_1, \dots, e_p)$ et $G = \text{Vect}(e_{q+1}, \dots, e_n)$.

Soit $x \in F$. On a $X_{p+1} = \dots = X_n = 0$ donc $f(x) = \frac{1}{2} (\lambda_1 X_1^2 + \dots + \lambda_p X_p^2) + (X_1^2 + \dots + X_p^2) \varepsilon(x)$

La restriction $f|_F$ de f à F présente donc en 0 un minimum local strict.

Soit $x \in G$. On a $X_1 = \dots = X_p = 0$ donc $f(x) = -\frac{1}{2} (\lambda_{p+1} X_{p+1}^2 + \dots + \lambda_n X_n^2) + (X_{p+1}^2 + \dots + X_n^2) \varepsilon(x)$

La restriction $f|_G$ de f à G présente donc en 0 un maximum local strict.

On a donc prouvé

THEOREME 13.8.4

Soit Ω un ouvert de \mathbb{R}^n , $f : \Omega \rightarrow \mathbb{R}$ de classe C^2 et $a \in \Omega$ un point critique de type selle. Il existe une décomposition de \mathbb{R}^n en somme directe de deux sous espaces orthogonaux $\mathbb{R}^n = F \oplus G$ telle que $f|_{a+F}$ (resp. $f|_{a+G}$) présente en a un minimum (resp. un maximum) local strict.

Ceci justifie la dénomination de point col ou selle. Dans le cas $n = 2$, et $a = (0, 0)$, $f(0, 0) = 0$ pour simplifier, on peut trouver deux directions orthogonales OX et OY telles que $f(x, 0) > 0$ pour $x \neq 0$ assez petit et $f(0, y) < 0$ pour $y \neq 0$ assez petit. La surface Σ d'équation $z = f(x, y)$ ressemble à une selle de cheval au voisinage de $(0, 0)$. Un prototype est la quadrique d'équation $z = x^2 - y^2$ (paraboloïde hyperbolique).

13.8.6 Application : position d'une surface par rapport à son plan tangent

On se donne Ω un ouvert de \mathbb{R}^2 , $f : \Omega \rightarrow \mathbb{R}$ une fonction de classe C^2 . Soit $(x_0, y_0) \in \Omega$ et $z_0 = f(x_0, y_0)$. Soit $M_0 = (x_0, y_0, z_0)$ et Σ la surface d'équation $z = f(x, y)$. On a vu que le plan tangent à Σ en M_0 était le plan affine Π d'équation $z = z_0 + p(x - x_0) + q(y - y_0)$ en utilisant les notations de Monge au point (x_0, y_0) . On se propose d'étudier la position de Σ par rapport à Π au voisinage de M_0 . Notons $m = (x, y)$ un point de Ω et M le point de Σ de coordonnées $(x, y, z_M = f(x, y))$. Le point P de Π qui se projette sur m a pour cote $z_P = z_0 + p(x - x_0) + q(y - y_0)$. On a donc

$$z_M - z_P = f(x, y) - f(x_0, y_0) - p(x - x_0) - q(y - y_0)$$

D'après la formule de Taylor et l'étude de signe faite précédemment, on a, les notations de Monge r, s, t étant prises en (x_0, y_0)

Si $rt - s^2 > 0$ et $r > 0$, $z_M - z_P > 0$ pour (x, y) assez voisin de (x_0, y_0) . La surface est au dessus de son plan tangent au voisinage de M_0 .

Si $rt - s^2 > 0$ et $r < 0$, $z_M - z_P < 0$ pour (x, y) assez voisin de (x_0, y_0) . La surface est au dessous de son plan tangent au voisinage de M_0 .

Si $rt - s^2 < 0$ $z_M - z_P$ n'est pas de signe constant dans un voisinage de M_0 et dans tout voisinage de ce point, il y a des points de Σ au dessus de Π et des points de Σ au dessous de Π .

En utilisant le complément précédent, on voit qu'il existe deux droites affines D et D' de \mathbb{R}^2 , passant par $m_0 = (x_0, y_0)$ et perpendiculaires entre elles telles que les points de la surface se projetant en un point de D (resp. de D') soit au dessus (resp. au dessous) du plan tangent Π en M_0 .

Dans le cas $rt - s^2 = 0$, il faut faire une étude spécifique.

Appendice : preuve de quelques théorèmes

Les démonstrations ci dessous ne sont pas au programme. Elles sont cependant instructives mettant en particulier en évidence l'utilisation des fonctions partielles et de l'inégalité usuelle des accroissements finis.

A Preuve du théorème de caractérisation des fonctions de classe C^1

On va ici prouver le résultat suivant :

Soit Ω un ouvert de \mathbb{R}^2 et $f : \Omega \rightarrow \mathbb{R}$ admettant des dérivées partielles par rapport aux deux variables continues sur Ω . Alors f est différentiable sur Ω .

On munit \mathbb{R}^2 de la norme $\|(x, y)\| = \max\{|x|, |y|\}$. Soit $a = (x_0, y_0) \in \Omega$ et $r > 0$ tel que $B := B((x_0, y_0), r) \subset \Omega$. Avec cette norme $B =]x_0 - r, x_0 + r[\times]y_0 - r, y_0 + r[$.

Il s'agit de montrer que f est différentiable en (x_0, y_0) . Si c'est le cas, sa différentielle est nécessairement l'application linéaire $L : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par

$$L(h, k) = h \frac{\partial f}{\partial x}(x_0, y_0) + k \frac{\partial f}{\partial y}(x_0, y_0)$$

Posons, pour $|h| < r$ et $|k| < r$

$$\Delta(h, k) = f(x_0 + h, y_0 + k) - f(x_0, y_0) - \left(h \frac{\partial f}{\partial x}(x_0, y_0) + k \frac{\partial f}{\partial y}(x_0, y_0) \right)$$

Soit $\varepsilon > 0$. Le résultat sera montré si on trouve un $\eta > 0$ tel que

$$(|h| < \eta \text{ et } |k| < \eta) \Rightarrow |\Delta(h, k)| \leq \varepsilon \|(h, k)\| = \varepsilon \max\{|h|, |k|\}$$

• On dispose, par continuité des dérivées partielles d'un $\eta \in]0, r[$ tel que

$$(1) \quad \max\{|u|, |v|\} < \eta \Rightarrow \left| \frac{\partial f}{\partial x}(x_0 + u, y_0 + v) - \frac{\partial f}{\partial x}(x_0, y_0) \right| \leq \frac{\varepsilon}{2} \quad \text{et} \quad \left| \frac{\partial f}{\partial y}(x_0 + u, y_0 + v) - \frac{\partial f}{\partial y}(x_0, y_0) \right| \leq \frac{\varepsilon}{2}$$

• On écrit $\Delta(h, k) = \Delta_1(h, k) + \Delta_2(h, k)$ avec

$$\Delta_1(h, k) = f(x_0 + h, y_0 + k) - f(x_0, y_0 + k) - h \frac{\partial f}{\partial x}(x_0, y_0)$$

$$\Delta_2(k) = f(x_0, y_0 + k) - f(x_0, y_0) - k \frac{\partial f}{\partial y}(x_0, y_0)$$

et on va majorer séparément ces deux termes.

On suppose à partir de maintenant que $|h| < \eta$ et $|k| < \eta$.

• Considérons la fonction $\varphi :]-\eta, \eta[\rightarrow \mathbb{R}$ définie par

$$\varphi(t) = f(x_0 + t, y_0 + k) - f(x_0, y_0 + k) - t \frac{\partial f}{\partial x}(x_0, y_0)$$

Cette fonction est dérivable et $\varphi'(t) = \frac{\partial f}{\partial x}(x_0 + t, y_0 + k) - \frac{\partial f}{\partial x}(x_0, y_0)$.

Puisque $|k| < \eta$, on a, d'après l'inégalité (1) $\forall t \in]-\eta, \eta[\quad |\varphi'(t)| \leq \frac{\varepsilon}{2}$. Pour $|h| < \eta$, l'inégalité des accroissements finis donne $|\varphi(h) - \varphi(0)| \leq |h| \frac{\varepsilon}{2}$ soit

$$|\Delta_1(h, k)| \leq \frac{\varepsilon}{2} |h| \leq \frac{\varepsilon}{2} \|(h, k)\|$$

• Considérons maintenant la fonction $\psi :]-\eta, \eta[\rightarrow \mathbb{R}$ définie par

$$\psi(t) = f(x_0, y_0 + t) - f(x_0, y_0) - t \frac{\partial f}{\partial y}(x_0, y_0)$$

ψ est dérivable et $\psi'(t) = \frac{\partial f}{\partial y}(x_0, y_0 + t) - \frac{\partial f}{\partial y}(x_0, y_0)$. Toujours d'après (1), on a $|\psi'(t)| \leq \frac{\varepsilon}{2}$ et d'après l'inégalité des accroissements finis $|\psi(k) - \psi(0)| \leq \frac{\varepsilon}{2} |k|$ soit encore

$$|\Delta_2(k)| \leq \frac{\varepsilon}{2} |k| \leq \frac{\varepsilon}{2} \|(h, k)\|$$

• En ajoutant, on obtient $\|(h, k)\| < \eta \Rightarrow |\Delta(h, k)| \leq \varepsilon \|(h, k)\|$. ■

La preuve dans le cas où $p > 2$ est analogue : on écrit $f(x_1 + h_1, \dots, x_p + h_p) - f(x_1, \dots, x_p)$ comme une somme de p différences où dans chacune de ces différences seule une coordonnée varie et on majore de la même manière en appliquant p fois l'inégalité des accroissements finis à des fonctions d'une variable réelle.

Si f est à valeurs dans \mathbb{R}^n , on applique ce qui précède à chacune de ces composantes pour prouver que f est différentiable. Enfin, si les espaces de départ et d'arrivée sont des espaces quelconques E et F , on se ramène au cas de \mathbb{R}^p et \mathbb{R}^n en fixant des bases de E et de F .

B Preuve du théorème de Schwarz

On peut supposer, en travaillant sur les composantes de f que f est à valeurs réelles. D'autre part, comme les dérivées partielles croisées en a en les variables x_i et x_j ne font intervenir que la fonction de deux variables

$$(x_i, x_j) \rightarrow f(a_1, \dots, a_{i-1}, x_i, a_{i+1}, \dots, a_{j-1}, x_j, a_{j+1}, \dots, a_p)$$

on peut supposer que f est une fonction de deux variables seulement. Enfin quitte à composer avec une translation, on peut supposer $a = (0, 0)$

On va donc prouver le résultat suivant :

Soient Ω un ouvert de \mathbb{R}^2 contenant $(0, 0)$, $f : \Omega \rightarrow \mathbb{R}$ admettant des dérivées partielles secondes croisées $\frac{\partial^2 f}{\partial y \partial x}$ et $\frac{\partial^2 f}{\partial x \partial y}$ dans un voisinage de $(0, 0)$ et continues en $(0, 0)$. Alors $\frac{\partial^2 f}{\partial y \partial x}(a, 0) = \frac{\partial^2 f}{\partial x \partial y}(a, 0)$

Posons $A = \frac{\partial^2 f}{\partial x \partial y}(0, 0)$ et $B = \frac{\partial^2 f}{\partial y \partial x}(0, 0)$. On munit \mathbb{R}^2 de la norme $\|(x, y)\| = \max\{|x|, |y|\}$. Soit $r > 0$ tel que $B(0, r) =]-r, r[\times]-r, r[\subset \Omega$ et telles que les deux dérivées partielles croisées soient définies et continues sur $B(0, r)$.

• Principe de la démonstration

On définit une fonction $g : B(0, r) \rightarrow \mathbb{R}$ par $g(x, y) = f(x, y) - f(x, 0) - f(0, y) + f(0, 0)$.

On va montrer que pour tout $\varepsilon > 0$ il existe $\eta > 0$ tel que $\|(u, v)\| < \eta \Rightarrow \begin{cases} |g(u, v) - Auv| \leq \varepsilon |uv| \\ |g(u, v) - Buv| \leq \varepsilon |uv| \end{cases}$

Il en résultera par différence que $\forall \varepsilon > 0 \quad |A - B| \leq 2\varepsilon$ donc que $A = B$.

• Mise en oeuvre

L'hypothèse de continuité des dérivées secondes assure l'existence d'un $\eta > 0$ tel qu'on ait

$$\left| \frac{\partial^2 f}{\partial x \partial y}(u, v) - A \right| < \varepsilon \quad \text{et} \quad \left| \frac{\partial^2 f}{\partial y \partial x}(u, v) - B \right| < \varepsilon$$

pour $\|(x, y)\| < \eta$

Soit $W = B(0, \eta)$. Définissons, pour $(x, y) \in W$, $\varphi(x, y) = g(x, y) - Bxy$. On a

$$\frac{\partial \varphi}{\partial x}(x, y) = \frac{\partial f}{\partial x}(x, y) - \frac{\partial f}{\partial x}(x, 0) - By$$

donc en particulier

$$\frac{\partial \varphi}{\partial x}(x, 0) = 0$$

puis

$$\left| \frac{\partial}{\partial y} \left(\frac{\partial \varphi}{\partial x} \right) (x, y) \right| = \left| \frac{\partial^2 f}{\partial y \partial x} (x, y) - B \right| < \varepsilon$$

Fixons $(u, v) \in W$.

Soit x tel que $|x| < \eta$. Considérons la fonction $t \rightarrow h(t) = \frac{\partial \varphi}{\partial x}(x, t)$. Cette fonction est dérivable et sa dérivée pour t entre 0 et v est majorée par ε . Donc d'après l'inégalité des accroissements finis, $|h(v) - h(0)| \leq \varepsilon|v|$ soit $\left| \frac{\partial \varphi}{\partial x}(x, v) \right| \leq \varepsilon|v|$. et cette inégalité est valable pour tout x tel que $|x| < \eta$.

Considérons maintenant la fonction $t \rightarrow k(t) = \varphi(t, v)$. k est dérivable, $|k'(t)| = \left| \frac{\partial \varphi}{\partial x}(t, v) \right| \leq \varepsilon|v|$ pour $t \in]-\eta, \eta[$. Donc, d'après l'inégalité des accroissements finis $|k(u) - k(0)| \leq \varepsilon|v||u|$ soit $|g(u, v) - Buv| \leq \varepsilon|u||v|$.

On procède de la même manière avec la fonction $\psi(x, y) = g(x, y) - Axy$ en dérivant d'abord en y puis en x ce qui fournit l'autre inégalité. ■

C Preuve de la formule de Taylor

En prenant des coordonnées, il suffit d'établir la formule pour $\mathbb{R}^p \supset \Omega \xrightarrow{f} F$. Il s'agit d'un développement limité. La formule sera vraie si elle l'est pour chacune des composantes de f dans une base \mathcal{B}_F de F . On va donc montrer

Soient Ω un ouvert de \mathbb{R}^p , $a \in \Omega$ et $f : \Omega \rightarrow \mathbb{R}$ de classe C^2 . Alors

$$f(a+h) = f(a) + \sum_{j=1}^{j=p} h_j \frac{\partial f}{\partial x_j}(a) + \frac{1}{2} \sum_{1 \leq i, j \leq p} h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(a) + \|h\|^2 \varepsilon(h)$$

Munissons \mathbb{R}^p de la norme $\|(x_1, \dots, x_p)\| = \max\{|x_j|\}$. Soit $r > 0$ tel que $B(a, r) \subset \Omega$.

Définissons la fonction $\varepsilon : B(0, r) \rightarrow \mathbb{R}$ par $\varepsilon(0) = 0$ et

$$\varepsilon(h) = \frac{1}{\|h\|^2} \left(f(a+h) - f(a) + \sum_{j=1}^{j=p} h_j \frac{\partial f}{\partial x_j}(a) - \frac{1}{2} \sum_{1 \leq i, j \leq p} h_i h_j \frac{\partial^2 f}{\partial x_i \partial x_j}(a) \right) \quad \text{pour } h = (h_1, \dots, h_p) \neq 0$$

Il s'agit de montrer que $\lim_{h \rightarrow 0} \varepsilon(h) = 0$.

Fixons h tel que $\|h\| < r$. Définissons $\varphi : [0, 1] \rightarrow \mathbb{R}$ par $\varphi(t) = f(a+th)$. La formule de Taylor avec reste intégrale s'écrit

$$\varphi(t) = \varphi(0) + t\varphi'(0) + \int_0^1 (1-t)\varphi''(t)dt$$

Or

$$\varphi'(t) = \sum_{j=1}^{j=p} \frac{\partial f}{\partial x_j}(a+th)h_j \quad \text{et} \quad \varphi''(t) = \sum_{1 \leq i, j \leq p} \frac{\partial^2 f}{\partial x_i \partial x_j}(a+th)h_i h_j$$

Soit $\varepsilon_0 > 0$. Par continuité des dérivées partielles secondes, il existe $\eta > 0$ tel que

$$\|x\| < \eta \Rightarrow \forall i, j \quad \left| \frac{\partial^2 f}{\partial x_i \partial x_j}(x) - \frac{\partial^2 f}{\partial x_i \partial x_j}(a) \right| \leq \frac{\varepsilon_0}{p^2}$$

Supposons $\|h\| < \eta$, on a alors $|\varphi''(t) - \varphi''(0)| \leq \frac{\varepsilon_0}{p^2} \sum_{1 \leq i, j \leq p} |h_i h_j| \leq \varepsilon_0 \|h\|^2$. On en déduit

$$\left| \int_0^1 (1-t)(\varphi''(t) - \varphi''(0))dt \right| \leq \varepsilon_0 \|h\|^2 \int_0^1 (1-t)dt \leq \varepsilon_0 \|h\|^2$$

ce qui s'écrit aussi

$$\left| \varphi(1) - \varphi(0) - \varphi'(0) - \frac{1}{2}\varphi''(0) \right| \leq \varepsilon_0 \|h\|^2$$

soit en remplaçant $\|h\|^2 |\varepsilon(h)| \leq \varepsilon_0 \|h\|^2$. Finalement, pour tout $\varepsilon_0 > 0$, on a trouvé $\eta > 0$ tel que $\|h\| < \eta \Rightarrow |\varepsilon(h)| \leq \varepsilon_0$, ce qu'il fallait démontrer.

14

Difféomorphismes, fonctions implicites

14.1 Difféomorphismes

DEFINITION 14.1.1

Soient E et F deux espaces vectoriels normés de dimension finie, U un ouvert non vide de E et V un ouvert non vide de F . Soit $k \in \mathbb{N}^* \cup \{\infty\}$. Un C^k difféomorphisme de U sur V est une bijection $\Phi : U \rightarrow V$ de classe C^k telle que la bijection réciproque $\Phi^{-1} : V \rightarrow U$ soit aussi de classe C^k .
 U et V sont dits C^k difféomorphes si il existe un difféomorphisme de classe C^k de U sur V .

Si Φ est un difféomorphisme de U sur V , Φ et Φ^{-1} sont continues, donc Φ est un homéomorphisme de U sur V . Deux ouverts difféomorphes sont donc nécessairement homéomorphes.

THEOREME 14.1.1

Soient E et F deux espaces vectoriels normés de dimension finie, U un ouvert non vide de E et V un ouvert non vide de F . Si U et V sont difféomorphes, on a $\dim E = \dim F$

preuve

Soit $\Phi : U \rightarrow V$ un difféomorphisme. Soient $a \in U, b = \Phi(a) \in V$. Par hypothèse Φ et Φ^{-1} sont de classe C^1 au moins et

$$\Phi^{-1} \circ \Phi = Id_U$$

donc en différentiant en a

$$d\Phi^{-1}(b) \circ d\Phi(a) = d(Id_U) = id_E$$

et de même, en différentiant $\Phi \circ \Phi^{-1} = id_V$ en b on obtient $d\Phi(a) \circ d\Phi^{-1}(b) = d(id_V) = id_F$.

Par conséquent $d\Phi(a)$ est un isomorphisme de E sur F d'inverse $d\Phi^{-1}(\Phi(a))$. En particulier $\dim E = \dim F$.

Notons qu'au passage on a montré que si Φ est un difféomorphisme de U sur V avec $\Phi(x) = y$ on a

$$(d\Phi(x))^{-1} = d\Phi^{-1}(\Phi(x)) \quad \text{ou} \quad d\Phi^{-1}(y) = (d\Phi(\Phi^{-1}(y)))^{-1} \quad (14.1)$$

THEOREME 14.1.2

Soient E et F deux espaces vectoriels normés de dimension finie, U un ouvert non vide de E et V un ouvert non vide de F . Soit $\varphi : U \rightarrow V$ une application de classe C^k qui est un C^1 difféomorphisme de U sur V . Alors φ est un C^k difféomorphisme de U sur V .

preuve

Posons $\psi = \varphi^{-1}$. Par hypothèse, ψ existe et est de classe C^1 sur Ω . Il s'agit de vérifier que ψ est C^k . Or $d\psi = (d\varphi \circ \psi)^{-1}$. Si on suppose que ψ est C^j avec $j < k$, comme $d\varphi$ est C^{k-1} la composée $d\varphi \circ \psi$ est C^j . Par ailleurs, l'application $Iso(E, F) \rightarrow Iso(F, E)$ qui à un isomorphisme linéaire f associe f^{-1} est C^∞ (exemple 13.6.1) donc $d\psi$ est C^j et ψ est C^{j+1} . Par récurrence finie, on en déduit que ψ est C^k .

Changement de variables

Soient U et V comme ci dessus et Φ un C^k difféomorphisme de U sur V . Soit G un espace vectoriel normé de dimension finie et $C^k(U, G)$ (resp. $C^k(V, G)$) l'espace vectoriel des applications de classe C^k de U (resp. V) dans G de classe C^k . L'application $C^k(V, G) \rightarrow C^k(U, G)$ qui à $f \in C^k(V, G)$ associe $f \circ \Phi$ est une bijection de $C^k(V, G)$ sur $C^k(U, G)$, la bijection réciproque étant l'application $g \rightarrow g \circ \Phi^{-1}$.

Exemple d'utilisation : équation des ondes.

Soit $c > 0$. On se propose de trouver toutes les fonctions $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ vérifiant l'équation dite équation des ondes

$$(E) \quad \forall (x, t) \in \mathbb{R}^2 \quad \frac{\partial^2 f}{\partial t^2}(x, t) - c^2 \frac{\partial^2 f}{\partial x^2}(x, t) = 0$$

A cette fin, on considère l'application $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ définie par $\Phi(x, t) = (x + ct, x - ct)$. Φ est un isomorphisme linéaire de \mathbb{R}^2 , donc un C^∞ difféomorphisme et $\Phi^{-1}(u, v) = \left(\frac{u+v}{2}, \frac{u-v}{2c}\right)$. Φ induit donc une bijection de $C^2(\mathbb{R}^2, \mathbb{R})$ sur $C^2(\mathbb{R}^2, \mathbb{R})$ définie par $G \rightarrow g = G \circ \Phi$. Soit $f \in C^2(\mathbb{R}, \mathbb{R})$ et $F = f \circ \Phi^{-1}$. On a donc

$$f(x, y) = F(x + ct, x - ct) = F(u, v)$$

en posant $u = x - ct, v = x + ct$.

$$\frac{\partial f}{\partial x}(x, t) = \frac{\partial F}{\partial u}(u, v) + \frac{\partial F}{\partial v}(u, v) \quad \frac{\partial f}{\partial t}(x, t) = -c \frac{\partial F}{\partial u}(u, v) + c \frac{\partial F}{\partial v}(u, v)$$

et, en omettant (x, t) et (u, v)

$$\begin{aligned} \frac{\partial^2 f}{\partial x^2} &= \frac{\partial^2 F}{\partial u^2} + 2 \frac{\partial^2 F}{\partial u \partial v} + \frac{\partial^2 F}{\partial v^2} \\ \frac{\partial^2 f}{\partial t^2} &= -c^2 \left(-\frac{\partial^2 F}{\partial u^2} + \frac{\partial^2 F}{\partial u \partial v} \right) + c^2 \left(-\frac{\partial^2 F}{\partial u \partial v} + \frac{\partial^2 F}{\partial v^2} \right) = c^2 \left(\frac{\partial^2 F}{\partial u^2} - 2 \frac{\partial^2 F}{\partial u \partial v} + \frac{\partial^2 F}{\partial v^2} \right) \end{aligned}$$

d'où

$$\frac{\partial^2 f}{\partial t^2} - c^2 \frac{\partial^2 f}{\partial x^2} = -4c^2 \frac{\partial^2 F}{\partial u \partial v}$$

On en déduit que f est solution de (E) ssi F est solution de

$$(E') \quad \forall (u, v) \in \mathbb{R}^2 \quad \frac{\partial^2 F}{\partial u \partial v}(u, v) = 0$$

Mais $\frac{\partial}{\partial u} \left(\frac{\partial F}{\partial v} \right) (u, v) = 0 \Leftrightarrow \frac{\partial F}{\partial v}(u, v)$ ne dépend pas de u donc ssi il existe une fonction h_1 d'une variable telle que $\frac{\partial F}{\partial v}(u, v) = h_1(v)$. Comme F doit être C^2 , h_1 est nécessairement C^1 . Soit h une primitive de h_1 . h est C^2 . La fonction $(u, v) \rightarrow F(u, v) - h(v)$ doit avoir une dérivée par rapport à v nulle, donc doit n'être fonction que de u , soit $g(u)$ avec nécessairement g de classe C^2 . Réciproquement, si g et h sont deux fonctions de classe C^2 d'une seule variable, la fonction $(u, v) \rightarrow g(u) + h(v)$ vérifie (E').

En conclusion, les solutions de l'équation des ondes (E) sont les fonctions de la forme

$$(x, t) \rightarrow g(x - ct) + h(x + ct)$$

où g et h sont des fonctions arbitraires définies et C^2 sur \mathbb{R} .

Cas de la dimension 1

PROPOSITION 14.1.1

Soit U un intervalle ouvert de \mathbb{R} et $\varphi : U \rightarrow \mathbb{R}$ une application de classe C^k , $k \geq 1$ telle que $\forall x \in U \varphi'(x) \neq 0$. Alors $\varphi(U)$ est un intervalle ouvert et φ est un C^k difféomorphisme de U sur $\varphi(U)$.

En effet, φ' est continue et ne s'annule pas sur U connexe, donc elle est de signe constant. Il s'ensuit que φ est strictement monotone sur U . Par conséquent, c'est un homéomorphisme de U sur $\varphi(U)$ qui est nécessairement un intervalle ouvert. Ensuite, le théorème sur la dérivée des fonctions réciproques garantit que φ^{-1} est C^1 avec $(\varphi^{-1})' = 1/\varphi' \circ \varphi^{-1}$. De cette dernière relation on déduit facilement que φ^{-1} est C^k .

Remarque : l'équation (14.1) est une généralisation de la formule $(\varphi^{-1})' = 1/\varphi' \circ \varphi^{-1}$.

14.2 Le théorème d'inversion locale

En dimension 1, si une application φ d'un intervalle U dans \mathbb{R} vérifie $\varphi'(x) \neq 0$ pour tout x dans U , cette application est un difféomorphisme de U sur l'intervalle $\varphi(U)$. En dimension supérieure ou égale à 2, ce résultat ne subsiste pas : si U est un ouvert de \mathbb{R}^p et si $\varphi : U \rightarrow \mathbb{R}^p$ de classe C^1 vérifie $d\varphi(x)$ inversible pour tout x dans U , φ n'est pas nécessairement injective sur U , donc n'est pas un difféomorphisme de U sur $\varphi(U)$ en général.

DEFINITION 14.2.1

Soient E, F deux espaces vectoriels de même dimension p , Ω un ouvert de E et $\varphi : \Omega \rightarrow F$ de classe C^k , $k \geq 1$. Soit enfin $a \in \Omega$. On dit que φ est un C^k -difféomorphisme local en a si il existe

- un voisinage ouvert U de a dans Ω
- un voisinage ouvert V de $b = \varphi(a)$ dans \mathbb{R}^p

tels que la restriction de φ à U soit un C^k difféomorphisme de U sur V .

Il faut noter que cette définition est locale. En particulier, le point $b = \varphi(a)$ peut avoir par φ d'autres antécédents que a , mais ces antécédents seront bien entendu dans $\Omega \setminus U$.

En appliquant les résultats de la section précédente à $\varphi|_U$, on voit que si φ est un difféomorphisme local en a , $d\varphi(a)$ est un isomorphisme linéaire de E sur F .

Le théorème d'inversion locale affirme que la réciproque est vraie.

THEOREME 14.2.1 (Théorème d'inversion locale)

Soient E, F deux \mathbb{R} -espaces vectoriels de dimension finie, Ω un ouvert de E et $\varphi : \Omega \rightarrow F$ de classe C^k , $k \geq 1$. Soit $a \in \Omega$. Si $d\varphi(a)$ est un isomorphisme de E sur F , φ est un difféomorphisme local en a .

Ce théorème est admis.

Remarquons que l'hypothèse sur $d\varphi(a)$ implique $\dim E = \dim F$.

Si φ est donné en coordonnées par

$$\varphi(x_1, \dots, x_p) = (\varphi_1(x_1, \dots, x_p), \dots, \varphi_p(x_1, \dots, x_p))$$

on appelle jacobien de φ au point x le déterminant de la matrice jacobienne, i.e. de la matrice $d\varphi(x)$ c'est à dire

$$J_\varphi(x) = \det \begin{pmatrix} \frac{\partial \varphi_1}{\partial x_1}(x) & \cdots & \frac{\partial \varphi_1}{\partial x_p}(x) \\ \vdots & & \vdots \\ \frac{\partial \varphi_p}{\partial x_1}(x) & \cdots & \frac{\partial \varphi_p}{\partial x_p}(x) \end{pmatrix}$$

On note souvent

$$J_\varphi(x) = \frac{D(\varphi_1, \dots, \varphi_p)}{D(x_1, \dots, x_p)}(x)$$

Le théorème s'énonce donc en abrégé sous la forme $J_\varphi(a) \neq 0 \Rightarrow \varphi$ est un difféomorphisme local en a .

EXEMPLE 14.2.1 (Coordonnées polaires)

Soit $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ l'application définie par

$$\varphi(\rho, \theta) = (\rho \cos \theta, \rho \sin \theta)$$

Le jacobien vaut

$$J_\varphi(\rho, \theta) = \begin{vmatrix} \cos \theta & -\rho \sin \theta \\ \sin \theta & \rho \cos \theta \end{vmatrix} = \rho$$

φ est donc un difféomorphisme local en tout point (ρ, θ) avec $\rho \neq 0$.

Ceci se traduit concrètement de la manière suivante : soit $(x_0, y_0) \neq (0, 0)$. Choisissons un couple (ρ_0, θ_0) tel que $x_0 = \rho_0 \cos \theta_0, y_0 = \rho_0 \sin \theta_0$. Alors il existe $\alpha > 0, \eta > 0$ tel que l'application φ soit un C^∞ difféomorphisme de $] \rho_0 - \eta, \rho_0 + \eta[\times] \theta_0 - \alpha, \theta_0 + \alpha[$ sur un voisinage V de (x_0, y_0) .

Il est bien évident que φ n'est pas une bijection de $\mathbb{R}^2 \setminus \{(0, 0)\}$ sur $\mathbb{R}^2 \setminus \{(0, 0)\}$.

Exercice : démontrer directement (de manière élémentaire) que la restriction de φ est un C^∞ difféomorphisme de $\Omega_0 = \mathbb{R}_+^* \times]-\pi, \pi[$ sur $U_0 = \mathbb{R}^2 \setminus \{(x, 0) ; x \leq 0\}$.

THEOREME 14.2.2 (Théorème d'inversion globale)

Soient E et F deux espaces vectoriels réels de dimension finie, Ω un ouvert de E et $\varphi : \Omega \rightarrow F$ de classe C^k vérifiant les deux propriétés suivantes :

- (1) φ est injectif.
- (2) $\forall x \in \Omega$ $d\varphi(x)$ est un isomorphisme de E sur F .

Alors $\varphi(\Omega)$ est un ouvert de F et φ est un C^k difféomorphisme de E sur F .

preuve

Posons $W = \varphi(\Omega)$.

• W est ouvert

Soit $b \in W$. Il existe $a \in \Omega$ tel que $\varphi(a) = b$. $d\varphi(a)$ est un isomorphisme linéaire, donc φ est un difféomorphisme local en a . Il existe des ouverts U et V contenant respectivement a et b tels que φ induise un difféomorphisme φ_U de U sur V . On a donc $V = \varphi(U) \subset \varphi(\Omega) = W$ ce qui prouve que W est ouvert.

• φ est un C^1 difféomorphisme de Ω sur W .

Puisque φ est injectif, c'est une bijection de Ω sur W . Posons $\psi = \varphi^{-1}$. Soit $b \in W$. Montrons que ψ est C^k au voisinage de b . Avec les notations précédentes, on a, puisque φ est injective $\psi|_V = (\varphi_U)^{-1}$. Donc dans le voisinage ouvert V de b , ψ coïncide avec l'application de classe C^k $(\varphi_U)^{-1}$. Par conséquent ψ est C^k dans V . Comme b est arbitraire, ψ est C^k sur Ω .

EXEMPLE 14.2.2 (Coordonnées sphériques)

On se place dans un espace affine euclidien \mathcal{E} de dimension 3 orienté. $(O, \vec{i}, \vec{j}, \vec{k})$ est un repère orthonormé direct. On considère l'application $\Phi : \mathbb{R}^3 \rightarrow \mathcal{E}$ qui au triplet (r, θ, φ) associe le point M défini par

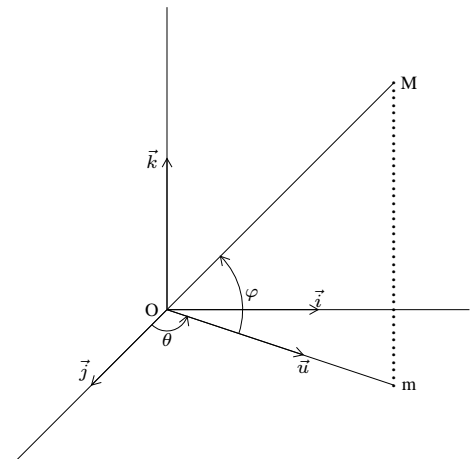
$$\overrightarrow{OM} = r \left(\cos \varphi \vec{u} + \sin \varphi \vec{k} \right) \quad \text{avec} \quad \vec{u} = \cos \theta \vec{i} + \sin \theta \vec{j}$$

On a donc, en notant (x, y, z) les coordonnées de M

$$\begin{cases} x = r \cos \varphi \cos \theta \\ y = r \cos \varphi \sin \theta \\ z = r \sin \varphi \end{cases}$$

(r, θ, φ) est appelé système de coordonnées sphériques de M .

La matrice de $d\Phi$ est $\begin{pmatrix} \cos \varphi \cos \theta & -r \cos \varphi \sin \theta & -r \sin \varphi \cos \theta \\ \cos \varphi \sin \theta & r \cos \varphi \cos \theta & -r \sin \varphi \sin \theta \\ \sin \varphi & 0 & r \cos \varphi \end{pmatrix}$ de déterminant $J_\Phi(r, \theta, \varphi) = r^2 \cos \varphi$.



Soit $\Omega = \mathbb{R}_+^* \times]-\pi, \pi[\times]-\pi/2, \pi/2[$. On vérifie facilement que $\Phi|_\Omega$ est injective. Par ailleurs son jacobien est partout non nul sur Ω . Donc Φ est un C^∞ difféomorphisme de Ω sur $\Phi(\Omega)$. Des calculs élémentaires montrent que $\Phi(\Omega) = \mathcal{E} \setminus \Pi$ ou Π est le demi-plan $\Pi = \{m(x, 0, z) ; x \leq 0\}$.

NB: θ est la longitude, φ la latitude. En physique, on prend souvent pour φ l'angle (\vec{k}, \widehat{OM}) , c'est à dire $\pi/2 - \varphi$; on obtient alors un jacobien égal à $r^2 \sin(\varphi)$ et le nouveau φ varie entre 0 et π .

14.3 Théorème des fonctions implicites

14.3.1 Cas $f : \mathbb{R}^2 \rightarrow \mathbb{R}$

THEOREME 14.3.1

Soient Ω un ouvert de \mathbb{R}^2 , $f : \Omega \rightarrow \mathbb{R}$ de classe C^1 et $(a, b) \in \Omega$ tel que $f(a, b) = 0$. On suppose que $\frac{\partial f}{\partial y}(a, b) \neq 0$. Alors, il existe un intervalle ouvert U de \mathbb{R} contenant a , un intervalle ouvert V de \mathbb{R} contenant b et une fonction $\varphi : U \rightarrow V$ de classe C^1 telle qu'on ait l'équivalence

$$(F_1) \quad \begin{cases} (x, y) \in U \times V \\ f(x, y) = 0 \end{cases} \Leftrightarrow \begin{cases} x \in U \\ y = \varphi(x) \end{cases}$$

et

$$\forall (x, y) \in U \times V \quad f'_y(x, y) = \frac{\partial f}{\partial y}(x, y) \neq 0$$

De plus, pour $x \in U$ on a

$$\varphi'(x) = -\frac{f'_x(x, \varphi(x))}{f'_y(x, \varphi(x))}$$

Enfin, si f est C^k , φ est C^k .

Le théorème des fonctions implicites est admis par le programme. Nous en donnerons cependant une démonstration dans le cas particulier ci dessus, démonstration qui est une application du théorème d'inversion locale.

Considérons l'application $\Phi : \Omega \rightarrow \mathbb{R}^2$ définie par $\Phi(x, y) = (x, f(x, y))$. La matrice jacobienne au point (a, b) est $J = \begin{pmatrix} 1 & 0 \\ f'_x(a, b) & f'_y(a, b) \end{pmatrix}$.

Son jacobien en (a, b) est donc non nul. Par conséquent, Φ est un difféomorphisme local d'un voisinage W de (a, b) sur un voisinage de $(a, 0)$. Quitte à restreindre W , on peut le supposer de la forme $W = U \times V$ où U, V sont des intervalles ouverts contenant respectivement a et b . Soit $W' = \Phi(U \times V)$.

1. Comme $d\Phi(x, y)$ est inversible en tout point de W , on a $f'_y(x, y) \neq 0$ pour $(x, y) \in W$.

2. Pour $(x, z) \in W'$, on a $\Phi^{-1}(x, z) = (x, \psi(x, z))$ avec $\psi : W' \rightarrow V$ de classe C^1 . Alors, pour $(x, y) \in U \times V$ on a

$$f(x, y) = 0 \Leftrightarrow \Phi(x, y) = (x, 0) \Leftrightarrow (x, y) = \Phi^{-1}(x, 0) \Leftrightarrow y = \psi(x, 0)$$

La fonction $\varphi(x) = \psi(x, 0)$ est définie sur un voisinage U' de 0. Quitte à remplacer U par $U \cap U'$, on a le résultat voulu.

Puisque φ est C^1 , en dérivant en x l'égalité $\forall x \in U \quad f(x, \varphi(x)) = 0$ on obtient la relation donnant $\varphi'(x)$. Enfin, une récurrence finie classique à partir de cette formule montre que si f est C^k alors φ est C^k . ■

Il est important de noter que ce théorème est local aussi bien en x (U voisinage de a) qu'en y (V voisinage de b). En particulier, il peut parfaitement exister d'autres fonctions $\varphi_1 : U \rightarrow \mathbb{R}$ telles que $f(x, \varphi_1(x)) = 0$ pour tout $x \in U$ mais ces fonctions seront à valeurs dans $\mathbb{R} \setminus V$.

Exercice

Montrer que la relation $\arctan(xy) = \exp(x+y) - 1$ définit au voisinage de $(0, 0)$ une fonction $y = f(x)$ et déterminer un développement limité à l'ordre 4 de f en 0.

Posons $F(x, y) = \arctan(xy) - \exp(x+y) + 1$. F est de classe C^∞ et $F(0, 0) = 0$. On a $F'_y(x, y) = \frac{x}{1+x^2y^2} - \exp(x+y)$ donc $F'_y(0, 0) = -1$. On peut donc appliquer le théorème des fonctions implicites. Il existe deux intervalles ouverts U et V contenant 0 et $f : U \rightarrow V$ de classe C^∞ tels que $\begin{cases} (x, y) \in U \times V \\ F(x, y) = 0 \end{cases} \Leftrightarrow \begin{cases} x \in U \\ y = f(x) \end{cases}$

On a facilement $F'_x(0, 0) = -1$ donc $f'(0) = -1$. Puisque f est C^∞ elle admet un développement limité d'ordre 4 en 0 soit $f(x) = -x + ax^2 + bx^3 + cx^4 + o(x^4)$. (Le coefficient de x est $f'(0) = -1$). La fonction $x \rightarrow g(x) = \arctan(xf(x)) - \exp(x+f(x)) + 1$ admet donc elle aussi un DL₄ en 0. Comme cette fonction est nulle, ce DL est nul et on obtiendra a, b, c par identification.

$\arctan(u) = u - u^3/3 + o(u^3)$ avec $u = -x^2 + ax^3 + bx^4 + o(x^4) \sim -x^2$ donc $u^3 \sim -x^6 = o(x^4)$ et

$$\arctan(xf(x)) = \arctan(-x^2 + ax^3 + bx^4 + o(x^4)) = -x^2 + ax^3 + bx^4 + o(x^4)$$

$v = x + f(x) = ax^2 + bx^3 + cx^4$ donc

$$\exp(x+f(x)) = 1 + v + v^2/2 + o(x^4) = 1 + ax^2 + bx^3 + cx^4 + a^2x^4/2 + o(x^4)$$

$$g(x) = -(a+1)x^2 + (a-b)x^3 + (b-c-a^2/2)x^4 + o(x^4)$$

D'où $a = -1$, $b = -1$, $c = -3/2$ et $f(x) = -x - x^2 - x^3 - 3x^4/2 + o(x^4)$

Application géométrique

Soient P un plan affine euclidien rapporté à un repère orthonormé (O, \vec{i}, \vec{j}) et Ω un ouvert de P . Soit $f : \Omega \rightarrow \mathbb{R}$ une application. Si $\vec{OM} = x\vec{i} + y\vec{j}$ on notera $f(M) = f(x, y)$. On suppose f de classe C^1 . On se propose d'étudier les lignes de niveau $C_k = \{M ; f(M) = k\}$ pour $k \in \mathbb{R}$.

On dira qu'un sous ensemble Γ de P est un morceau de courbe de classe C^1 si il existe un intervalle ouvert non vide I et une application $\varphi : I \rightarrow \mathbb{R}$ de classe C^1 telle que $\Gamma = \{M(x, y) \mid x \in I \text{ et } y = \varphi(x)\}$ ou $\Gamma = \{M(x, y) \mid y \in I \text{ et } x = \varphi(y)\}$.

PROPOSITION 14.3.1

Soit $M_0 \in C_k$. Supposons que $\nabla f(M_0) \neq 0$. Il existe un voisinage ouvert W de M_0 dans P tel que $\Gamma = C_k \cap W$ soit un morceau de courbe de classe C^1 . De plus, en tout point $M \in \Gamma$ on a $\nabla f(M) \neq 0$ et la tangente à Γ en M est la droite passant par M orthogonale à $\nabla f(M)$.

preuve

Soit (x_0, y_0) les coordonnées de M_0 . On a $\nabla f(M_0) = f'_x(x_0, y_0)\vec{i} + f'_y(x_0, y_0)\vec{j}$. Puisque ce vecteur est non nul, on a par exemple $f'_y(x_0, y_0) \neq 0$. On peut alors appliquer le théorème des fonctions implicites à f en ce point. Il existe des voisinages ouverts U et x_0 et V de y_0 et une fonction $\varphi : U \rightarrow V$ de classe C^1 tels qu'on ait l'équivalence (F_1) et $f'_y(x, y) \neq 0$ dans $U \times V$. Soit $W = \{M(x, y) ; (x, y) \in U \times V\}$. C'est un voisinage de M_0 et $C_k \cap W$ est le graphe de φ .

La tangente en un point M de Γ est dirigée par le vecteur $\vec{i} + \varphi'(x)\vec{j}$. Compte tenu de la formule donnant $\varphi'(x)$, on voit que ce vecteur est orthogonal à $\nabla f(M)$.

Si $f'_y(x_0, y_0) = 0$ on a $f'_x(x_0, y_0) \neq 0$ et on obtient les mêmes résultats en échangeant les rôles de x et de y .

Exemple : tangente à l'ellipse

Dans le plan affine euclidien P on se donne deux points distincts F et F' et on considère l'ellipse $E = \{M \mid MF + MF' = 2a\}$ où $2a > 2c = FF'$. Soit f la fonction $f(M) = FM + F'M = \left\| \overrightarrow{FM} \right\| + \left\| \overrightarrow{F'M} \right\|$. La fonction f est différentiable en tout point de $P \setminus \{F', F\}$ (c.f. l'exemple 13.4.3), donc en tout point de E et

$$\nabla f(M) = \frac{\overrightarrow{FM}}{FM} + \frac{\overrightarrow{F'M}}{F'M}$$

Soient $\vec{u} = \frac{\overrightarrow{FM}}{FM}$ et $\vec{u}' = \frac{\overrightarrow{F'M}}{F'M}$. On a $\nabla f(M) = 0 \Leftrightarrow \vec{u} + \vec{u}' = 0$, c'est à dire ssi M appartient au segment FF' privé de ses extrémités, mais dans ce cas $FM + F'M' = 2c < 2a$ et $M \notin E$. On peut donc appliquer le résultat précédent en tout point de E . Au voisinage de chaque point l'ellipse est un morceau de graphe et la tangente en M est la perpendiculaire passant par M au vecteur $\vec{u} + \vec{u}'$. Mais ce vecteur est un vecteur directeur de la bissectrice intérieure de l'angle $\widehat{FMF'}$. Donc la tangente en un point M d'une ellipse E est la bissectrice extérieure de l'angle $\widehat{FMF'}$.

14.3.2 Cas des fonctions $f : \mathbb{R}^3 \rightarrow \mathbb{R}$

THEOREME 14.3.2

Soient Ω un ouvert de \mathbb{R}^3 , $f : \Omega \rightarrow \mathbb{R}$ de classe C^1 et $(a, b, c) \in \Omega$ tel que $f(a, b, c) = 0$. On suppose que $\frac{\partial f}{\partial z}(a, b, c) \neq 0$. Alors, il existe un ouvert U de \mathbb{R}^2 contenant (a, b) , un intervalle ouvert V de \mathbb{R} contenant c et une fonction $\varphi : U \rightarrow V$ de classe C^1 telle qu'on ait l'équivalence

$$(F_2) \quad \left\{ \begin{array}{l} (x, y, z) \in U \times V \\ f(x, y, z) = 0 \end{array} \right. \Leftrightarrow \left\{ \begin{array}{l} (x, y) \in U \\ z = \varphi(x, y) \end{array} \right.$$

et

$$\forall (x, y, z) \in U \times V \quad f'_z(x, y, z) = \frac{\partial f}{\partial z}(x, y, z) \neq 0$$

De plus, pour $(x, y) \in U$ on a

$$\frac{\partial \varphi}{\partial x}(x, y) = -\frac{f'_x(x, y, \varphi(x, y))}{f'_z(x, y, \varphi(x, y))} \quad \text{et} \quad \frac{\partial \varphi}{\partial y}(x, y) = -\frac{f'_y(x, y, \varphi(x, y))}{f'_z(x, y, \varphi(x, y))}$$

Enfin, si f est C^k , φ est C^k .

Application géométrique

Soit E un espace affine euclidien de dimension 3, $(O, \vec{i}, \vec{j}, \vec{k})$ un repère orthonormé direct. On dira qu'un sous ensemble Σ de E est un morceau de surface de classe C^1 si il existe un ouvert connexe non vide $U \subset \mathbb{R}^2$ et une application $\varphi : U \rightarrow \mathbb{R}$ de classe C^1 telle que $\Sigma = \{M(x, y, z) \mid (x, y) \in U \text{ et } z = \varphi(x, y)\}$ ou $\Sigma = \{M(x, y, z) \mid (y, z) \in U \text{ et } x = \varphi(y, z)\}$ ou $\Sigma = \{M(x, y, z) \mid (x, z) \in U \text{ et } y = \varphi(x, z)\}$.

Soit alors Ω un ouvert de E et $f : \Omega \rightarrow \mathbb{R}$ de classe C^1 et $S_k = \{M \in E \mid f(M) = k\}$.

PROPOSITION 14.3.2

Soit $M_0(x_0, y_0, z_0) \in S_k$. Supposons $\nabla f(M_0) \neq 0$. Alors il existe un voisinage ouvert W de M_0 dans E tel que $\Sigma = S_k \cap W$ soit un morceau de surface de classe C^1 . De plus en tout point $M \in \Sigma$ on a $\nabla f(M) \neq 0$ et le plan tangent à Σ en M est le plan passant par M et orthogonal à $\nabla f(M)$.

La démonstration est la même que dans le cas des courbes.

14.3.3 Cas des fonctions $f : \mathbb{R}^3 \rightarrow \mathbb{R}^2$

Si $g, h : \mathbb{R}^3 \rightarrow \mathbb{R}$ sont deux fonctions de classe C^1 , on notera, u et v désignant deux variables distinctes parmi x, y, z

$$\frac{D(g, h)}{D(u, v)} = \det \begin{vmatrix} g'_u & g'_v \\ h'_u & h'_v \end{vmatrix}$$

THEOREME 14.3.3

Soient Ω un ouvert de \mathbb{R}^3 , $f = (X, Y) : \Omega \rightarrow \mathbb{R}^2$ de classe C^1 et $(a, b, c) \in \Omega$ tel que $f(a, b, c) = 0$.

on suppose que le déterminant

$$\frac{D(X, Y)}{D(y, z)}(a, b, c) := \begin{vmatrix} \frac{\partial X}{\partial y}(a, b, c) & \frac{\partial X}{\partial z}(a, b, c) \\ \frac{\partial Y}{\partial y}(a, b, c) & \frac{\partial Y}{\partial z}(a, b, c) \end{vmatrix}$$

est non nul.

Alors, il existe un ouvert U de \mathbb{R} contenant a , un ouvert V de \mathbb{R}^2 contenant (b, c) et une fonction $(\varphi, \psi) : U \rightarrow V$ de classe C^1 telle qu'on ait l'équivalence

$$(F_3) \quad \begin{cases} (x, (y, z)) \in U \times V \\ X(x, y, z) = 0 \\ Y(x, y, z) = 0 \end{cases} \Leftrightarrow \begin{cases} x \in U \\ y = \varphi(x) \\ z = \psi(x) \end{cases}$$

et

$$\forall (x, (y, z)) \in U \times V \quad \frac{D(X, Y)}{D(y, z)}(x, y, z) \neq 0$$

De plus, pour $x \in U$, $\varphi'(x)$ et $\psi'(x)$ sont solutions du système de Cramer

$$\begin{cases} X'_x(x, \varphi(x), \psi(x)) + X'_y(x, \varphi(x), \psi(x))\varphi'(x) + X'_z(x, \varphi(x), \psi(x))\psi'(x) = 0 \\ Y'_x(x, \varphi(x), \psi(x)) + Y'_y(x, \varphi(x), \psi(x))\varphi'(x) + Y'_z(x, \varphi(x), \psi(x))\psi'(x) = 0 \end{cases}$$

Enfin, si f est C^k , φ et ψ sont C^k .

Application géométrique

Soit E un espace affine euclidien de dimension 3, $(O, \vec{i}, \vec{j}, \vec{k})$ un repère orthonormé direct.

On dira qu'un sous ensemble Γ de E est un morceau de courbe de classe C^1 si il existe un intervalle ouvert non vide $I \subset \mathbb{R}$ et deux applications $\alpha : I \rightarrow \mathbb{R}$, $\beta : I \rightarrow \mathbb{R}^2$ de classe C^1 telle que $\Gamma = \{M(x, y, z) \mid z \in I \text{ et } x = \alpha(z), y = \beta(z)\}$ ou $\Gamma = \{M(x, y, z) \mid y \in I \text{ et } x = \alpha(y), z = \beta(y)\}$ ou $\Gamma = \{M(x, y, z) \mid x \in I \text{ et } y = \alpha(x), z = \beta(x)\}$.

Soient Ω un ouvert de E et $f = (X, Y) : \Omega \rightarrow \mathbb{R}^2$ de classe C^1 . Soit $C = \{M \mid f(M) = 0\}$

PROPOSITION 14.3.3

Soit $M_0 \in C$. On suppose que les vecteurs $\nabla X(M_0)$ et $\nabla Y(M_0)$ sont indépendants. Alors, il existe un voisinage ouvert W de M_0 dans E tel que $\Gamma = C \cap W$ soit un morceau de courbe de classe C^1 . De plus, pour tout $M \in \Gamma$, les vecteurs $\nabla X(M)$ et $\nabla Y(M)$ sont indépendants et la tangente à Γ en M est la droite passant par M et orthogonale à $\nabla X(M)$ et à $\nabla Y(M)$, c'est à dire colinéaire à $\nabla X(M) \wedge \nabla Y(M)$ si on oriente l'espace.

Si $\nabla X(M)$ et $\nabla Y(M)$ sont indépendants, ils ne sont pas nuls. On peut donc appliquer la proposition précédente. Il existe un voisinage W de M_0 , tel que $\Sigma = W \cap \{M \mid X(M) = 0\}$ et $\Sigma' = W \cap \{M \mid Y(M) = 0\}$ soient des morceaux de surface. On a alors $C \cap W = \Sigma \cap \Sigma'$.

La condition $\nabla X(M_0)$ et $\nabla Y(M_0)$ indépendants exprime le fait que les plans tangents à Σ et à Σ' en M_0 sont non parallèles, donc en fait distincts. La proposition affirme que cette condition implique que C est un morceau de courbe et que la tangente à C en M_0 est l'intersection des plans tangents en M_0 à Σ et Σ' .

preuve

Orientons l'espace directeur de E de sorte que le repère donné soit orthonormé dircet. Dire que $\nabla X(M)$ et $\nabla Y(M)$ sont

indépendants revient à dire que leur produit vectoriel est non nul. Or ce produit vectoriel a pour composantes

$$\left(\frac{D(X, Y)}{D(y, z)}, \frac{D(X, Y)}{D(z, x)}, \frac{D(X, Y)}{D(x, y)} \right)$$

Soit $M_0 \in C$. L'une des composantes de ce produit vectoriel est donc non nulle. Supposons par exemple que ce soit la première. On peut alors appliquer le théorème des fonctions implicites. On obtient un intervalle ouvert U de \mathbb{R} , contenant x_0 , un ouvert V de \mathbb{R}^2 contenant (y_0, z_0) et deux fonctions $\varphi : U \rightarrow \mathbb{R}$ et $\psi : U \rightarrow \mathbb{R}$ telles que (φ, ψ) soit à valeurs dans V et qu'on ait l'équivalence (F_3) . Si $W = \{M(x, y, z) \mid x \in U \text{ et } (y, z) \in V\}$ on voit que $C \cap W$ est le morceau de courbe défini par $x \rightarrow (x, \varphi(x), \psi(x))$. Un vecteur tangent \vec{T} en M_0 à (C) a pour composantes $(1, \varphi'(x_0), \psi'(x_0))$ et le système de Cramer donnant $(\varphi'(x_0), \psi'(x_0))$ exprime exactement le fait que \vec{T} est orthogonal à $\nabla X(M_0)$ et à $\nabla Y(M_0)$.

14.3.4 Énoncé du théorème des fonctions implicites dans le cas général

Soient Ω un ouvert de $\mathbb{R}^p \times \mathbb{R}^q$ et $f : \Omega \rightarrow \mathbb{R}^q$ une application de classe C^1 . Soit $a \in \mathbb{R}^p$, $b \in \mathbb{R}^q$ tels que $f(a, b) = 0$. Notons $(x, y) \in \mathbb{R}^p \times \mathbb{R}^q$ et $x = (x_1, \dots, x_p)$, $y = (y_1, \dots, y_q)$. Enfin, notons $f = (f_1, \dots, f_q)$. On suppose que

$$\frac{D(f_1, \dots, f_q)}{D(y_1, \dots, y_q)}(a, b) := \det \begin{vmatrix} \frac{\partial f_1}{\partial y_1}(a, b) & \cdots & \frac{\partial f_1}{\partial y_q}(a, b) \\ \vdots & & \vdots \\ \frac{\partial f_q}{\partial y_1}(a, b) & \cdots & \frac{\partial f_q}{\partial y_q}(a, b) \end{vmatrix}$$

est non nul.

Alors il existe un ouvert U contenant a , un ouvert V contenant b et une application $\varphi : U \rightarrow V$ de classe C^1 tels que

$$\forall (x, y) \in U \times V \quad \frac{D(f_1, \dots, f_q)}{D(y_1, \dots, y_q)}(x, y) \neq 0$$

et

$$\left\{ \begin{array}{l} (x, y) \in U \times V \\ f(x, y) = 0 \end{array} \right\} \Leftrightarrow \left\{ \begin{array}{l} x \in U \\ y = \varphi(x) \end{array} \right.$$

Autrement dit, ce théorème donne une condition suffisante pour qu'on puisse localement résoudre le système

$$(S) \quad \begin{cases} f_1(x_1, \dots, x_p, y_1, \dots, y_q) = 0 \\ \vdots \\ f_q(x_1, \dots, x_p, y_1, \dots, y_q) = 0 \end{cases}$$

sous la forme

$$\begin{cases} y_1 = \varphi_1(x_1, \dots, x_p) \\ \vdots \\ y_q = \varphi_q(x_1, \dots, x_p) \end{cases}$$

Enfin, les dérivées partielles des φ_j par rapport à la variable x_k fixée sont obtenues en dérivant les q identités

$$f_m(x_1, \dots, x_p, \varphi_1(x_1, \dots, x_p), \dots, \varphi_q(x_1, \dots, x_p)) = 0 \quad 1 \leq m \leq q$$

par rapport à x_k . Le système obtenu est linéaire et a pour déterminant $\frac{D(f_1, \dots, f_q)}{D(y_1, \dots, y_q)}(x, \varphi(x)) \neq 0$ donc est de Cramer.

15

Equations et systèmes différentiels linéaires

Dans tout ce chapitre $\mathbb{K} = \mathbb{R}$ ou \mathbb{C} .

15.1 Equations différentielles linéaires scalaires du premier ordre

15.1.1 Théorème de Cauchy

Soit I un intervalle non trivial (i.e. non vide et non réduit à un point) de \mathbb{R} et $p, f : I \rightarrow \mathbb{K}$ deux fonctions. On considère les équations différentielles

$$(E) \quad y' + p(t)y = f(t)$$

et

$$(E_0) \quad y' + p(t)y = 0$$

Une solution de (E) sur un sous intervalle $J \subset I$ est une application dérivable $\varphi : J \rightarrow \mathbb{K}$ telle que

$$\forall t \in J \quad \varphi'(t) + p(t)\varphi(t) = f(t)$$

Soit $\varphi : J \rightarrow \mathbb{K}$ une solution de (E). Si $\psi : J \rightarrow \mathbb{K}$ est dérivable, ψ est solution de (E) ssi $\psi - \varphi$ est solution de (E₀). Si p et f sont continues, toute solution de (E) est C^1 .

THEOREME 15.1.1

Soient $p : I \rightarrow \mathbb{K}$ continue, $t_0 \in I$ et $y_0 \in \mathbb{K}$. Il existe une solution de (E₀) définie sur I tout entier, et une seule φ telle que $\varphi(t_0) = y_0$. Elle est donnée par

$$\forall t \in I \quad \varphi(t) = y_0 \exp\left(-\int_{t_0}^t p(s)ds\right)$$

preuve

Posons $P(t) = \int_{t_0}^t p(s)ds$. Puisque p est continue, P est C^1 . Soit $\varphi : I \rightarrow \mathbb{K}$ de classe C^1 et $\psi(t) = \varphi(t)e^{P(t)}$. On a $\psi'(t) = e^{P(t)}(\varphi'(t) + \varphi(t)p(t))$ donc φ solution de (E₀) ssi $\forall t \in I \quad \psi'(t) = 0 \Leftrightarrow \forall t \in I \quad \psi(t) = C$ où C est une constante. Donc $\varphi(t) = C \exp(-P(t))$ et la condition $\varphi(t_0) = y_0$ donne $C = y_0$. ■

Les solutions de (E₀) forment donc un \mathbb{K} -ev de dimension 1. Mise à part la fonction nulle, une solution de (E₀) ne s'annule jamais sur I .

Méthode de variation de la constante

Soit $p, f : I \rightarrow \mathbb{K}$ continue et φ_0 une solution sur I de l'équation homogène (E₀). Soit $t_0 \in I$. Soit $g : I \rightarrow \mathbb{K}$ une fonction de classe C^1 quelconque. Soit $z : I \rightarrow \mathbb{K}$ la fonction définie par $z(t) = \frac{g(t)}{\varphi_0(t)}$. Puisque φ_0 est C^1 et ne s'annule pas, z est C^1 et $\forall t \in I \quad g(t) = z(t)\varphi_0(t)$ de sorte que

$$g'(t) + p(t)g(t) = z'(t)\varphi_0(t) + z(t)(\varphi_0'(t) + p(t)\varphi_0(t)) = z'(t)\varphi_0(t)$$

On en déduit que g est solution de (E) sur I ssi $z'(t) = \frac{f(t)}{\varphi_0(t)}$. On en déduit z par intégration. Finalement les solutions de (E) sur I sont données par

$$g(t) = \left(\int_{t_0}^t \frac{f(s)}{\varphi_0(s)} ds + C \right) \varphi_0(t)$$

où C est une constante arbitraire et $t_0 \in I$.

En prenant $P(t) = \int_{t_0}^t p(s) ds$ et $\varphi_0(t) = e^{-P(t)}$ on obtient

$$g(t) = C e^{-P(t)} + e^{-P(t)} \int_{t_0}^t f(s) e^{P(s)} ds$$

THEOREME 15.1.2

Soient $p, f : I \rightarrow \mathbb{K}$ continue, $t_0 \in I$ et $y_0 \in \mathbb{K}$. L'équation (E) admet une solution sur I φ et une seule vérifiant $\varphi(t_0) = y_0$. Elle est donnée par

$$\varphi(t) = e^{-P(t)} \left(y_0 + \int_{t_0}^t f(s) e^{P(s)} ds \right) = y_0 e^{-P(t)} + \int_{t_0}^t f(s) e^{\int_t^s p(u) du} \quad \text{avec} \quad P(t) = \int_{t_0}^t p(s) ds$$

D'un point de vue pratique, pour résoudre (E), on commence par résoudre (E₀). On trouve un espace vectoriel de solutions $\{\lambda \varphi_0 ; \lambda \in \mathbb{K}\}$ où φ_0 est une solution non nulle de (E₀). Ensuite, on recherche une solution particulière ψ de (E). Si il n'y en a pas d'évidente, on cherche ψ sous la forme $\psi(t) = z(t)\varphi_0(t)$ donc en remplaçant la constante λ par une fonction inconnue z , d'où le nom de méthode de variation de la constante.

15.1.2 Problèmes de raccordement

Soient maintenant I un intervalle non trivial de \mathbb{R} et $a, b, c : I \rightarrow \mathbb{K}$ trois fonctions continues sur I . On considère les équations différentielles

$$(E'_0) \quad a(t)y' + b(t)y = 0 \quad \text{et} \quad (E') \quad a(t)y + b(t)y' = c(t)$$

L'ensemble des solutions de (E₀) sur I forme un \mathbb{K} -ev, soit $\text{Sol}_I(E'_0)$. L'ensemble des solutions de (E') sur I est soit vide, soit un espace affine d'espace directeur $\text{Sol}_I(E'_0)$.

Si J est un sous intervalle de I sur lequel a ne s'annule pas, l'équation (E') s'écrit, sur J $y' + p(t)y = f(t)$ avec $p = b/a$ et $f = c/a$. Les résultats précédents s'appliquent donc sur tout intervalle J sur lequel a ne s'annule pas. Mais l'ensemble des solutions sur I tout entier peut être vide, ou de dimension strictement supérieure à 1.

Exemple 1

Considérons l'équation $xy' + y = 0$. Les solutions sur \mathbb{R}_+^* (resp. sur \mathbb{R}_-^*) sont les fonctions $x \rightarrow C_+/x$ (resp. $x \rightarrow C_-/x$). Soit φ une solution sur \mathbb{R} . $\varphi|_{\mathbb{R}_+^*}$ est solution sur \mathbb{R}_+^* , donc il existe une constante C_+ telle que $\forall x > 0 \varphi(x) = C_+/x$. De même, il existe une constante C_- telle que $\forall x < 0 \varphi(x) = C_-/x$. Puisque φ est solution sur \mathbb{R} , elle est continue en 0. On a donc $C_+ = C_- = 0$ et $\varphi = 0$ qui est bien évidemment solution sur \mathbb{R} . L'espace des solutions sur \mathbb{R} est ici réduit à $\{0\}$. En employant la même méthode, le lecteur vérifiera sans peine que l'équation $xy' + y = x$ admet une unique solution sur \mathbb{R} qui est la fonction $x \rightarrow \frac{1}{2}x$.

Exemple 2

Considérons maintenant l'équation $xy' - y = 0$. Les solutions sur \mathbb{R}_+^* (resp. sur \mathbb{R}_-^*) sont les fonctions $x \rightarrow C_+x$ (resp. $x \rightarrow C_-x$).

Soit φ une solution sur \mathbb{R} . $\varphi|_{\mathbb{R}_+^*}$ est solution sur \mathbb{R}_+^* , donc il existe une constante C_+ telle que $\forall x > 0 \varphi(x) = C_+x$. De même, il existe une constante C_- telle que $\forall x < 0 \varphi(x) = C_-x$. φ doit être dérivable en 0, donc $\varphi'_d(0) = \varphi'_g(0)$ d'où $C_+ = C_-$. Réciproquement, pour tout $C \in \mathbb{R}$, la fonction $x \rightarrow Cx$ est solution sur \mathbb{R} .

Exemple 3

Considérons enfin l'équation $xy' - 2y = 0$. Les solutions sur \mathbb{R}_+^* (resp. sur \mathbb{R}_-^*) sont les fonctions $x \rightarrow C_+x^2$ (resp. $x \rightarrow C_-x^2$).

Soit φ une solution sur \mathbb{R} . $\varphi|_{\mathbb{R}_+^*}$ est solution sur \mathbb{R}_+^* , donc il existe une constante C_+ telle que $\forall x > 0 \varphi(x) = C_+x^2$. De même, il existe une constante C_- telle que $\forall x < 0 \varphi(x) = C_-x^2$.

Réciproquement, pour tout couple C_+, C_- de réels, la fonction $\varphi(x) = \begin{cases} C_+x^2 & \text{si } x > 0 \\ 0 & \text{si } x = 0 \\ C_-x^2 & \text{si } x < 0 \end{cases}$ est solution sur \mathbb{R} . En effet, elle

est solution sur \mathbb{R}^* et en 0, $\varphi(0) = 0$ et φ est dérivable avec $\varphi'(0) = 0$, donc l'équation différentielle est vérifiée en $x = 0$. Dans cet exemple, l'espace des solutions sur \mathbb{R} est de dimension 2, une base étant formée des fonctions φ_0, φ_1 où $\varphi_0(x) = \begin{cases} x^2 & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases}$ et $\varphi_1(x) = \begin{cases} 0 & \text{si } x \geq 0 \\ x^2 & \text{si } x \leq 0 \end{cases}$.

15.2 Systèmes linéaires d'ordre 1

15.2.1 Énoncé du problème

On se donne un intervalle non trivial I de \mathbb{R} , E un \mathbb{K} -ev de dimension finie, $A : I \rightarrow L(E)$ et $B : I \rightarrow E$ deux applications continues. On considère le système différentiel

$$(S) \quad Y' = A(t)Y + B(t)$$

et le système homogène associé

$$(S_0) \quad Y' = A(t)Y$$

Si J est un sous intervalle de I , une J -solution de (S) est une application $Y : J \rightarrow E$ dérivable telle que

$$\forall t \in J \quad Y'(t) = A(t)Y(t) + B(t)$$

Si Y est une solution, $t \rightarrow A(t)Y(t) + B(t)$ est une fonction continue, donc Y est nécessairement de classe C^1 . Si A et B sont de classe C^m , toute solution est de classe C^{m+1} .

Nous noterons $\text{Sol}_J(S)$ (resp. $\text{Sol}_J(S_0)$) l'ensemble des solutions sur J de (S) (resp. de (S_0)). Soient f et g deux fonctions continues sur J avec f solution de (S). Alors g est solution de (S) ssi $g - f$ est solution de (S_0) .

Il est immédiat que $\text{Sol}_J(S_0)$ est un \mathbb{K} -espace vectoriel et que $\text{Sol}_J(S)$ est soit vide, soit un espace affine d'espace directeur $\text{Sol}_J(S_0)$.

Problème de Cauchy

Soient $t_0 \in I$ et $Y_0 \in E$. Le problème de Cauchy relatif à la donnée initiale (t_0, Y_0) consiste à rechercher toutes les J -solutions de (S) telles que $t_0 \in J$ et $Y(t_0) = Y_0$.

15.2.2 Théorème de Cauchy-Lipschitz linéaire

THEOREME 15.2.1

Soient E un \mathbb{K} -ev de dimension finie, I un intervalle non trivial de \mathbb{R} , $A : I \rightarrow L(E)$ et $B : I \rightarrow E$ continues.

1. Pour toute donnée initiale $(t_0, Y_0) \in I \times E$ il existe une I -solution Y et une seule du système $Y' = A(t)Y + B(t)$ vérifiant $Y(t_0) = Y_0$.

2. Pour tout sous intervalle J de I contenant t_0 , il existe une unique J -solution du problème de Cauchy avec la condition initiale (t_0, Y_0) qui est la restriction à J de Y .

Ce théorème est admis par le programme. Une preuve est donnée en fin de chapitre. Dans la suite, on notera $\text{Sol}(S_0) = \text{Sol}_I(S_0)$ et de même pour (S). Le terme solution de (S) désignera une I -solution de (S).

THEOREME 15.2.2

Soient E un \mathbb{K} -ev de dimension finie, I un intervalle non trivial de \mathbb{R} et $A : I \rightarrow L(E)$ continue. Soit $t_0 \in I$ fixé. L'application

$$e_{t_0} : \text{Sol}_I(S_0) \rightarrow E \quad \varphi \rightarrow \varphi(t_0)$$

est un isomorphisme d'espaces vectoriels.

preuve

L'application e_{t_0} est clairement linéaire. Le théorème de Cauchy-Lipschitz affirme que tout $Y_0 \in E$ a un unique antécédent par e_{t_0} . C'est donc une bijection.

COROLLAIRE 15.2.1

L'espace des solutions de (S_0) est de dimension $\dim E$.

Si $B : I \rightarrow E$ est continue, $\text{Sol}(S)$ est un espace affine de dimension $\dim E$.

COROLLAIRE 15.2.2

Soient toujours $A : I \rightarrow L(E)$ continue. Soient $\varphi_1, \dots, \varphi_p$ p solutions de $Y' = A(t)Y$. Les propriétés suivantes sont équivalentes :

- (1) $(\varphi_1, \dots, \varphi_p)$ est un système libre dans $\text{Sol}(S_0)$.
- (2) $\exists t_0 \in I$ $(\varphi_1(t_0), \dots, \varphi_p(t_0))$ est un système libre de E .
- (3) $\forall t \in I$ $(\varphi_1(t), \dots, \varphi_p(t))$ est un système libre de E .

preuve

D'après le théorème de Cauchy-Lipschitz, pour tout $t \in I$, ε_t est un isomorphisme de $\text{Sol}(S_0)$ sur E , d'où I étant non vide, les équivalences entre (1) et (3) et entre (1) et (2).

COROLLAIRE 15.2.3

Soient $A : I \rightarrow L(E)$ continue et $\varphi_1, \dots, \varphi_n$ $n = \dim E$ solutions de $Y' = A(t)Y$. Les propriétés suivantes sont équivalentes :

- (1) $(\varphi_1, \dots, \varphi_n)$ est une base de $\text{Sol}(S_0)$.
- (2) $\exists t_0 \in I$ $(\varphi_1(t_0), \dots, \varphi_n(t_0))$ est une base de E .
- (3) $\forall t \in I$ $(\varphi_1(t), \dots, \varphi_n(t))$ est une base de E .

Wronskien**DEFINITION 15.2.1**

Soient $A : I \rightarrow L(E)$ continue et $\mathcal{B} = (e_1, \dots, e_n)$ une base fixée de E . Soient $\varphi_1, \dots, \varphi_n$ n solutions de $Y' = A(t)Y$. On appelle wronskien dans la base \mathcal{B} du système $(\varphi_1, \dots, \varphi_n)$ la fonction

$$t \rightarrow w_{\mathcal{B}}(t) = \det_{\mathcal{B}}(\varphi_1(t), \dots, \varphi_n(t))$$

Ecrivons $\varphi_j(t) = \sum_{i=1}^{i=n} \varphi_{i,j}(t)e_i$ on a donc

$$w_{\mathcal{B}}(t) = \det(W_{\mathcal{B}}(t)) \quad \text{avec} \quad W_{\mathcal{B}}(t) = \begin{pmatrix} \varphi_{1,1}(t) & \cdots & \varphi_{1,n}(t) \\ \vdots & & \vdots \\ \varphi_{n,1}(t) & \cdots & \varphi_{n,n}(t) \end{pmatrix}$$

Si $E = \mathbb{K}^n$ on appelle wronskien du système le déterminant dans la base canonique de \mathbb{K}^n .

COROLLAIRE 15.2.4

Soient $A : I \rightarrow L(E)$ continue, $\mathcal{B} = (e_1, \dots, e_n)$ une base fixée de E et $\varphi_1, \dots, \varphi_n$ $n = \dim E$ solutions de $Y' = A(t)Y$. Les propriétés suivantes sont équivalentes :

- (1) $(\varphi_1, \dots, \varphi_n)$ est une base de $\text{Sol}(S_0)$
- (2) $\exists t_0 \in I$ $w_{\mathcal{B}}(t_0) \neq 0$
- (3) $\forall t \in I$ $w_{\mathcal{B}}(t) \neq 0$.

THEOREME 15.2.3

Soient $A : I \rightarrow L(E)$ continue, $\mathcal{B} = (e_1, \dots, e_n)$ une base fixée de E et $\varphi_1, \dots, \varphi_n$ $n = \dim E$ solutions de $Y' = A(t)Y$ et w leur wronskien relatif à la base \mathcal{B} . w vérifie l'équation différentielle scalaire

$$\forall t \in I \quad w'(t) = \text{tr}(A(t))w(t)$$

et donc, si $t_0 \in I$

$$\forall t \in I \quad w(t) = w(t_0) \exp\left(\int_{t_0}^t \text{tr}[A(s)]ds\right)$$

preuve

On a en omettant d'indiquer la base \mathcal{B}

$$\begin{aligned} w'(t) &= \det(\varphi_1'(t), \varphi_2(t), \dots, \varphi_n(t)) + \det(\varphi_1(t), \varphi_2'(t), \varphi_3(t), \dots, \varphi_n(t)) + \dots + \det(\varphi_1(t), \varphi_2(t), \dots, \varphi_n'(t)) \\ &= \det(A(t)\varphi_1(t), \varphi_2(t), \dots, \varphi_n(t)) + \det(\varphi_1(t), A(t)\varphi_2(t), \dots, \varphi_n(t)) + \dots + \det(\varphi_1(t), \dots, A(t)\varphi_n(t)) \end{aligned}$$

La fin de la preuve résulte du lemme d'algèbre linéaire suivant :

Lemme 15.2.1

Soient E un \mathbb{K} -ev de dimension n et $f \in L(F)$. Soient v_1, \dots, v_n n vecteurs de E et \mathcal{B} une base de E . On a

$$\sum_{j=1}^{j=n} \det_{\mathcal{B}}(v_1, \dots, v_{j-1}, f(v_j), v_{j+1}, \dots, v_n) = \text{tr}(f) \det_{\mathcal{B}}(v_1, \dots, v_n)$$

Preuve du lemme

L'application Φ qui à (v_1, \dots, v_n) associe le premier membre de l'égalité à prouver est manifestement n linéaire. Il est facile de voir qu'elle est alternée. Or l'espace vectoriel des formes n -linéaires alternées sur un espace vectoriel de dimension n est de dimension 1. Il existe donc un scalaire λ tel que $\Phi = \lambda \det_{\mathcal{B}}$. Pour calculer λ , on applique cette égalité au système $\mathcal{B} = (e_1, \dots, e_n)$. Le second membre vaut alors λ . Si $A = (a_{i,j}) = \text{Mat}_{\mathcal{B}}(f)$, on a

$$\det_{\mathcal{B}}(e_1, \dots, e_{j-1}, f(e_j), e_{j+1}, \dots, e_n) = \det_{\mathcal{B}}\left(e_1, \dots, e_{j-1}, \sum_{i=1}^{i=n} a_{i,j} e_i, e_{j+1}, \dots, e_n\right) = a_{j,j}$$

donc $\lambda = \sum_{i=1}^{i=n} a_{i,i} = \text{tr}(f)$.

15.2.3 Méthode de variation des constantes

Soient $(\varphi_1, \dots, \varphi_n)$ une base de solutions du système (S_0) .

Lemme 15.2.2

Soient $\psi : I \rightarrow E$ continue (resp. de classe C^1). Il existe des fonctions $u_k : I \rightarrow \mathbb{K}$, $1 \leq k \leq n$ continues (resp. de classes C^1) telles que

$$\forall t \in I \quad \psi(t) = \sum_{k=1}^{k=n} u_k(t) \varphi_k(t)$$

preuve

L'existence des fonctions u_k est évidente, puisque pour tout t , $(\varphi_1(t), \dots, \varphi_n(t))$ est une base de E . Il faut vérifier qu'elles sont continues (resp. de classe C^1). Soit $\mathcal{B} = (e_1, \dots, e_n)$ une base fixe de E . Soit $W(t) = \text{Mat}_{\mathcal{B}}(\varphi_1(t), \dots, \varphi_n(t))$.

Dans la base \mathcal{B} on a $\psi(t) = \sum_{k=1}^{k=n} \psi_k(t) e_k$ avec ψ_k continue (resp. de classe C^1) pour tout k . La formule de changement de base donne

$$\begin{pmatrix} u_1(t) \\ \vdots \\ u_n(t) \end{pmatrix} = (W_{\mathcal{B}}(t))^{-1} \begin{pmatrix} \psi_1(t) \\ \vdots \\ \psi_n(t) \end{pmatrix}$$

Comme l'application $GL(n, \mathbb{K}) \rightarrow M_n(\mathbb{K})$ qui à P associe P^{-1} est de classe C^1 , il en est de même de $t \rightarrow (W_{\mathcal{B}}(t))^{-1}$ par composition. Donc les fonctions u_k sont continues (resp. C^1).

Soit alors $\psi = \sum_{k=1}^{k=n} u_k \varphi_k$ de classe C^1 . On a

$$\psi' - A\psi = \sum_{k=1}^{k=n} (u_k' \varphi_k + u_k \varphi_k' - u_k A \varphi_k) = \sum_{k=1}^{k=n} u_k' \varphi_k$$

Le lemme garantit l'existence de fonctions continues $\beta_k : I \rightarrow \mathbb{K}$ telles que $\forall t \in I \quad B(t) = \sum_{k=1}^{k=n} \beta_k(t) \varphi_k(t)$. Puisque $(\varphi_1(t), \dots, \varphi_n(t))$ est une base de E pour tout t , on en déduit que ψ est solution de (S) ssi $u'_k = \beta_k$. Par quadrature, on peut donc en déduire une solution de (S) .

Donc si on connaît une base $(\varphi_1, \dots, \varphi_n)$ de solutions de (S_0) , on peut par ce procédé trouver les solutions de (S) . La solution générale de (S_0) est $\sum_{k=1}^{k=n} \lambda_k \varphi_k$. On cherche une solution de (S) en remplaçant chacune des constantes λ_k par une fonction numérique inconnue u_k et on remplace dans (S) ; d'où le nom de la méthode. Le problème est alors ramené à décomposer $B(t)$ sur la base $(\varphi_1, \dots, \varphi_n)$ et à effectuer n quadratures.

15.3 Equations différentielles linéaires d'ordre n

15.3.1 Lien avec un système linéaire d'ordre 1

Dans toute cette partie on se donne un intervalle non trivial I de \mathbb{R} et $(n + 1)$ fonctions continues a_1, \dots, a_n et b de I dans \mathbb{K} . On considère les équations différentielles

$$(\mathcal{E}) \quad y^{(n)} + a_1(t)y^{(n-1)} + \dots + a_{n-1}(t)y' + a_n(t)y = b(t)$$

et

$$(\mathcal{E}_0) \quad y^{(n)} + a_1(t)y^{(n-1)} + \dots + a_{n-1}(t)y' + a_n(t)y = 0$$

On appelle solution sur un intervalle $J \subset I$ de (\mathcal{E}) une fonction $\varphi : I \rightarrow \mathbb{K}$ n -fois dérivable sur J telles que

$$\forall t \in I \quad \varphi^{(n)}(t) + a_1(t)\varphi^{(n-1)}(t) + \dots + a_{n-1}(t)\varphi'(t) + a_n(t)\varphi(t) = b(t)$$

Toute fonction solution est nécessairement de classe C^n .

Il est clair que l'ensemble des solutions sur J $\text{Sol}_J(\mathcal{E}_0)$ de (\mathcal{E}_0) est un \mathbb{K} -ev et que l'ensemble des solutions sur J de (\mathcal{E}) est soit vide, soit un espace affine dirigé par $\text{Sol}_J(\mathcal{E}_0)$

Définissons une application $\theta : C^n(I, \mathbb{K}) \rightarrow C^1(I, \mathbb{K}^n)$ en posant, pour $\varphi : I \rightarrow \mathbb{K}$ de classe C^n ,

$$\Phi = \theta(\varphi) = \begin{pmatrix} \varphi \\ \varphi' \\ \vdots \\ \varphi^{(n-1)} \end{pmatrix}$$

θ est linéaire. Introduisons les matrices

$$A(t) = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 1 \\ -a_n(t) & \dots & \dots & -a_2(t) & -a_1(t) \end{pmatrix} \quad \text{et} \quad B(t) = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ b(t) \end{pmatrix}$$

Ceci étant, on vérifie facilement la propriété suivante :

Si φ est solution de (\mathcal{E}) la fonction $\Phi = \theta(\varphi)$ est solution du système $(S) \quad Y' = A(t)Y + B(t)$.

Réciproquement, si Φ est solution de (S) , la première coordonnée φ_1 de Φ est solution de \mathcal{E} et de plus, $\Phi = \theta(\varphi_1)$.

L'étude des équations linéaires scalaires d'ordre n est donc ramenée à l'étude de systèmes linéaires d'ordre 1 particuliers.

15.3.2 Conséquences

Le problème de Cauchy

Le problème de Cauchy pour (S) consiste à se donner $t_0 \in I$ et $Y_0 \in \mathbb{K}^n$ et à chercher les solutions de (S) , Φ telles que $\Phi(t_0) = Y_0$.

L'analogie pour (\mathcal{E}) consiste donc à se donner $t_0 \in I$ et n scalaires y_0, y_1, \dots, y_{n-1} et à rechercher les solutions φ de (S) telles que $\theta(\varphi)$ soit solution du problème de Cauchy pour (S) et la donnée $t_0, Y_0 = \begin{pmatrix} y_0 \\ \vdots \\ y_{n-1} \end{pmatrix}$.

Autrement dit, le problème de Cauchy pour (\mathcal{E}) pour la condition initiale $(t_0, (y_0, \dots, y_{n-1}))$ consiste à chercher les solutions φ de (\mathcal{E}) telles que

$$\varphi(t_0) = y_0, \varphi'(t_0) = y_1, \dots, \varphi^{(n-1)}(t_0) = y_{n-1}$$

Théorème de Cauchy

Le théorème de Cauchy pour les systèmes donne alors

THEOREME 15.3.1

Soient $a_1, \dots, a_n, b : I \rightarrow \mathbb{K}$ $n + 1$ fonctions continues, $t_0 \in I$ et $(y_0, y_1, \dots, y_{n-1}) \in \mathbb{K}^n$. Il existe une solution φ de (\mathcal{E}) définie sur I et une seule φ telle que

$$\varphi(t_0) = y_0, \varphi'(t_0) = y_1, \dots, \varphi^{(n-1)}(t_0) = y_{n-1}$$

THEOREME 15.3.2

Soient $a_1, \dots, a_n : I \rightarrow \mathbb{K}$ n fonctions continues, $t_0 \in I$. L'application

$$\text{Sol}(\mathcal{E}_0) \rightarrow \mathbb{K}^n \quad \varphi \rightarrow \left(\varphi(t_0), \varphi'(t_0), \dots, \varphi^{(n-1)}(t_0) \right)$$

est un isomorphisme de l'espace vectoriel $\text{Sol}(\mathcal{E}_0)$ sur \mathbb{K}^n .

COROLLAIRE 15.3.1

L'espace $\text{Sol}(\mathcal{E}_0)$ est de dimension n . L'espace $\text{Sol}(\mathcal{E})$ est un espace affine de dimension n , dirigé par $\text{Sol}(\mathcal{E}_0)$.

COROLLAIRE 15.3.2

Soient $\varphi_1, \dots, \varphi_n$ n solutions de l'équation (\mathcal{E}_0) et $\Phi_k = \theta(\varphi_k)$. Les propriétés suivantes sont équivalentes :

- (1) $(\varphi_1, \dots, \varphi_n)$ est une base de $\text{Sol}(\mathcal{E}_0)$.
- (2) $\exists t_0 \in I$ $(\Phi_1(t_0), \dots, \Phi_n(t_0))$ est une base de \mathbb{K}^n .
- (3) $\forall t \in I$ $(\Phi_1(t), \dots, \Phi_n(t))$ est une base de \mathbb{K}^n .

La matrice dans la base canonique de \mathbb{K}^n de $(\Phi_1(t), \dots, \Phi_n(t))$ est la matrice

$$W(\varphi_1, \dots, \varphi_n)(t) = \begin{pmatrix} \varphi_1(t) & \cdots & \varphi_n(t) \\ \varphi_1'(t) & \cdots & \varphi_n'(t) \\ \vdots & \ddots & \vdots \\ \varphi_1^{(n-1)}(t) & \cdots & \varphi_n^{(n-1)}(t) \end{pmatrix}$$

Cette matrice est appelée la matrice fondamentale du système de solutions $(\varphi_1, \dots, \varphi_n)$. C'est la matrice wronskienne du système (Φ_1, \dots, Φ_n) de solutions de $Y' = A(t)Y$. Son déterminant est le wronskien de ce système. Comme $\text{tr}(A(t)) = -a_1(t)$ la fonction $w(t) = \det(W(\varphi_1, \dots, \varphi_n)(t))$ vérifie l'équation linéaire scalaire du premier ordre

$$w'(t) = -a_1(t)w(t)$$

et donc

$$w(t) = w(t_0) \exp\left(-\int_{t_0}^t a_1(s) ds\right)$$

15.4 Equations différentielles linéaires scalaires du second ordre

Nous allons reprendre et détailler les résultats généraux obtenus au paragraphe précédents dans le cas des équations linéaires scalaires du second ordre. Dans toute cette partie, I est un intervalle non trivial de \mathbb{R} , $p, q, f : I \rightarrow \mathbb{K}$ sont des fonctions continues. On considère les équations différentielles

$$\begin{aligned} (E_0) \quad & y'' + p(x)y' + q(x)y = 0 \\ (E) \quad & y'' + p(x)y' + q(x)y = f(x) \end{aligned}$$

15.4.1 Résultats généraux

On notera \mathcal{S}_0 (resp. \mathcal{S}) le \mathbb{K} -espace vectoriel (resp. le \mathbb{K} -espace affine) des solutions sur I de l'équation (E_0) (resp. (E)). On sait que (\mathcal{S}_0) est de dimension 2.

Le système linéaire d'ordre 1 associé à (E_0) est

$$(S_0) \quad Y' = A(x)Y \quad \text{avec} \quad A(x) = \begin{pmatrix} 0 & 1 \\ -q(x) & -p(x) \end{pmatrix}$$

et celui associé à (E) est

$$(S) \quad Y' = A(x)Y + B(x) \quad \text{avec} \quad B(x) = \begin{pmatrix} 0 \\ f(x) \end{pmatrix}$$

Si une fonction $y : I \rightarrow \mathbb{K}$ de classe C^2 est solution de (E_0) (resp. de (E)) la fonction vectorielle $Y = \begin{pmatrix} y \\ y' \end{pmatrix}$ est solution de (S_0) (resp. de (S)). Réciproquement, si $Y : I \rightarrow \mathbb{K}^2$ est solution de (S_0) (resp. de (S)), la fonction Y est de la forme $Y = \begin{pmatrix} y \\ y' \end{pmatrix}$ et y est solution de (E_0) (resp. de (E)).

L'énoncé du théorème de Cauchy-Lipschitz devient

Pour tout $x_0 \in I$ et $(y_0, y'_0) \in \mathbb{K}^2$ il existe une solution φ de (E) et une seule définie sur I tout entier et telle que $\varphi(x_0) = y_0$ et $\varphi'(x_0) = y'_0$.

En particulier, pour tout $x_0 \in I$, l'application $\mathcal{S}_0 \rightarrow \mathbb{K}^2$ qui à φ associe $(\varphi(x_0), \varphi'(x_0))$ est un isomorphisme de \mathcal{S}_0 sur \mathbb{K}^2 .

15.4.2 Méthode de variation des constantes

Soit (u, v) une base de solutions de (E_0) . Les fonctions $U = \begin{pmatrix} u \\ u' \end{pmatrix}$ et $V = \begin{pmatrix} v \\ v' \end{pmatrix}$ forment une base de l'espace des solutions de (S_0) dont le wronskien est

$$w = \det(U, V) = uv' - u'v$$

Dans ce cas particulier, puisque u et v sont solutions de (E_0) on a de suite

$$w' = uv'' - u''v = u(-pv' - qv) - (-pu' - qu)v = -pw$$

et donc

$$w(x) = w(x_0) \exp\left(-\int_{x_0}^x p(s)ds\right)$$

Soit $\psi : I \rightarrow \mathbb{K}$ de classe C^2 et $\Psi = \begin{pmatrix} \psi \\ \psi' \end{pmatrix}$. Le lemme 15.2.2 garantit l'existence de deux fonctions λ et μ de I dans \mathbb{K} de classe C^1 telles que $\Psi = \lambda U + \mu V$ soit

$$\begin{cases} \psi & = \lambda u + \mu v \\ \psi' & = \lambda u' + \mu v' \end{cases}$$

ce qui équivaut à

$$\begin{cases} \psi & = \lambda u + \mu v \\ 0 & = \lambda' u + \mu' v \end{cases}$$

On en déduit

$$\psi'' + p\psi' + q\psi = \lambda'u' + \mu'v'$$

Par conséquent, ψ est solution de (E) ssi $\psi = \lambda u + \mu v$ avec

$$\forall x \in I \quad \begin{cases} \lambda'(x)u(x) + \mu'(x)v(x) & = & 0 \\ \lambda'(x)u'(x) + \mu'(x)v'(x) & = & f(x) \end{cases}$$

Pour tout x de I , le système linéaire aux inconnues $\lambda'(x), \mu'(x)$ est de Cramer. Son déterminant n'est autre que le wronskien $w(x) = u(x)v'(x) - u'(x)v(x) \neq 0$. On obtient de suite

$$\lambda'(x) = -\frac{f(x)v(x)}{w(x)}; \quad \mu'(x) = \frac{u(x)f(x)}{w(x)}$$

D'où

$$\lambda(x) = -\int_{x_0}^x \frac{f(t)v(t)}{w(t)} dt + a; \quad \mu(x) = \int_{x_0}^x \frac{u(t)f(t)}{w(t)} dt + b$$

a et b étant des constantes arbitraires. En particulier, la solution ψ_0 de (E) nulle en x_0 ainsi que sa dérivée est donnée par

$$\psi_0(x) = \int_{x_0}^x \frac{u(t)v(x) - u(x)v(t)}{u(t)v'(t) - u'(t)v(t)} f(t) dt$$

Remarquons que puisque $\psi' = \lambda u' + \mu v'$, on a en prime, sans utiliser de théorème de dérivation sous le signe somme

$$\psi_0'(x) = \int_{x_0}^x \frac{u(t)v'(x) - u'(x)v(t)}{u(t)v'(t) - u'(t)v(t)} f(t) dt$$

La solution générale de (E) est alors $\psi = \psi_0 + au + bv$, a et b étant des constantes arbitraires.

Exemple

Considérons l'équation (E) $y'' + y = \frac{2}{\cos^3 x}$ sur l'intervalle $I =]-\frac{\pi}{2}, \frac{\pi}{2}[$.

L'équation linéaire associée est $y'' + y = 0$. Une base de solutions est (u, v) avec $u(x) = \cos x$ et $v(x) = \sin x$. Le wronskien est constant égal à 1. On cherche une solution particulière de l'équation (E) sous la forme $\psi(x) = \lambda(x) \cos x + \mu(x) \sin x$ en imposant

$$\lambda'(x) \cos x + \mu'(x) \sin x = 0$$

De sorte que $\psi'(x) = -\lambda(x) \sin x + \mu(x) \cos x$. On dérive et on reporte dans l'équation (E). On obtient

$$-\lambda'(x) \sin x + \mu'(x) \cos x = \frac{2}{\cos^3 x}$$

En résolvant le système de Cramer formé par les deux dernières équations, il vient

$$\lambda'(x) = -\frac{2 \sin x}{\cos^3 x}; \quad \mu'(x) = \frac{2}{\cos^2 x}$$

On peut choisir

$$\lambda(x) = -\frac{1}{\cos^2 x} \quad \text{et} \quad \mu(x) = 2 \tan x$$

ce qui donne une solution particulière de (E) :

$$\psi(x) = -\frac{1}{\cos x} + 2 \frac{\sin^2 x}{\cos x} = -\frac{\cos 2x}{\cos x}$$

La solution générale de (E) sur I est donc $x \rightarrow -\frac{\cos 2x}{\cos x} + a \cos x + b \sin x$, a et b étant des constantes arbitraires.

15.4.3 Cas où l'on connaît une solution ne s'annulant jamais de l'équation homogène

Supposons que $u : I \rightarrow \mathbb{K}$ soit une solution de l'équation $y'' + p(x)y' + q(x)y = 0$ telle que $\forall x \in I \quad u(x) \neq 0$. Soit $\psi : I \rightarrow \mathbb{K}$ de classe C^2 . Puisque u est de classe C^2 et ne s'annule pas, la fonction $z = \frac{\psi}{u}$ est aussi de classe C^2 . On peut donc poser $\psi = uz$ avec z de classe C^2 . En dérivant, il vient $\psi' = u'z + uz'$; $\psi'' = u''z + 2u'z' + uz''$. On a donc

$$\psi'' + p\psi' + q\psi = (u'' + pu' + qu)z + (2u' + pu)z' + uz'' = (2u' + pu)z' + uz''$$

Donc ψ est solution de (E) ssi $(2u' + pu)z' + uz'' = f$ ce qui s'écrit aussi puisque u ne s'annule pas

$$z'' + \left(2\frac{u'}{u} + p\right)z' = f$$

On est donc ramené à une équation linéaire du premier ordre en la fonction inconnue z' . On trouve ensuite z par quadrature.

15.4.4 Propriétés élémentaires des zéros des solutions

On prend ici $\mathbb{K} = \mathbb{R}$. On considère l'équation $(E_0) : y'' + p(x)y' + q(x)y = 0$ où p et q sont des fonctions continues définies sur I à valeurs réelles et on étudie les solutions réelles de cette équation.

PROPOSITION 15.4.1

Soit u une solution non identiquement nulle de (E_0) . Les zéros de u sont isolés.

preuve

Soit $x_0 \in I$ tel que $u(x_0) = 0$. On a

$$u'(x_0) \neq 0$$

En effet, dans le cas contraire, u serait une solution telle que $u(x_0) = u'(x_0) = 0$. Or d'après le théorème de Cauchy, la fonction nulle est l'unique solution qui vérifie cette condition initiale et u est supposée non nulle.

On a alors $u(x) = (x - x_0)(u'(x_0) + \varepsilon(x))$ avec $\lim_{x \rightarrow x_0} \varepsilon(x) = 0$. Soit $\eta > 0$ tel que $|x - x_0| < \eta \Rightarrow |\varepsilon(x)| < |u'(x_0)|$. Pour $x \in]x_0 - \eta, x_0 + \eta[$ et $x \neq x_0$, on a $u(x) \neq 0$.

En fait, dans $]x_0 - \eta, x_0 + \eta[$, $u(x)$ est du signe de $u'(x_0)(x - x_0)$. Donc u change de signe en x_0 .

COROLLAIRE 15.4.1

Soit u une solution non identiquement nulle de (E_0) . u n'a qu'un nombre fini de zéros dans tout intervalle compact contenu dans I .

preuve

Soit $K = [a, b] \subset I$ et $Z = \{x \in I \mid u(x) = 0\}$. L'ensemble Z est fermé dans I , donc $Z_K = Z \cap K$ est fermé dans K compact, donc est compact. Or un compact formé de points isolés est fini. ■

COROLLAIRE 15.4.2

Soit u une solution non identiquement nulle de (E_0) . L'ensemble Z des zéros de u est au plus dénombrable.

preuve

Si I est compact, Z est fini. Sinon I peut s'écrire comme réunion d'une suite d'intervalles compacts K_n (le vérifier !). On a alors $Z = \bigcup_{n \geq 0} (Z \cap K_n)$ et chacun des $Z \cap K_n$ est fini. Or une réunion dénombrable d'ensembles finis est au plus dénombrable. ■

COROLLAIRE 15.4.3

Soit u une solution non identiquement nulle de (E_0) sur un intervalle I de la forme $I = [a, b[$, $b \leq +\infty$. Si u a une infinité de zéros sur I on peut les numéroter en une suite strictement croissante $(x_j)_{j \geq 0}$.

preuve

Soit $(b_n)_{n \geq 0}$ une suite strictement croissante de points de I de limite b avec $b_0 = a$. Soit $I_n = [b_0, b_n]$. u n'a qu'un nombre fini (éventuellement nul) de zéros dans I_n , donc aussi dans $J_n = I_n \setminus I_{n-1}$ pour $n \geq 1$. Soit m_n le nombre de zéros de u dans J_n et m_0 le nombre de zéros de u dans I_0 . On numérote les zéros de u dans I_0 de manière croissante en $x_0 < x_1 < \dots < x_{m_0-1}$, les zéros de u dans J_1 en $x_{m_0} < \dots < x_{m_0+m_1-1}$ etc...

Remarque

Si A est un sous ensemble dénombrable de \mathbb{R}_+ par exemple, on peut toujours numéroter les éléments de A en une suite (a_n) , par définition d'un ensemble dénombrable. Mais il n'est pas toujours possible de le faire de manière croissante comme le prouve l'exemple de \mathbb{Q}_+ .

PROPOSITION 15.4.2

Soient u et v deux solutions indépendantes de (E_0) sur I .

1. u et v n'ont pas de zéros communs.

2. Si a et b sont deux zéros consécutifs de u , il existe un zéro de v dans l'intervalle $]a, b[$.

preuve

1. Toute solution de (E_0) , est combinaison linéaire de u et v . Si on avait $u(x_0) = v(x_0) = 0$ pour un $x_0 \in I$, toute solution de

(E₀) s'annulerait en x_0 ce qui est en contradiction avec le théorème de Cauchy.

2. Supposons $u(a) = u(b) = 0$ avec $a < b$ et $u(x) \neq 0$ pour $a < x < b$. La fonction u garde un signe constant sur l'intervalle $]a, b[$. Quitte à remplacer u par $-u$, on peut supposer $u(x) > 0$ pour $a < x < b$. Comme $u'(a) = \lim_{x \rightarrow a} \frac{u(x) - u(a)}{x - a}$ on a

$u'(a) \geq 0$. Par ailleurs on sait que $u'(a) \neq 0$ donc $u'(a) > 0$ et de même $u'(b) < 0$.

Considérons alors le wronskien $w(x) = u(x)v'(x) - u'(x)v(x)$. On a $w(a) = -u'(a)v(a)$ et $w(b) = -u'(b)v(b)$. Or w ne s'annule pas, donc est de signe constant sur $[a, b]$. On en déduit $w(a)w(b) > 0$ donc $u'(a)u'(b)v(a)v(b) > 0$. Comme $u'(a)u'(b) < 0$, il vient $v(a)v(b) < 0$ donc v s'annule entre a et b .

15.4.5 Exemples de méthode de résolution

Nous allons passer en revue différentes méthodes de résolution sur un exemple, l'équation

$$(E) \quad 4xy'' + 2y' - y = 0$$

Notons d'abord que le théorème de Cauchy-Lipschitz s'applique sur \mathbb{R}_+^* et sur \mathbb{R}_-^* .

1. Utilisation de séries entières

Nous allons rechercher les solutions de (E) développables en série entière.

• Soit $f(x) = \sum_{n \geq 0} a_n x^n$ la somme d'une série entière de rayon de convergence $R > 0$. Pour $x \in]-R, R[$ on a

$$\begin{aligned} 4xf''(x) + 2f'(x) - f(x) &= \sum_{n \geq 0} 4(n+1)na_{n+1}x^n + \sum_{n \geq 0} 2(n+1)a_{n+1}x^n - \sum_{n \geq 0} a_n x^n \\ &= \sum_{n \geq 0} [2(n+1)(2n+1)a_{n+1} - a_n] x^n \end{aligned}$$

f est solution de (E) ssi la suite (a_n) vérifie la relation de récurrence $(\mathcal{R}) \quad \forall n \in \mathbb{N} \quad a_{n+1} = \frac{1}{(2n+1)(2n+2)} a_n$

• Si $a_0 = 0$ la relation \mathcal{R} implique $a_n = 0$ pour tout n . Pour $a_0 \neq 0$ fixé la relation (\mathcal{R}) définit une suite unique dont tous les termes sont non nuls. L'application de la règle de d'Alembert prouve que la série entière associée à un rayon de convergence infini et donc que la somme f de cette série est solution sur \mathbb{R} de (E).

• Un calcul immédiat donne $a_n = \frac{1}{(2n)!} a_0$. Pour $a_0 = 1$, on obtient $f_0(x) = \sum_{n \geq 0} \frac{x^n}{(2n)!}$. Il vient

$$f_0(x) = \begin{cases} \sum_{n \geq 0} \frac{(\sqrt{x})^{2n}}{(2n)!} = \cosh(\sqrt{x}) & \text{si } x > 0 \\ \sum_{n \geq 0} (-1)^n \frac{(\sqrt{-x})^{2n}}{(2n)!} = \cos(\sqrt{-x}) & \text{si } tx < 0 \\ 1 & \text{si } x = 0 \end{cases}$$

(E) admet donc des solutions développables en série entière sur \mathbb{R} tout entier. L'ensemble \mathcal{D} des solutions développables en série entière de (E) forme donc un espace vectoriel de dimension 1 engendré par la fonction f_0 .

2. Méthode de variation d'une constante (15.4.3)

Posons $I = \mathbb{R}_+^*$. La fonction $\varphi(x) = \cosh(\sqrt{x})$ est une solution de (E) sur I qui ne s'annule pas. On peut donc chercher les solutions y de (E) sur I sous la forme $y = z\varphi$. Il vient

$$4xy'' + 2y' - y = 4x(z''\varphi + 2z'\varphi' + z\varphi'') + 2(z'\varphi + z\varphi') - z\varphi = 4x\varphi z'' + (8x\varphi' - 2\varphi)z'$$

donc y est solution de (E) sur I ssi z est solution de l'équation

$$z'' + \left(2 \frac{\varphi'(x)}{\varphi(x)} - \frac{1}{2x} \right) z' = 0$$

d'où

$$z'(x) = C \frac{\sqrt{x}}{\varphi(x)^2} = C \frac{\sqrt{x}}{\cosh^2(\sqrt{x})} \text{ et } z(x) = \frac{C}{2} \tanh(\sqrt{x}) + D$$

où C et D sont des constantes.

On en déduit les solutions de (E) sur \mathbb{R}_+^* :

$$f(x) = A \sinh(\sqrt{x}) + D \cosh(\sqrt{x}), \quad A, D \text{ constantes}$$

3. Utilisation de changement de variables

Posons $J = \mathbb{R}_-^*$. On dispose d'une solution sur J , la fonction $\psi(x) = \cos(\sqrt{-x})$. Nous allons résoudre l'équation sur J au moyen du changement de variables $t = \sqrt{-x}$. Plus précisément, l'application $\theta : \mathbb{R}_-^* \rightarrow \mathbb{R}_+^*$ définie par $\theta(x) = \sqrt{-x}$ est un C^∞ difféomorphisme qui induit une bijection $\Theta : C^2(\mathbb{R}_+^*, \mathbb{R}) \rightarrow C^2(\mathbb{R}_-^*, \mathbb{R})$ définie par $\Theta(g) = g \circ \theta$ la bijection réciproque étant $\Theta^{-1}(f) = f \circ \theta^{-1}$ où $\theta^{-1}(t) = -t^2$.

Soit $f \in C^2(\mathbb{R}_-, \mathbb{R})$ et $F = f \circ \theta^{-1}$. Posons $t = \theta(x) = \sqrt{-x}$. On a $\theta'(x) = -\frac{1}{2\sqrt{-x}} = -\frac{1}{2t}$ d'où

$$f(x) = F(\theta(x)) = F(t)$$

$$f'(x) = F'(\theta(x))\theta'(x) = -\frac{1}{2t}F'(t)$$

$$f''(x) = \left(\frac{1}{2t^2}F'(t) - \frac{1}{2t}F''(t) \right) \theta'(x) = -\frac{1}{2t} \left(\frac{1}{2t^2}F'(t) - \frac{1}{2t}F''(t) \right) = \frac{1}{4t^2}F''(t) - \frac{1}{4t^3}F'(t)$$

$$4xf''(x) + 2f'(x) - f(x) = -F''(t) - F(t)$$

Donc f est solution de (E) sur \mathbb{R}_-^* ssi F est solution sur \mathbb{R}_+^* de $y'' + y = 0$ ce qui donne $F(t) = A \cos t + B \sin t$, avec A, B constantes. Les solutions sur \mathbb{R}_-^* de (E) sont donc les fonctions

$$f(x) = A \cos(\sqrt{-x}) + B \sin(\sqrt{-x})$$

4. Raccordements

Pour finir, cherchons les solutions sur \mathbb{R} de (E). On sait déjà que les fonctions λf_0 , $\lambda \in \mathbb{R}$ sont solutions. Soit u une solution sur \mathbb{R} de (E). La restriction de u à \mathbb{R}_+^* (resp. à \mathbb{R}_-^*) est solution sur \mathbb{R}_+^* (resp. sur \mathbb{R}_-^*). Donc il existe des constantes A, B, C, D telles que

$$u(x) = \begin{cases} A \cos(\sqrt{-x}) + B \sin(\sqrt{-x}) & \text{si } x < 0 \\ C \cosh(\sqrt{x}) + D \sinh(\sqrt{x}) & \text{si } x > 0 \end{cases}$$

La continuité de u en 0 implique $A = C$. Mais la fonction $v = u - Af_0$ est solution sur \mathbb{R} . Or dans un voisinage à droite (resp. à gauche) de 0, on a $v(x) \sim D\sqrt{x}$ (resp. $v(x) \sim B\sqrt{-x}$). Comme v doit être dérivable en 0, on en déduit (en regardant le taux d'accroissement de v) que nécessairement $B = D = 0$.

Finalement, les seules solutions sur \mathbb{R} de (E) sont les fonctions λf_0 , $\lambda \in \mathbb{R}$.

15.5 Appendice : preuve du théorème de Cauchy

Cette démonstration n'est pas au programme. Mais c'est un bon exercice sur l'utilisation de suites de fonctions.

E est un \mathbb{K} espace vectoriel de dimension finie et on se donne $A : I \rightarrow L(E)$ et $B : I \rightarrow E$ continues.

1. Soit $\varphi : I \rightarrow E$ continue. On a l'équivalence

$$\left. \begin{array}{l} \varphi \text{ solution de } Y' = A(t)Y + B(t) \\ \varphi(t_0) = Y_0 \end{array} \right\} \Leftrightarrow \forall t \in I \quad \varphi(t) = Y_0 + \int_{t_0}^t (A(s)Y(s) + B(s))ds$$

En effet, si φ est solution de (S) vérifiant $\varphi(t_0) = Y_0$ φ est C^1 donc

$$\varphi(t) = \varphi(t_0) + \int_{t_0}^t \varphi'(x)dx = Y_0 + \int_{t_0}^t (A(s)\varphi(s) + B(s))ds$$

Réciproquement, si $\varphi : I \rightarrow E$ est une fonction continue vérifiant

$$(S') \quad \varphi(t) = Y_0 + \int_{t_0}^t (A(s)\varphi(s) + B(s)) ds$$

puisque la fonction sous le signe intégrale est une fonction continue de s , φ est C^1 et $\varphi'(t) = A(t)\varphi(t) + B(t)$ pour tout $t \in I$. D'autre part, on a manifestement $\varphi(t_0) = Y_0$.

Donc résoudre le problème de Cauchy considéré sur I revient à chercher une fonction continue $I \rightarrow E$ vérifiant (S').

2. Existence :

On va utiliser la méthode des approximations successives : on définit une suite de fonctions $\varphi_n : I \rightarrow E$ en posant $\varphi_0(t) = 0$ pour tout t et pour $n \geq 0$

$$\forall t \in I \quad \varphi_{n+1}(t) = Y_0 + \int_{t_0}^t (A(s)\varphi_n(s) + B(s)) ds$$

• La suite (φ_n) est uniformément convergente sur tout compact de I . En effet, soit $K = [a, b]$ un intervalle compact contenu dans I et contenant t_0 . Munissons E d'une norme $\| \cdot \|$ et $L(E)$ de la norme subordonnée que nous noterons aussi $\| \cdot \|$. Posons

$$C = \sup_{t \in K} \|A(t)\| \quad M = \sup_{t \in K} \|\varphi_1(t) - \varphi_0(t)\|$$

Ces quantités existent puisque les fonctions $t \rightarrow \|A(t)\|$ et $t \rightarrow \|\varphi_1(t) - \varphi_0(t)\|$ sont continues sur le compact K . (Avec le choix fait ici de φ_0 on a $M = \|Y_0\|$).

Montrons par récurrence sur $n \in \mathbb{N}^*$ que

$$\forall t \in K \quad \|\varphi_n(t) - \varphi_{n-1}(t)\| \leq M \frac{(C|t - t_0|)^{n-1}}{(n-1)!}$$

Cette relation est vérifiée pour $n = 1$ par définition de C et M . Supposons la établie pour un entier n . On a

$$\varphi_{n+1}(t) - \varphi_n(t) = \int_{t_0}^t [A(s)(\varphi_n(s) - \varphi_{n-1}(s))] ds$$

d'où

$$\|\varphi_{n+1}(t) - \varphi_n(t)\| \leq \left| \int_{t_0}^t \|A(s)\| \|\varphi_n(s) - \varphi_{n-1}(s)\| ds \right| \leq \left| \int_{t_0}^t C \frac{(C|s - t_0|)^{n-1}}{(n-1)!} ds \right| = M \frac{(C|t - t_0|)^n}{n!}$$

Si on note ℓ la longueur de l'intervalle compact K , on en déduit

$$\forall t \in K \quad \|\varphi_n(t) - \varphi_{n-1}(t)\| \leq M \frac{(C\ell)^{n-1}}{(n-1)!}$$

On en déduit la convergence normale, donc uniforme de la série de fonctions de terme général $\varphi_n - \varphi_{n-1}$ et par conséquent la convergence uniforme de la suite des sommes partielles φ_n .

• La suite (φ_n) converge donc uniformément sur tout compact K de I vers une fonction $\varphi : I \rightarrow E$ continue.

Sur l'intervalle compact K considéré ci dessus, la suite $\psi_n = A\varphi_n + B$ converge uniformément vers $\psi = A\varphi + B$. En effet, pour $t \in K$, on a $\|\psi(t) - \psi_n(t)\| \leq \|A(t)\| \|\varphi(t) - \varphi_n(t)\| \leq C \sup_{t \in K} \|\varphi(t) - \varphi_n(t)\|$ qui tend vers 0 quand n tend vers l'infini. Par conséquent

$\int_{t_0}^t \psi_n(s) ds$ converge vers $\int_{t_0}^t \psi(s) ds$. De $\varphi_{n+1}(t) = Y_0 + \int_{t_0}^t \psi_n(s) ds$ on déduit, pour tout $t \in K$

$$\varphi(t) = Y_0 + \int_{t_0}^t \psi(s) ds = Y_0 + \int_{t_0}^t (A(s)\varphi(s) + B(s)) ds$$

Ceci a lieu pour tout intervalle compact K contenu dans I et contenant t_0 . L'égalité obtenue est donc vrai pour tout $t \in I$.

3. Unicité

Soient φ_2 et φ_1 deux solutions sur I de (S) telles que $\varphi_1(t_0) = \varphi_2(t_0)$ et $h = \varphi_2 - \varphi_1$. On a donc $h(t) = \int_{t_0}^t A(s)h(s) ds$. Soit K un intervalle compact contenu dans I et contenant t_0 . Posons $m = \sup_{t \in K} \|h(t)\|$. On démontre comme au point 2., par récurrence sur n l'inégalité

$$\forall t \in K \quad \|h(t)\| \leq m \frac{(C|t - t_0|)^n}{n!}$$

Le second membre tend vers 0 quand n tend vers l'infini, donc $h(t) = 0$ pour tout $t \in K$. Par conséquent, K étant arbitraire $h = 0$.

16

Systemes différentiels linéaires à coefficients constants

$\mathbb{K} = \mathbb{R}$ ou \mathbb{C} .

16.1 Exponentielle d'endomorphisme

16.1.1 Définition générale de l'exponentielle

Rappelons qu'une \mathbb{K} -algèbre de Banach est une \mathbb{K} -algèbre A associative, unitaire, normée complète et dont la norme vérifie $\forall u, v \in A \quad \|vu\| \leq \|v\| \|u\|$ et $\|1_A\| = 1$. L'exemple qui nous sera utile dans la suite est $A = M_n(\mathbb{K})$ muni d'une norme sous-multiplicative ou $A = L(E)$ où E est un \mathbb{K} -evn de dimension finie, la norme de $L(E)$ étant la norme subordonnée à la norme de E .

THEOREME 16.1.1

Soient A une algèbre de Banach. La série de terme général $\frac{1}{n!}u^n$ converge pour tout $u \in A$ et converge normalement sur tout sous ensemble borné de A . Sa somme est appelée exponentielle de u et notée $\exp(u)$.

Par convention dans cet énoncé $u^0 = 1_A$. Dans ce chapitre, on réservera la notation e^z à l'exponentielle complexe.

preuve

Soit $R > 0$. Pour $\|u\| \leq R$ on a $\|u^n\| \leq \|u\|^n \leq R^n$ donc $\left\| \frac{u^n}{n!} \right\| \leq \frac{R^n}{n!}$ qui est le terme général d'une série numérique convergente, d'où la convergence normale annoncée. Comme A est complète, toute série absolument convergente d'éléments de A est convergente. ■

Notons que si $\lambda \in \mathbb{K}$, on a $\exp(\lambda 1_A) = e^\lambda 1_A$.

COROLLAIRE 16.1.1

L'application $\exp : A \rightarrow A$ est continue.

C'est la somme d'une série de fonctions continues ($u \rightarrow u^n/n!$) normalement convergente sur tout compact.

Dans la suite, pour $x \in A$ ou $x \in \mathbb{K}$ on posera $S_n(x) = \sum_{k=0}^{k=n} \frac{x^k}{k!}$.

PROPOSITION 16.1.1

Pour tout $u \in A$ et tout $P \in \mathbb{K}[X]$, $P(u)$ et $\exp(u)$ commutent.

preuve

On a pour tout n , $uS_n(u) = S_n(u)u$ donc $u \exp(u) = \exp(u)u$. On utilise ici la continuité des applications linéaires $v \rightarrow uv$

et $v \rightarrow vu$, continuité qui résulte des inégalités $\|uv\| \leq \|u\| \|v\|$ et $\|vu\| \leq \|u\| \|v\|$. Puisque u commute avec $\exp(u)$, il en est de même de toute puissance de u , et de tout polynôme en u .

THEOREME 16.1.2

Soient $u, v \in A$ tels que $uv = vu$. On a $\exp(u + v) = \exp(u) \exp(v)$.

preuve

Posons $T_n = \{(i, j) \mid 0 \leq i \leq n, 0 \leq j \leq n, i + j > n\}$. Puisque u et v commutent, on peut calculer $(u + v)^k$ au moyen de la formule du binôme. Il vient,

$$\begin{aligned} S_n(u + v) - S_n(u)S_n(v) &= \sum_{k=0}^{k=n} \frac{1}{k!} \left(\sum_{r=0}^{r=k} C_k^r u^r v^{k-r} \right) - \left(\sum_{i=0}^{i=n} \frac{u^i}{i!} \right) \left(\sum_{j=1}^{j=n} \frac{v^j}{j!} \right) \\ &= \sum_{r+s \leq n} \frac{u^r v^s}{r! s!} - \sum_{\substack{0 \leq i \leq n \\ 0 \leq j \leq n}} \frac{u^i v^j}{i! j!} \\ &= \sum_{(i,j) \in T_n} \frac{u^i v^j}{i! j!} \end{aligned}$$

On en déduit

$$\|S_n(u + v) - S_n(u)S_n(v)\| \leq \sum_{(i,j) \in T_n} \frac{\|u\|^i \|v\|^j}{i! j!}$$

La second membre, par le même calcul que ci dessus vaut $S_n(\|u\| + \|v\|) - S_n(\|u\|)S_n(\|v\|)$. Quand n tend vers l'infini, il tend vers $e^{\|u\| + \|v\|} - e^{\|u\|} e^{\|v\|} = 0$ d'après les propriétés déjà établies de l'exponentielle réelle.

COROLLAIRE 16.1.2

Pour tout $u \in A$, $\exp(u)$ est inversible d'inverse $\exp(-u)$.

En effet, u et $-u$ commutent et $\exp(0_A) = 1_A$.

16.1.2 Exponentielle d'endomorphismes

Dans la suite E est un \mathbb{K} -espace vectoriel de dimension finie n , $\mathcal{B} = (e_1, \dots, e_n)$ est une base de E . On munit E d'une norme et $L(E)$ de la norme subordonnée.

On munit aussi $M_n(\mathbb{K})$ de la norme subordonnée à une norme de \mathbb{K}^n . Une telle norme est sous-multiplicative i.e. telle que $\|PQ\| \leq \|P\| \|Q\|$ pour toutes $P, Q \in M_n(\mathbb{K})$.

L'application $u \rightarrow \text{Mat}_{\mathcal{B}}(u)$ de $L(E)$ dans $M_n(\mathbb{K})$ est un isomorphisme linéaire. Comme $L(E)$ est de dimension finie elle est continue. $L(E)$ et $M_n(\mathbb{K})$ sont des algèbres de Banach.

Lemme 16.1.1

1. Soit $u \in L(E)$ et $M = \text{Mat}_{\mathcal{B}}(u)$. Alors $\text{Mat}_{\mathcal{B}}(\exp(u)) = \exp(M)$.
2. Soient $A \in M_n(\mathbb{K})$ et $P \in GL(n, \mathbb{K})$. On a $\exp(P^{-1}AP) = P^{-1} \exp(A)P$.

En effet, pour tout p on a $\text{Mat}_{\mathcal{B}}(S_p(u)) = S_p(M)$. La conclusion en résulte en prenant la limite quand p tend vers l'infini et en utilisant la continuité de l'application $u \rightarrow \text{Mat}_{\mathcal{B}}(u)$. De même, on déduit 2. de l'égalité $S_p(P^{-1}AP) = P^{-1}S_p(A)P$.

THEOREME 16.1.3

Soit $u \in L(E)$.

1. L'application $\varphi : \mathbb{R} \rightarrow L(E)$ définie par $\varphi(t) = \exp(tu)$ est un morphisme du groupe $(\mathbb{R}, +)$ dans le groupe $(GL(E), \circ)$.
 2. L'application φ est de classe C^1 et $\forall t \in \mathbb{R} \varphi'(t) = u \circ \exp(tu) = \exp(tu) \circ u$.
- On a un énoncé analogue pour $t \rightarrow \exp(tA)$ quand $A \in M_n(\mathbb{K})$.

preuve

1. Pour $t, s \in \mathbb{R}$, su et tu commutent donc $\varphi(t+s) = \exp((t+s)u) = \exp(tu)\exp(su) = \varphi(t)\varphi(s)$.
2. Montrons que φ est dérivable en 0 avec $\varphi'(0) = u$. On a

$$\frac{\varphi(t) - \varphi(0)}{t} - u = \frac{1}{t}(\exp(tu) - Id) - u = \sum_{n \geq 2} \frac{t^{n-1}}{n!} u^n = tH(t)$$

avec $H(t) = \sum_{n \geq 0} \frac{t^n}{n!} u^{n+2}$. Pour $|t| \leq 1$ on a $\|H(t)\| \leq \sum_{n \geq 0} \frac{1}{n!} \|u\|^{n+2}$ qui est une constante, donc $\lim_{t \rightarrow 0} tH(t) = 0$ ce qui prouve la conclusion.

Maintenant, si $t_0 \in \mathbb{R}$ on a

$$\frac{1}{t-t_0}(\exp(tu) - \exp(t_0u)) = \left(\frac{1}{t-t_0} \exp((t-t_0)u) - id \right) \exp(t_0u)$$

quantité qui tend vers $u \exp(t_0u)$ quand t tend vers 0.

Ce théorème admet une réciproque

THEOREME 16.1.4

Soit $\varphi : \mathbb{R} \rightarrow GL(E)$ un morphisme de classe C^1 et $u = \varphi'(0)$. On a pour tout $t \in \mathbb{R}$, $\varphi(t) = \exp(tu)$.

preuve

On a $\varphi(s+t_0) = \varphi(t_0)\varphi(s) = \varphi(s)\varphi(t_0)$ donc en dérivant par rapport à s en $s=0$, $\varphi'(t_0) = \varphi(t_0)\varphi'(0) = \varphi'(0)\varphi(t_0)$, donc pour tout $t \in \mathbb{R}$ $\varphi'(t) = \varphi(t)u = u\varphi(t)$.

Posons $\psi(t) = \exp(-tu)\varphi(t)$. ψ est C^1 car $t \rightarrow \exp(-tu)$ et φ le sont. On a $\psi'(t) = -\exp(-tu)u\varphi(t) + \exp(-tu)\varphi'(t) = \exp(-tu)(-u\varphi(t) + \varphi'(t)) = 0$ donc ψ est constante sur \mathbb{R} , $\psi(t) = \psi(0) = Id$ et $\varphi(t) = \exp(tu)$.

EXEMPLE 16.1.1

Soit $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. On a, pour tout $t \in \mathbb{R}$

$$\exp(tJ) = \begin{pmatrix} \cos t & -\sin t \\ \sin t & \cos t \end{pmatrix}$$

Première preuve

On a $J^2 = -I_2$ donc $J^{2k} = (-1)^k I_2$ et $J^{2k+1} = J^{2k}J = (-1)^k J$. On en déduit

$$S_{2n}(tJ) = \left(\sum_{k=0}^{k=n} (-1)^k \frac{t^{2k}}{(2k)!} \right) I_2 + \left(\sum_{k=1}^n (-1)^{k-1} \frac{t^{2k-1}}{(2k-1)!} \right) J$$

d'où le résultat en prenant la limite quand n tend vers l'infini.

Deuxième preuve

On identifie \mathbb{R}^2 orienté muni de sa structure euclidienne usuelle avec \mathbb{C} . L'endomorphisme u de \mathbb{R}^2 de matrice J dans la base canonique devient la multiplication par $i : u(z) = iz$ donc l'endomorphisme associé à tJ est $tu : z \rightarrow itz$. Par conséquent $S_n(u)(z) = S_n(it)z$ et en prenant la limite $\exp(tu)(z) = e^{it}z$. Donc $\exp(tu)$ est la rotation d'angle t . Sa matrice, $\exp(tJ)$ dans la base canonique est donc la matrice indiquée.

16.1.3 Exponentielle d'endomorphismes : propriétés algébriques

Soit $u \in L(E)$. Tout polynôme en u commute avec $\exp(u)$, donc les espaces propres et les espaces caractéristiques de u sont stables par $\exp(u)$.

THEOREME 16.1.5

Soient $u \in L(E)$ et $\lambda \in \mathbb{K}$ une valeur propre de u . Alors e^λ est valeur propre de $\exp(u)$ et $\ker(u - \lambda Id) \subset \ker(\exp(u) - e^\lambda id)$.

preuve

Si $v \in \ker(u - \lambda Id)$ on a $u^p(v) = \lambda^p v$ pour tout entier p , donc $S_n(u)(v) = S_n(\lambda)v$. Pour v fixé, l'application $L(E) \rightarrow E$ qui à f associe $f(v)$ est continue donc $\lim_{n \rightarrow \infty} (S_n(u)(v)) = \lim_{n \rightarrow \infty} (S_n(\lambda)(v)) = \exp(u)(v)$. Comme $S_n(\lambda)v$ tend vers $e^\lambda v$, le résultat est prouvé.

THEOREME 16.1.6

Soit $u \in L(E)$ dont le polynôme caractéristique χ_u est scindé sur \mathbb{K} .

Si $\chi_u(X) = \prod_{j=1}^{j=r} (\lambda_j - X)^{m_j}$ alors $\chi_{\exp(u)} = \prod_{j=1}^{j=r} (e^{\lambda_j} - X)^{m_j}$

preuve

Dans une base \mathcal{B} convenable, la matrice T de u est triangulaire supérieure avec une diagonale égale à

$$\left(\underbrace{\lambda_1, \dots, \lambda_1}_{m_1 \text{ fois}}, \dots, \underbrace{\lambda_r, \dots, \lambda_r}_{m_r \text{ fois}} \right)$$

La matrice $S_n(T)$ de $S_n(u)$ dans la même base est triangulaire supérieure avec une diagonale égale à

$$\left(\underbrace{S_n(\lambda_1), \dots, S_n(\lambda_1)}_{m_1 \text{ fois}}, \dots, \underbrace{S_n(\lambda_r), \dots, S_n(\lambda_r)}_{m_r \text{ fois}} \right)$$

En prenant la limite, la matrice $\exp(T)$ de $\exp(u)$ est triangulaire supérieure avec une diagonale égale à

$$\left(\underbrace{e^{\lambda_1}, \dots, e^{\lambda_1}}_{m_1 \text{ fois}}, \dots, \underbrace{e^{\lambda_r}, \dots, e^{\lambda_r}}_{m_r \text{ fois}} \right)$$

d'où la conclusion.

COROLLAIRE 16.1.3

Soit $u \in L(E)$. On a $\det(\exp(u)) = e^{\text{tr}(u)}$.

preuve

Soient \mathcal{B} une base de E et $A = \text{Mat}_{\mathcal{B}}(u) \in M_n(\mathbb{K})$. On peut appliquer le théorème qui précède en considérant A comme appartenant à $M_n(\mathbb{C})$. On en déduit $\det(\exp(u)) = \det(\exp(A)) = e^{\text{tr}(A)} = e^{\text{tr}(u)}$.

THEOREME 16.1.7

Soient $u \in L(E)$ dont le polynôme caractéristique est scindé sur \mathbb{K} et $u = D + N$ la décomposition de Dunford de u . La décomposition de Dunford de $\exp(u)$ est $\Delta + \nu$ avec $\Delta = \exp(D)$ et $\nu = \exp(D)(\exp(N) - I_n)$ avec $\exp(N) = \sum_{k=0}^{\beta-1} \frac{1}{k!} N^k$ où β est l'indice de nilpotence de N .

preuve

Par hypothèse D et N commutent, donc $\exp(u) = \exp(D) \exp(N)$. On peut donc écrire

$$\exp(u) = \exp(D) + \exp(D)(\exp(N) - I_n)$$

• Il existe une base \mathcal{B} telle que $\text{Mat}_{\mathcal{B}}(D)$ soit diagonale, soit $\text{diag}(\lambda_1, \dots, \lambda_n)$. Alors $\text{Mat}_{\mathcal{B}}(\exp(D)) = \exp(\text{Mat}_{\mathcal{B}}(D)) = \text{diag}(e^{\lambda_1}, \dots, e^{\lambda_n})$ donc $\Delta = \exp(D)$ est diagonalisable.

• Comme $N^\beta = 0$ on a $\exp(N) = \sum_{k=0}^{\beta-1} \frac{1}{k!} N^k$ donc $\exp(N) - I_n = NM$ avec $M = I + \frac{1}{2!}N + \dots + \frac{1}{(\beta-1)!}N^{\beta-1}$. Comme

N et M commutent, $(\exp(N) - I_n)^\beta = N^\beta M^\beta = 0$ donc $(\exp(N) - I_n)$ est nilpotente.

Ensuite, D et N commutent, donc $\exp(D)$ et $\exp(N)$ aussi, donc $\exp(D)$ commute avec $\exp(N) - I_n$. Comme précédemment, on en déduit que leur produit est nilpotent, d'indice de nilpotence au plus β .

• Enfin, $\exp(D)$ et $\exp(D)(\exp(N) - I_n)$ commutent. ■

COROLLAIRE 16.1.4

Soit $u \in L(E)$ dont le polynôme caractéristique est scindé sur \mathbb{K} . $\exp(u)$ est diagonalisable ssi u est diagonalisable.

preuve

$\exp(u)$ diagonalisable $\Leftrightarrow \nu = 0 \Leftrightarrow \exp(D)(\exp(N) - I_n) = 0 \Leftrightarrow \exp(N) = I_n$ car $\exp(D)$ est inversible. Le corollaire sera prouvé si on montre que $\exp(N) = I_n \Leftrightarrow N = 0$. Supposons $N \neq 0$. L'ordre de nilpotence β est donc au moins égal à 2. C'est alors un exercice classique de montrer que $(N, N^2, \dots, N^{\beta-1})$ est un système libre. Il en résulte que $\exp(N) - I_n \neq 0$. ■

16.2 Systèmes différentiels linéaires du premier ordre à coefficients constants

16.2.1 Système homogène

Dans cette partie, on considère un \mathbb{K} -ev de dimension finie E , $A \in L(E)$ et le système différentiel $Y' = AY$. Bien entendu, tous les théorèmes du chapitre précédent s'appliquent, mais on a ici des résultats plus précis.

Pour $A \in L(E)$ et $v \in E$ on notera Av pour $A(v)$.

THEOREME 16.2.1 (Théorème de Cauchy)

Soient $A \in L(E)$, $t_0 \in \mathbb{R}$, $Y_0 \in E$. Le système $Y' = AY$ admet une unique solution φ sur \mathbb{R} vérifiant $\varphi(t_0) = Y_0$. Elle est donnée par

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

preuve

Soit $\varphi(t) = \exp((t - t_0)A)Y_0$. On a $\varphi(t_0) = Y_0$ et en utilisant le théorème 16.1.3 $\varphi'(t) = A \exp((t - t_0)A)Y_0 = A\varphi(t)$. Si ψ est une solution vérifiant $\psi(t_0) = Y_0$ on dérive $h(t) = \exp(-(t - t_0)A)\psi(t)$. Il vient $h'(t) = \exp(-(t - t_0)A)(-A\psi(t) + \psi'(t)) = 0$ donc $\forall t$ $h(t) = h(t_0) = Y_0$ et donc $\psi(t) = \exp((t - t_0)A)Y_0$.

Remarques

1. Si φ est une solution quelconque de $Y' = AY$ et si $Y_0 = \varphi(0)$ on a $\varphi(t) = \exp(tA)Y_0$ par unicité.
2. Soient $\mathcal{B} = (u_1, \dots, u_n)$ une base de E et $\varphi_j(t) = \exp(tA)u_j$. $(\varphi_1, \dots, \varphi_n)$ est une base de solutions de $Y' = AY$.
3. Soit $A \in M_n(\mathbb{K})$. Les colonnes de la matrice $\exp(tA)$ forment une base de solutions du système $Y' = AY$. En effet, la i -ième colonne de cette matrice est $\exp(tA)e_i$ où (e_1, \dots, e_n) est la base canonique de $M_n(\mathbb{K})$.
4. Soit $A \in M_n(\mathbb{K})$ et $(\varphi_1, \dots, \varphi_n)$ une base de solutions de $Y' = AY$. Soit $W(t) = \text{Mat}(\varphi_1(t), \dots, \varphi_n(t))$ la matrice dont les colonnes sont les $\varphi_j(t)$. On a en utilisant la remarque 1. $W(t) = \exp(tA)W(0)$ donc $\exp(tA) = W(t)W(0)^{-1}$. La connaissance d'une base de solutions du système permet donc de calculer $\exp(tA)$.
5. Si A est diagonalisable, et si on choisit pour \mathcal{B} une base de vecteurs propres de A : $Ae_j = \lambda_j u_j$, on obtient une base de solutions de $Y' = AY$: $\varphi_1(t) = e^{\lambda_1 t} u_1, \dots, \varphi_n(t) = e^{\lambda_n t} u_n$.

16.2.2 Structure de l'ensemble des solutions

Commençons par étudier le cas particulier où $A = \lambda Id + \nu$ avec $\nu \in L(E)$ nilpotent. On a $\exp(tA) = e^{\lambda t} \exp(t\nu) = e^{\lambda t} \left(Id + t\nu + \dots + \frac{t^{\beta-1}}{(\beta-1)!} \nu^{\beta-1} \right)$ où β est l'indice de nilpotence de ν . En posant $Y_j = \nu^j(Y_0)$, on obtient la solution qui prend la valeur Y_0 pour $t = 0$:

$$Y(t) = e^{\lambda t} \left(Y_0 + tY_1 + \dots + \frac{t^{\beta-1}}{(\beta-1)!} Y_{\beta-1} \right)$$

DEFINITION 16.2.1

Nous dirons qu'une fonction $F : \mathbb{R} \rightarrow E$ est polynômiale si il existe un entier m et des vecteurs $V_0, \dots, V_m \in E$ tels que $\forall t$ $F(t) = V_0 + tV_1 + \dots + t^m V_m$. Une telle écriture est alors unique. Il revient au même de dire que les composantes F_j de F dans une base (e_1, \dots, e_n) de E sont des fonctions polynômes usuelles de la variable t . Si F n'est pas nulle, on définira son degré comme le plus grand des entiers k tels que $V_k \neq 0$. C'est aussi le maximum des degrés des composantes F_j de F .

On voit donc que si $A = \lambda I_n + \nu$ avec ν nilpotent d'indice β , toute solution de $Y' = AY$ est de la forme $Y(t) = e^{\lambda t} P(t)$ où P est une fonction polynômiale de degré au plus $\beta - 1$.

Attention! Il est faux que toute fonction de cette forme soit solution du système !

Revenons au cas général et plaçons nous toujours dans le cas où le polynôme caractéristique de u est scindé sur \mathbb{K} , ce qui est toujours le cas si $\mathbb{K} = \mathbb{C}$. Ecrivons ce polynôme sous la forme $\chi_u = \prod_{j=1}^{j=r} (\lambda_j - X)^{m_j}$. Soit $F_j = \ker(u - \lambda_j Id)^{m_j}$ l'espace caractéristique associé à la valeur propre λ_j . On sait qu'on a $E = \bigoplus_{j=1}^{j=r} F_j$.

Notons A_j l'endomorphisme induit par A sur F_j . Il est de la forme $A_j = \lambda_j Id_{A_j} + \nu_j$ où ν_j est nilpotent.

Soit $Y_0 \in E$. On le décompose en $Y_0 = V_1 + \dots + V_r$ avec $V_j \in F_j$. On a alors

$$\exp(tA)Y_0 = \sum_{j=1}^{j=r} \exp(tA)V_j = \sum_{j=1}^{j=r} \exp(tA_j)V_j = \sum_{j=1}^{j=r} e^{\lambda_j t} \left(V_j + t\nu_j(V_j) + \dots + \frac{t^{\beta_j-1}}{(\beta_j-1)!} \nu_j^{\beta_j-1}(V_j) \right)$$

Conclusion : toute solution φ de $Y' = AY$ peut s'écrire comme somme $\varphi = \sum_{j=1}^{j=r} \varphi_j$ de solutions, chacune des φ_j étant à valeur dans l'espace caractéristique F_j . Chacune des $\varphi_j(t)$ est le produit de $e^{\lambda_j t}$ par une fonction polynomiale de degré au plus $\dim F_j - 1 = m_j - 1$.

16.2.3 Système avec second membre

On se donne $A \in L(E)$ fixée, I un intervalle non trivial de \mathbb{R} , $B : I \rightarrow E$ une application continue. On considère le système

$$(S) \quad Y' = AY + B(t)$$

Si on connaît une solution $\psi_0 : I \rightarrow E$ de ce système, l'ensemble des solutions est $\{\psi = \psi_0 + \varphi ; \varphi \text{ solution de } Y' = AY\}$.

Méthode de variation des constantes

On va détailler la méthode en se plaçant dans le cas matriciel. Soit donc $A \in M_n(\mathbb{K})$ et $B : I \rightarrow M_{n,1}(K) = \mathbb{K}^n$ continue. Soit $(\varphi_1, \dots, \varphi_n)$ une base de solutions de $Y' = AY$ et $W(t)$ la matrice dont les colonnes sont les $\varphi_j : W(t) = \text{Mat}(\varphi_1(t), \dots, \varphi_n(t))$. Pour tout t , $W(t)$ est inversible et on a $W'(t) = AW(t)$. On cherche une solution particulière sous la forme $\Phi(t) = W(t)Z(t)$ avec $Z : I \rightarrow \mathbb{K}^n$ de classe C^1 . Il vient $\Phi'(t) = W'(t)Z(t) + W(t)Z'(t) = AW(t)Z(t) + W(t)Z'(t)$ donc $\Phi'(t) - A\Phi(t) = W(t)Z'(t)$ de sorte que Y est solution de (S) ssi Z vérifie $\forall t \in I \quad Z'(t) = W(t)^{-1}B(t)$. Si $t_0 \in I$, on peut choisir $Z(t) = \int_{t_0}^t W(s)^{-1}B(s)ds$. On obtient une solution particulière

$$Y_0(t) = W(t) \int_{t_0}^t W(s)^{-1}B(s)ds = \int_{t_0}^t W(t)W(s)^{-1}B(s)ds$$

Si on prend comme base de solutions $\varphi_j(t) = \exp(tA)e_j$ où (e_1, \dots, e_n) est la base canonique de \mathbb{K}^n , on a $W(t) = \exp(tA)$ et

$$Y_0(t) = \int_{t_0}^t \exp((t-s)A)B(s)ds$$

de sorte que la solution générale de (S) est

$$Y(t) = \exp(tA)X_0 + \int_{t_0}^t \exp((t-s)A)B(s)ds$$

16.3 Equations différentielles linéaires à coefficients constants

16.3.1 Equation homogène

Soit $n \geq 1$ un entier naturel, a_0, \dots, a_{n-1} n nombres complexes fixés ; on posera $a_n = 1$. On introduit le polynôme

$$P(X) = \sum_{k=0}^{k=n} a_k X^k \text{ qui se factorise dans } \mathbb{C}[X] \text{ en } P(X) = \prod_{j=1}^{j=r} (X - \lambda_j)^{m_j}, \text{ les } \lambda_j \text{ étant deux à deux distincts.}$$

On considère l'équation différentielle

(\mathcal{E}_P) $y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_1y' + a_0y = 0$
 L'équation $P(x) = 0$ s'appelle équation caractéristique de (\mathcal{E}_P). Ces racines sont les λ_j , de multiplicité m_j .

THEOREME 16.3.1

La solution générale de (\mathcal{E}_P) est donnée par

$$\varphi(t) = \sum_{j=1}^{j=r} e^{\lambda_j t} p_j(t)$$

où p_j est un polynôme quelconque à coefficients complexes, de degré au plus $m_j - 1$.

preuve

• Le système différentiel associé est $Y' = AY$ avec

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 1 \\ -a_0 & \dots & \dots & -a_{n-2} & -a_{n-1} \end{pmatrix}$$

et une fonction φ est solution de \mathcal{E}_P ssi $Y = \begin{pmatrix} \varphi \\ \varphi' \\ \vdots \\ \varphi^{(n-1)} \end{pmatrix}$ est solution de $Y' = AY$. Le polynôme caractéristique de A est

égal au signe près au polynôme P .

• Il résulte de l'étude faite au paragraphe 16.2.2 que toute solution de ce système est de la forme $Y = \sum_{j=1}^{j=r} Y_j$ où $Y_j(t) = e^{\lambda_j t} \vec{P}_j(t)$ les \vec{P}_j étant des fonction $I \rightarrow F_j \subset E$ polynomiales. En prenant la première coordonnée de Y , on voit que toute solution de (\mathcal{E}_P) est de la forme indiquée dans l'énoncé.

- Soit \mathcal{S} l'espace des solutions de \mathcal{E}_P . On sait que c'est un \mathbb{C} -ev de dimension n .
- Soit $H = \mathbb{C}_{m_1-1}[X] \times \mathbb{C}_{m_2-1}[X] \times \dots \times \mathbb{C}_{m_r-1}[X]$. H est un \mathbb{C} -ev de dimension $m_1 + \dots + m_r = n$. Considérons enfin l'application $\Phi : H \rightarrow C^\infty(\mathbb{R}, \mathbb{C})$ qui à (p_1, \dots, p_r) associe la fonction $t \rightarrow e^{\lambda_1 t} p_1(t) + \dots + e^{\lambda_r t} p_r(t)$. Cette application est linéaire. On vient de voir que $\mathcal{S} \subset \Phi(H)$. On a $\dim \mathcal{S} = n$ et $\dim H = n$. Le théorème du rang donne $\dim(\Phi(H)) = n - \dim \ker \Phi$. On en déduit $\ker \Phi = \{0\}$ et $\mathcal{S} = \Phi(H)$ ce qui est précisément ce que l'on voulait démontrer.

NB : Une autre preuve de ce résultat qui n'utilise pas les systèmes différentiels linéaires ni l'exponentielle de matrices figure dans le cours d'algèbre (chapitre réduction des endomorphismes).

Le théorème de Cauchy s'applique bien entendu. Pour tout $t_0 \in \mathbb{R}$ et $(a_0, \dots, a_{n-1}) \in \mathbb{C}^n$, il existe une unique solution φ telle que $\varphi^{(j)}(t_0) = a_j$ pour $0 \leq j \leq n - 1$.

16.3.2 Equation avec second membre exponentielle-polynôme

Soient $\alpha \in \mathbb{C}$ et $R \in \mathbb{C}[X]$. On considère maintenant l'équation

(E) $y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_1y' + a_0y = e^{\alpha t} R(t)$

Si α est racine de l'équation caractéristique $\sum_{k=0}^{k=n} a_k x^k = 0$ on notera m sa multiplicité. Sinon, on pose $m = 0$.

THEOREME 16.3.2

Il existe une solution particulière de (E) de la forme $t \rightarrow e^{\alpha t} Q(t)$ où Q est un polynôme et $\deg Q = \deg R + m$. De plus si $m = 0$, c'est à dire si α n'est pas racine de l'équation caractéristique, Q est unique.

preuve

La preuve repose sur deux lemmes

Lemme 1

Soient $\gamma_0, \dots, \gamma_n \in \mathbb{C}$ et $R \in \mathbb{C}[X]$ de degré p . On suppose en outre $\gamma_0 \neq 0$. L'équation $\gamma_n y^{(n)} + \dots + \gamma_1 y' + \gamma_0 y = R$ admet une unique solution polynôme Q et $\deg Q = \deg R$.

preuve

Si Q est un polynôme de degré q , $\gamma_n Q^{(n)} + \dots + \gamma_0 Q$ est un polynôme de degré q puisque $\gamma_0 \neq 0$. Si l'équation considérée admet une solution polynôme celle-ci est donc nécessairement de degré égal à p . Considérons alors l'application $F : \mathbb{C}_p[X] \rightarrow \mathbb{C}_p[X]$ qui à un polynôme Q associe le polynôme $F(Q) = \gamma_n Q^{(n)} + \dots + \gamma_0 Q$. Sa matrice dans la base canonique de $\mathbb{C}_n[X]$ est triangulaire, les éléments de la diagonale étant tous égaux à γ_0 . C'est donc un isomorphisme de $\mathbb{C}_p[X]$ sur lui-même. Le polynôme R admet un unique antécédent par F , ce qui prouve le lemme 1.

Lemme 2

Soit $Q \in \mathbb{C}[X]$ de degré q et $f(t) = e^{\alpha t} Q(t)$. On a, en posant $P(X) = X^n + a_{n-1}X^{n-1} + \dots + a_0$.

$$f^{(n)}(t) + a_1 f^{(n-1)}(t) + \dots + a_1 f'(t) + a_0 f(t) = e^{\alpha t} \sum_{k=0}^{k=n} \frac{P^{(k)}(\alpha)}{k!} Q^{(k)}(t)$$

preuve

La formule de Leibnitz donne $f^{(k)}(t) = \sum_{j=0}^{j=k} C_k^j \alpha^{k-j} e^{\alpha t} Q^{(j)}(t)$. On en déduit (en posant $a_n = 1$)

$$\begin{aligned} \sum_{k=0}^{k=n} a_k f^{(k)}(t) &= \sum_{k=0}^{k=n} \left(\sum_{j=0}^{j=k} a_k C_k^j \alpha^{k-j} Q^{(j)}(t) e^{\alpha t} \right) \\ &= e^{\alpha t} \sum_{j=0}^{j=n} \left(\sum_{k=j}^{k=n} k(k-1) \dots (k-j+1) a_k \alpha^{k-j} \right) \frac{Q^{(j)}(t)}{j!} \\ &= e^{\alpha t} \left(\sum_{j=0}^{j=n} \frac{P^{(j)}(\alpha)}{j!} Q^{(j)}(t) \right) \end{aligned}$$

Montrons pour finir que ces deux lemmes impliquent le théorème. On cherche une solution de l'équation de la forme $f(t) = e^{\alpha t} Q(t)$. Le lemme 2 montre que f est solution de (E) ssi Q vérifie $\sum_{j=0}^{j=n} \frac{P^{(j)}(\alpha)}{j!} Q^{(j)}(t) = R(t)$. Compte tenu de la définition de m , cette équation s'écrit $\sum_{j=m}^{j=n} \frac{P^{(j)}(\alpha)}{j!} Q^{(j)}(t) = R(t)$ et le coefficient $\frac{P^{(m)}(\alpha)}{m!}$ est non nul. Il résulte du lemme 1 qu'il existe un unique polynôme S de degré égal au degré de R tel que $\sum_{j=m}^{j=n} \frac{P^{(j)}(\alpha)}{j!} Q^{(j)}(t) = R(t) \Leftrightarrow Q^{(m)} = S$. Par conséquent, si $m = 0$, $Q = S$ est unique et si $m > 0$ toute solution vérifie $\deg Q^{(m)} = \deg R$ donc $\deg Q = \deg R + m$.

16.4 Systèmes linéaires d'ordre 2 à coefficients réels

On se donne $A = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \in M_2(\mathbb{R})$ et on considère le système linéaire $Z' = AZ$ soit en posant $Z = \begin{pmatrix} x \\ y \end{pmatrix}$

$$(S) \quad \begin{cases} x' &= ax + cy \\ y' &= bx + dy \end{cases}$$

Une solution φ de ce système est une application $\mathbb{R} \rightarrow \mathbb{R}^2$, c'est à dire une courbe paramétrée. On appelle trajectoire le support de cette courbe paramétrée. On se propose d'étudier dans ce paragraphe l'allure des trajectoires des solutions de (S). La solution nulle a une trajectoire réduite à un point, l'origine. Si la matrice A est nulle, toutes les solutions sont constantes et tout point de \mathbb{R}^2 est une trajectoire. Dans la suite on supposera $A \neq 0$.

Soient (\vec{i}, \vec{j}) la base canonique de \mathbb{R}^2 , (\vec{I}, \vec{J}) une autre base, P la matrice de passage. Soit Z une solution du système (S). Posons $U = \begin{pmatrix} X \\ Y \end{pmatrix} = P^{-1}Z$. U vérifie de $U' = P^{-1}Z' = P^{-1}AZ = (P^{-1}AP)U$. Pour déterminer l'allure des trajectoires de (S), on peut donc se placer dans une autre base, mieux adaptée. On étudiera donc des systèmes plus simples $Y' = BY$ en choisissant une matrice B semblable à A .

Le polynôme caractéristique de A est $\chi_A(T) = T^2 - \text{tr}(A)T + \det A$. Son discriminant vaut $\Delta = (\text{tr}(A))^2 - 4 \det A$.

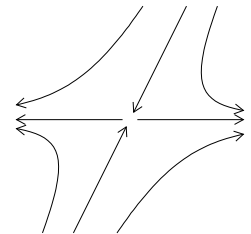
1. Cas $\det A = 0$.

- (a) $\text{tr}(A) = 0$. A est semblable à $B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Le système associé s'écrit $X' = Y$, $Y' = 0$ donc les solutions sont $X = y_0 t + x_0$, $Y = y_0$. Les trajectoires sont les droites parallèles à l'axe des X et distinctes de cet axe et les points de l'axe des X .
- (b) $\text{tr}(A) \neq 0$. Soit $\lambda = \text{tr}(A)$. A est semblable à $B = \begin{pmatrix} \lambda & 0 \\ 0 & 0 \end{pmatrix}$. Le système associé s'écrit $X' = \lambda X$, $Y' = 0$ les solutions sont $X(t) = x_0 e^{\lambda t}$, $Y(t) = y_0$. Les trajectoires sont les points de l'axe des Y (cas $x_0 = 0$) et pour chaque valeur de y_0 les deux demi-droites contenues dans la droite $Y = y_0$ d'origine le point $(0, y_0)$.

2. Cas $\det A \neq 0$.

- (a) $\det(A) < 0$

La matrice A admet deux valeurs propres réelles, l'une strictement positive, l'autre strictement négative. A est semblable à $B = \begin{pmatrix} \lambda & 0 \\ 0 & -\mu \end{pmatrix}$ où $\lambda > 0, \mu > 0$. Les solutions dans la base correspondante sont données par $X(t) = x_0 e^{\lambda t}$, $Y(t) = y_0 e^{-\mu t}$. On obtient, outre l'origine, quatre trajectoires particulières, les demi-axes de coordonnées. Les deux demi-axes des X sont parcourus de l'origine vers l'infini, les deux demi-axes des Y de l'infini à l'origine. Les autres trajectoires ont pour équation cartésienne



$$\left(\frac{X}{x_0}\right)^\mu \left(\frac{Y}{y_0}\right)^\lambda = 1 \quad \text{ou} \quad Y = y_0 \left(\frac{x_0}{X}\right)^{\mu/\lambda}$$

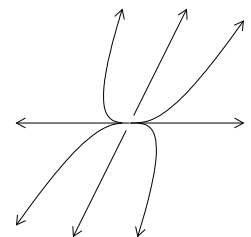
Si $\lambda = \mu$, c'est à dire $\text{tr}(A) = 0$, on obtient des branches d'hyperboles. Chaque trajectoire est contenue dans l'un des quatre quarts de plan limité par les axes de coordonnées.

- (b) $\det(A) > 0$. Ce cas se subdivise en trois

- i. $\Delta > 0$ i.e. $\text{tr}(A)^2 > 4 \det A$

A admet deux valeurs propres réelles distinctes de même signe λ et μ et A est semblable à $B = \begin{pmatrix} \lambda & 0 \\ 0 & \mu \end{pmatrix}$.

On supposera $|\mu| > |\lambda|$. Dans une base ad-hoc, les solutions du système sont $X(t) = x_0 e^{\lambda t}$, $Y(t) = y_0 e^{\mu t}$. Comme dans le cas précédent, les quatre demi-axes sont des trajectoires, le sens de parcours dépendant du signe commun de λ et μ . Chacune des autres trajectoires, autres que le singleton $\{0\}$ est contenue dans l'un des quatre quarts de plan limités par les axes de coordonnées. L'équation cartésienne d'une telle trajectoire est



$$Y = y_0 \left(\frac{X}{x_0}\right)^{\mu/\lambda}$$

avec cette fois $\mu/\lambda > 1$. Sur la figure, on a pris $\lambda > 0$ et $\mu > 0$. Si ces nombres sont négatifs, on a les mêmes trajectoires, mais le sens de parcours est renversé.

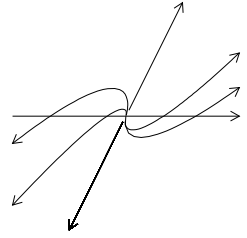
- ii. $\Delta = 0$ i.e. $\text{tr}(A)^2 = 4 \det A$.

A admet une valeur propre double λ . Il y a deux cas :

- A est diagonalisable, donc semblable à $B = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$. Dans la base des vecteurs propres, les solutions sont $X = x_0 e^{\lambda t}$, $Y = y_0 e^{\lambda t}$. Les trajectoires sont toutes les demi-droites issues de l'origine, parcourues de l'origine vers l'infini si $\lambda > 0$ dans l'autre sens sinon.
- A n'est pas diagonalisable.

Elle est semblable à $B = \begin{pmatrix} \lambda & 0 \\ 1 & \lambda \end{pmatrix}$. Le système, dans la nouvelle base a pour solutions $X(t) = x_0 e^{\lambda t}$, $Y = y_0 e^{\lambda t} + t x_0 e^{\lambda t}$. Outre l'origine, on obtient deux trajectoires particulières, les demi-axes des y parcourus de l'origine à l'infini si $\lambda > 0$, dans l'autre sens sinon. Les autres trajectoires ont pour équation cartésienne

$$Y = \frac{y_0}{x_0} X + \frac{X}{\lambda} \ln \frac{X}{x_0}$$

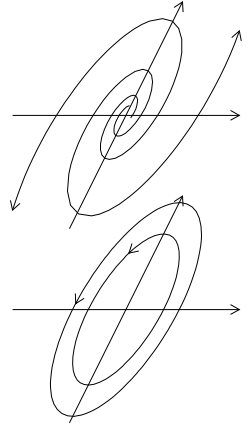


iii. $\Delta < 0$ i.e. $\text{tr}(A)^2 < 4 \det A$

A admet deux valeurs propres complexes conjuguées non réelles $\lambda = s + i\omega$, $\bar{\lambda} = s - i\omega$ avec $\omega \neq 0$. Il existe $P \in GL(2, \mathbb{R})$ telle que $P^{-1}AP = B = \begin{pmatrix} s & -\omega \\ \omega & s \end{pmatrix} = sI_2 + \omega J$ avec $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ (le vérifier !). On a $\exp(tB) = e^{ts} \exp(t\omega J) = e^{ts} \begin{pmatrix} \cos \omega t & -\sin \omega t \\ \sin \omega t & \cos \omega t \end{pmatrix}$. (c.f. l'exemple 16.1.1).

Dans la base de \mathbb{R}^2 formée par les colonnes de P , les solutions du système s'écrivent donc $X(t) = e^{ts}(x_0 \cos \omega t - y_0 \sin \omega t)$, $Y(t) = e^{ts}(x_0 \sin \omega t + y_0 \cos \omega t)$. En posant $x_0 + iy_0 = r_0 e^{i\theta_0}$, il vient $X(t) + iY(t) = r_0 e^{ts} e^{i(\omega t + \theta_0)}$.

Pour $s = 0$, i.e. $\text{tr}(A) = 0$, les trajectoires sont des ellipses de centre O , sinon on obtient des spirales. (Si on munit \mathbb{R}^2 d'un produit scalaire telle que la base formée des colonnes de P soit orthonormée, les courbes obtenues sont précisément des spirales logarithmiques, d'équation polaire $\rho = K e^{k\theta}$ avec $K = r_0 e^{-s\theta_0/\omega}$, $k = s/\omega$).



17

Equations différentielles

17.1 Systèmes d'ordre 1

17.1.1 Définitions

Soient Ω un ouvert non vide de $\mathbb{R} \times \mathbb{R}^p$ et $F : \Omega \rightarrow \mathbb{R}^p$ une application. On considère l'équation différentielle

$$(E) \quad X' = F(t, X)$$

Une solution de (E) est un couple (I, φ) où I est un intervalle non trivial de \mathbb{R} et φ une application dérivable de I dans \mathbb{R}^p vérifiant

$$\begin{cases} \forall t \in I & (t, \varphi(t)) \in \Omega \\ \forall t \in I & \varphi'(t) = F(t, \varphi(t)) \end{cases}$$

On dira aussi que φ est une I -solution pour signifier que (I, φ) est une solution.

Si (I, φ) et (J, ψ) sont deux solutions, on dit que (J, ψ) prolonge (I, φ) si $I \subset J$ et $\forall t \in I \quad \psi(t) = \varphi(t)$.

Une solution (I, φ) est dite maximale si toute solution qui la prolonge lui est égale. Autrement dit, (I, φ) est une solution maximale si c'est une solution et si il n'existe aucune solution (J, ψ) avec $I \subsetneq J$ et $\psi(t) = \varphi(t)$ pour tout $t \in I$.

Remarquons que si F est continue (resp. de classe C^k), toute solution de (E) est de classe C^1 (resp C^{k+1}).

On appelle donnée initiale pour le problème de Cauchy un point $(t_0, X_0) \in \Omega$. Le problème de Cauchy pour cette donnée initiale consiste à chercher les solutions (I, φ) avec $t_0 \in I$ et $\varphi(t_0) = X_0$.

17.1.2 Théorème de Cauchy-Lipschitz

THEOREME 17.1.1 (Unicité locale)

Soient Ω un ouvert non vide de $\mathbb{R} \times \mathbb{R}^p$, $F : \Omega \rightarrow \mathbb{R}^p$ une application de classe C^1 . Soient (I, φ) et (J, ψ) deux solutions telles que $\exists t_0 \in I \cap J \quad \varphi(t_0) = \psi(t_0)$. Alors $\forall t \in I \cap J \quad \psi(t) = \varphi(t)$.

THEOREME 17.1.2 (Existence et unicité d'une solution maximale)

Soient Ω un ouvert non vide de $\mathbb{R} \times \mathbb{R}^p$, $F : \Omega \rightarrow \mathbb{R}^p$ une application de classe C^1 et $(t_0, X_0) \in \Omega$.

1. Il existe une solution maximale (I, φ) de (E) telle que $\varphi(t_0) = X_0$.
2. Soit (J, ψ) une solution de (E) telle que $t_0 \in J$ et $\psi(t_0) = X_0$. On a $J \subset I$ et $\psi = \varphi|_J$. En particulier il n'y a qu'une solution maximale vérifiant $\varphi(t_0) = X_0$.
3. L'intervalle I de définition de la solution maximale est ouvert.

Voici une méthode d'usage très fréquent pour étudier l'intervalle de définition d'une solution maximale. Nous l'utiliserons dans les exemples qui suivent.

PROPOSITION 17.1.1

Soient $F : \Omega \rightarrow \mathbb{R}^p$ de classe C^1 et $\psi : J \rightarrow \mathbb{R}^p$ une solution de (E) définie sur un intervalle ouvert $J =]\alpha, \beta[$ avec $\beta < +\infty$. Supposons qu'existe $\lim_{x \rightarrow \beta} \varphi(x)$, soit λ et que $(\beta, \lambda) \in \Omega$. Alors ψ se prolonge en une solution sur $] \alpha, \beta]$ et par conséquent ψ n'est pas une solution maximale.

On a bien entendu un énoncé analogue pour α .

preuve

Notons d'abord que pour $x \in J$, $(x, \psi(x)) \in \Omega$ donc si la limite de ψ en β existe, on a toujours $(\beta, \lambda) \in \bar{\Omega}$. Ici, on suppose que $(\beta, \lambda) \in \Omega$. Notons Ψ le prolongement de ψ à $J' =]\alpha, \beta]$ défini par $\Psi(\beta) = \lambda$. Ψ est continue sur J' .

Pour $x \in J$, on a $\psi'(x) = F(x, \psi(x))$, donc F étant continue en (β, λ) (c'est ici qu'on utilise le fait que ce point est dans Ω), on a $\lim_{x \rightarrow \beta} \psi'(x) = F(\beta, \lambda)$. On sait que l'existence de cette limite implique que Ψ est C^1 sur J' et que $\Psi'(\beta) = \lim_{x \rightarrow \beta} \psi'(x) = F(\beta, \lambda) = F(\beta, \Psi(\beta))$. Par conséquent Ψ est bien une solution de (E) sur J' qui prolonge ψ .

Notons que l'hypothèse $(\beta, \lambda) \in \Omega$ est automatiquement vérifiée si $\Omega = \mathbb{R} \times \mathbb{R}^p$.

17.2 Equation différentielle scalaire du premier ordre

C'est le cas $p = 1$, Ω est un ouvert de \mathbb{R}^2 et $F : \Omega \rightarrow \mathbb{R}$. Si (I, φ) est une solution, le graphe $C_\varphi = \{(x, \varphi(x)) ; x \in I\}$ est contenu dans Ω et s'appelle une courbe intégrale de l'équation différentielle (E). Si F est de classe C^1 , il résulte du théorème de Cauchy Lipschitz que deux courbes intégrales distinctes associées à des solutions maximales ne se coupent pas.

Nous allons sur deux exemples montrer comment utiliser le théorème de Cauchy-Lipschitz pour résoudre une équation différentielle ou pour obtenir l'allure des courbes intégrales.

17.2.1 Etude de l'équation $y' = x \sin y$

1. On a $\Omega = \mathbb{R}^2$ et $F(x, y) = x \sin y$. La fonction F est de classe C^1 sur Ω donc le théorème de Cauchy-Lipschitz s'applique pour toute condition initiale (x_0, y_0) .

2. Pour tout $k \in \mathbb{Z}$, la fonction constante $c_k : \mathbb{R} \rightarrow \mathbb{R}$, $c_k(x) = k\pi$ est solution de (E), nécessairement maximale.

3. Montrons que les solutions maximales sont définies sur \mathbb{R} tout entier.

Soit (I, φ) une solution maximale. On sait que I est un intervalle ouvert, soit $I =]\alpha, \beta[$. Montrons par l'absurde que $\beta = \infty$.

Supposons $\beta < \infty$ et soit $x_0 \in I$. Pour $x \in [x_0, \beta[$ on a $|\varphi'(x)| \leq |x| \leq C = \max(|x_0|, |\beta|)$. La fonction φ est de classe C^1 sur $[x_0, \beta[$ et a une dérivée bornée. Elle se prolonge donc par continuité en β : $\varphi(x)$ a une limite finie λ quand x tend vers β . Comme $(\beta, \lambda) \in \Omega = \mathbb{R}^2$, φ se prolonge (proposition 17.1.1) en une solution sur $] \alpha, \beta]$ ce qui contredit le caractère maximal de φ .

On démontre de la même manière que $\alpha = -\infty$.

Il en résulte que toute solution maximale est définie en 0, donc que pour déterminer toutes les solutions maximales, il suffit de déterminer celles qui vérifient une condition du type $\varphi(0) = y_0$.

4. Toute solution maximale est paire.

Soit $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ une solution maximale et $\psi(x) = \varphi(-x)$. On a $\psi'(x) = -\varphi'(-x) = -[(-x) \sin(\varphi(-x))] = x \sin(\varphi(x))$ donc ψ est aussi une solution maximale (car définie sur \mathbb{R}). Mais $\psi(0) = \varphi(0)$ donc par le théorème d'unicité $\forall x \in \mathbb{R} \psi(x) = \varphi(x)$.

5. Si φ est une solution maximale, il en est de même de la fonction $x \rightarrow \varphi(x) + 2k\pi$ pour $k \in \mathbb{Z}$ et de la fonction $-\varphi$.

Notons alors φ_{y_0} la solution maximale qui prend la valeur y_0 en 0. On déduit de ce qui précède que $\varphi_{y_0+2k\pi} = \varphi_{y_0} + 2k\pi$ et que $\varphi_{-y_0} = -\varphi_{y_0}$. Par conséquent, pour étudier les courbes intégrales autres que les droites d'équations $y = k\pi$ on peut supposer $0 < y_0 < \pi$. Les autres s'en déduisent par les translations de vecteur $2k\pi \vec{j}$ éventuellement suivie d'une symétrie par rapport à l'axe des x .

6. Allure des courbes intégrales.

Soit $\varphi = \varphi_{y_0}$ avec $0 < y_0 < \pi$. Les courbes intégrales de deux solutions maximales distinctes ne se rencontrent pas, donc φ ne peut pas prendre les valeurs 0 et π . Il résulte du théorème des valeurs intermédiaires que $\forall x \in \mathbb{R} \ 0 < \varphi(x) <$

π . Donc $\sin[\varphi(x)] > 0$ pour tout $x \in \mathbb{R}$. Par conséquent, $\forall x \in \mathbb{R}_+, \varphi'(x) > 0$. φ induit une bijection strictement croissante de $[0, +\infty[$ sur un intervalle $[y_0, \lambda[$ avec $\lambda = \lim_{x \rightarrow \infty} \varphi(x) \in [y_0, \pi]$. Mais on a $x' = \frac{\varphi'(x)}{\sin[\varphi(x)]}$ donc

$$\int_{x_0}^x t dt = \int_{x_0}^x \frac{\varphi'(t)}{\sin \varphi(t)} dt = \int_{y_0}^{\varphi(x)} \frac{du}{\sin u}$$

Lorsque x tend $+\infty$, le premier membre tend vers l'infini, donc l'intégrale $\int_{y_0}^{\lambda} \frac{du}{\sin u}$ diverge. Par conséquent $\lambda = \pi$. Donc sur $[0, +\infty[$, φ croît de y_0 à π . Par parité, sur $] -\infty, 0]$, φ décroît de π à y_0 .

7. Résolution complète.

Dans cet exemple, on peut calculer complètement la solution $\varphi = \varphi_{y_0}$. Supposons $0 < y_0 < \pi$. On a vu que $\sin(\varphi(x))$ ne s'annulait pas. On a donc

$$\frac{\varphi'(x)}{\sin \varphi(x)} = x$$

donc en intégrant de 0 à x

$$\forall x \in \mathbb{R} \quad \frac{x^2}{2} = \int_0^x \frac{\varphi'(t)}{\sin \varphi(t)} dt = \ln \left(\tan \frac{\varphi(x)}{2} \right) - \ln \left(\tan \frac{y_0}{2} \right)$$

On en déduit $\tan \frac{\varphi(x)}{2} = \tan \frac{y_0}{2} e^{x^2/2}$ et comme $0 < \frac{\varphi(x)}{2} < \frac{\pi}{2}$

$$\varphi_{y_0}(x) = 2 \arctan \left(\tan \frac{y_0}{2} e^{x^2/2} \right)$$

Les autres solutions sont les fonctions $\varphi_{y_0} + 2k\pi$, $-\varphi_{y_0} + 2k\pi$ et les fonctions constantes $c_k = k\pi$.

L'équation étudiée est un exemple d'équation à variables séparables. Une équation à variables séparables est une équation qui peut s'écrire, sous réserve d'une discussion plus ou moins compliquée, sous la forme $v(y)y' = u(x)$ où u et v sont des fonctions d'une variable.

17.2.2 Etude de l'équation $y' = \exp(xy) - 1$

Considérons l'équation différentielle

$$(E) \quad y' = e^{xy} - 1$$

1. $\Omega = \mathbb{R}^2$, $F : \Omega \rightarrow \mathbb{R}$ définie par $F(x, y) = e^{xy} - 1$ est de classe C^1 sur \mathbb{R}^2 . Le théorème de Cauchy-Lipschitz s'applique en chaque point.

Par ailleurs, il est clair que la fonction nulle est solution sur \mathbb{R} . Il en résulte que toute autre solution maximale est soit strictement positive pour tout x , soit strictement négative pour tout x .

Soit $\varphi : I \rightarrow \mathbb{R}$ une solution maximale. Soit $J = -I = \{-x ; x \in I\}$ et $\psi : J \rightarrow \mathbb{R}$ définie par $\psi(x) = -\varphi(-x)$. On a $\psi'(x) = \varphi'(-x) = e^{(-x)\varphi(-x)} - 1 = e^{x\psi(x)} - 1$ donc ψ est aussi une solution de (E). On vérifie facilement qu'elle est maximale. La courbe intégrable C_ψ est symétrique de C_φ par rapport à l'origine.

On peut donc se limiter à l'étude des solutions maximales à valeurs dans \mathbb{R}_+^* .

2. Soit $\varphi : I \rightarrow \mathbb{R}$ une solution maximale de (E) à valeurs strictement positives. $x\varphi(x)$ est du signe de x , donc $e^{x\varphi(x)} - 1$ aussi. Par conséquent φ est croissante sur $I \cap [0, +\infty[$ et décroissante sur $I \cap]-\infty, 0]$.

D'autre part, en dérivant $\varphi'(x) = e^{x\varphi(x)} - 1$ on obtient $\varphi''(x) = e^{x\varphi(x)}(x\varphi'(x) + \varphi(x))$. Mais $\varphi(x) > 0$ et $x\varphi'(x) \geq 0$ d'après ce qui précède.

Donc, toute solution strictement positive de (E) est strictement convexe sur I .

3. Etude de l'intervalle de définition d'une solution maximale.

Soit $I =]\alpha, \beta[$ l'intervalle de définition de la solution maximale strictement positive φ . On a $-\infty \leq \alpha < \beta \leq +\infty$.

- (a) Montrons d'abord que $\alpha < 0$. Supposons le contraire : $\alpha \geq 0$. Alors φ est croissante sur I . Etant positive, elle aurait une limite finie en α . En utilisant le raisonnement fait dans la preuve de la proposition 17.1.1 on voit qu'on prolonge φ en une solution de l'équation différentielle sur $[\alpha, \beta[$ ce qui contredit le fait que la solution est maximale. Donc $\alpha < 0$.
- (b) Montrons que $\beta > 0$. Supposons le contraire : on a donc $-\infty < \beta \leq 0$. $I \subset \mathbb{R}_-$ donc φ est décroissante sur I . Etant positive, elle admet une limite finie en β . Comme précédemment, on en déduit que φ se prolonge en une solution sur $] -\alpha, \beta]$. Contradiction.
- (c) Montrons que $\alpha = -\infty$. Soit $J =]\alpha, 0[$. Sur cet intervalle non vide, $x\varphi(x) \leq 0$ donc $-1 < e^{x\varphi(x)} - 1 \leq 0$ soit $-1 < \varphi'(x) \leq 0$. L'inégalité des accroissements finis donne $|\varphi(x) - \varphi(0)| \leq |x|$ donc $|\varphi(x)| \leq |\varphi(0)| + |x|$. Supposons $\alpha > -\infty$. Alors, pour $x \in J$ on a $|\varphi(x)| \leq |\varphi(0)| + |x|$. φ est décroissante sur J et majorée. Elle admet donc une limite finie en α . On en déduit comme précédemment que l'on peut prolonger φ en une solution définie sur $[\alpha, \beta[$ ce qui contredit le fait que φ est une solution maximale.
Conclusion : Nous avons montré que l'intervalle de définition d'une solution maximale strictement positive est de la forme $] -\infty, \beta[$ avec $0 < \beta \leq +\infty$.
- (d) Nous allons montrer que $\beta < +\infty$ en comparant φ avec la solution d'une autre équation différentielle. φ est définie en 0. Posons $y_0 = \varphi(0) > 0$.

Pour $t > 0$ on a $e^t - 1 > \frac{t^2}{2}$. Sur $]0, \beta[$ la fonction φ est positive donc $\varphi'(x) = e^{x\varphi(x)} - 1 \geq \frac{x^2\varphi^2(x)}{2}$
Considérons l'équation différentielle

$$(E') \quad y' = \frac{x^2 y^2}{2}$$

Le théorème de Cauchy Lipschitz s'applique en tout point de \mathbb{R}^2 à (E'). D'autre part la fonction nulle est solution, donc la solution maximale de (E') qui prend la valeur $y_0 = \varphi(0)$ en 0 ne s'annule pas. Elle vérifie donc

$$\frac{\psi'(x)}{\psi^2(x)} = \frac{x^2}{2}$$

donc en intégrant de 0 à x

$$\frac{1}{y_0} - \frac{1}{\psi(x)} = \frac{x^3}{6} \quad \text{d'où} \quad \psi(x) = \frac{6y_0}{6 - y_0 x^3}$$

La solution maximale de (E') qui prend la valeur y_0 en 0 est donc $\psi : I' \rightarrow \mathbb{R}$ où $I' =] -\infty, t_0 [$ avec $t_0 = \sqrt[3]{\frac{6}{y_0}}$.

Proposition

Pour $x \in]0, \beta[\cap I'$ on a $\varphi(x) > \psi(x)$.

preuve

On a d'abord $\varphi(0) = \psi(0) = y_0$, $\varphi'(0) = \psi'(0) = 0$ et $\varphi''(0) = y_0 > 0 = \psi''(0)$. Au voisinage de 0, on a donc $\varphi(x) - \psi(x) \sim \frac{y_0}{2}x^2$. Par conséquent, il existe $\eta > 0$ tel que $0 < x < \eta \Rightarrow \varphi(x) > \psi(x)$.

Supposons l'affirmation de la proposition fausse. D'après le théorème des valeurs intermédiaires, l'ensemble $F = \{x \in]0, \beta[\cap I' \mid \varphi(x) = \psi(x)\}$ est non vide. Il est minoré par η . Il admet donc une borne inférieure $\zeta \geq \eta$. D'autre part, il est fermé dans $]0, \beta[\cap I'$ donc $\zeta \in F$. Par conséquent, on a $\varphi(\zeta) = \psi(\zeta)$. Sur $]0, \zeta[$, $\varphi - \psi$ ne s'annule pas. Comme elle est positive sur $]0, \eta[$, on a $\forall x \in]0, \zeta[$, $\varphi(x) - \psi(x) > 0$.

On a $\varphi(\zeta) = \psi(\zeta)$ et $\varphi(0) = \psi(0)$; d'après le théorème de Rolle, il existe un $c \in]0, \zeta[$ tel que $\varphi'(c) - \psi'(c) = 0$.
Mais

$$\varphi'(c) - \psi'(c) = e^{c\varphi(c)} - 1 - \frac{c^2\psi^2(c)}{2} > \frac{c^2\varphi^2(c)}{2} - \frac{c^2\psi^2(c)}{2} > 0$$

car $\varphi(c) > \psi(c) > 0$. Contradiction.

On en déduit que $\beta \leq t_0$. En effet, supposons le contraire. Alors φ est continue en t_0 et on a $\psi(x) < \varphi(x)$ pour $0 < x < t_0$ ce qui contredit le fait que $\lim_{x \rightarrow t_0} \psi(x) = \infty$.

Finalement on a montré que l'intervalle de définition d'une solution maximale strictement positive φ de (E) était de la forme $] -\infty, \beta[$ avec $0 < \beta < \infty$.

Comme φ est croissante sur $]0, \beta[$, φ admet une limite $\lambda \in [y_0, +\infty[$ en β . Si cette limite était finie, φ se prolongerait en par continuité en β , puis par un argument déjà détaillé en une solution de (E) sur $] -\infty, \beta]$ ce qui est impossible puisque φ est une solution maximale. Donc $\lim_{x \rightarrow \beta} \varphi(x) = +\infty$.

4. Pour terminer, nous allons montrer que la courbe intégrale C_φ admet une asymptote oblique quand x tend vers $-\infty$.
D'abord, φ est décroissante sur \mathbb{R}_- donc pour $x \leq 0$ on a $\varphi(x) \geq y_0 = \varphi(0) > 0$. Donc $\lim_{x \rightarrow -\infty} x\varphi(x) = -\infty$ et

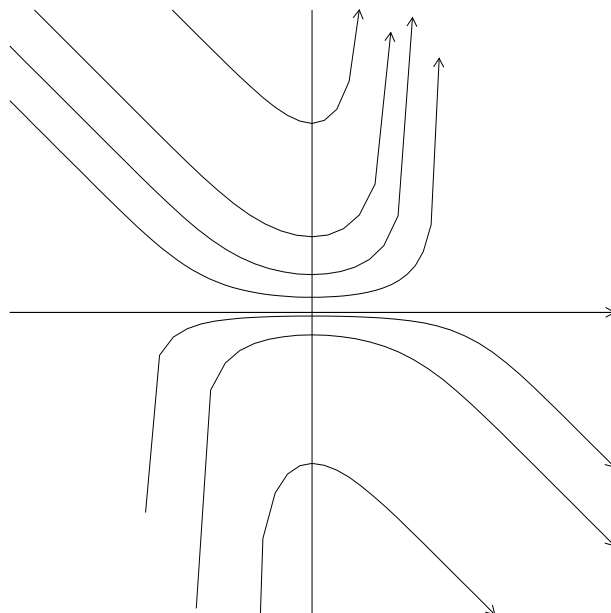
$\varphi'(x) = e^{x\varphi(x)} - 1$ tend vers -1 .

Ensuite, on a $\varphi(x) + x = y_0 + \int_0^x (\varphi'(t) + 1) dt = y_0 + \int_0^x e^{t\varphi(t)} dt$ mais pour $t \leq 0$, $t\varphi(t) \leq ty_0$ donc $0 \leq e^{t\varphi(t)} \leq e^{ty_0}$.

On en déduit que l'intégrale $\int_{-\infty}^0 e^{t\varphi(t)} dt$ converge. Par conséquent $\varphi(x) + x$ a une limite finie quand x tend vers $-\infty$

ce qui prouve que C_φ admet une asymptote oblique d'équation $Y = -X + C$ où $C = y_0 - \int_{-\infty}^0 e^{t\varphi(t)} dt$.

La figure ci dessous donne l'allure des courbes intégrales de l'équation (E').



17.3 Equations scalaires du second ordre

17.3.1 Généralités

Soit Ω un ouvert de \mathbb{R}^3 et $f : \Omega \rightarrow \mathbb{R}$ une application. On considère l'équation différentielle du second ordre

$$(E) \quad y'' = f(x, y, y')$$

DEFINITION 17.3.1

Soit I un intervalle non trivial de \mathbb{R} . Une I -solution de (E) est une application $\varphi : I \rightarrow \mathbb{R}$ deux fois dérivable telle que

$$\begin{cases} \forall x \in I & (x, \varphi(x), \varphi'(x)) \in \Omega \\ \forall x \in I & \varphi''(x) = f(x, \varphi(x), \varphi'(x)) \end{cases}$$

Si f est continue (resp. de classe C^k), toute solution est de classe C^2 (resp. C^{k+2}).

Si (I, φ) et (J, ψ) sont deux solutions, on dit que (J, ψ) prolonge (I, φ) si $I \subset J$ et $\forall t \in I \quad \psi(t) = \varphi(t)$.

Une solution (I, φ) est dite maximale si toute solution qui la prolonge lui est égale.

On appelle donnée initiale pour le problème de Cauchy un point $(x_0, y_0, y'_0) \in \Omega$. Le problème de Cauchy pour cette donnée initiale consiste à chercher les solutions (I, φ) avec $x_0 \in I$, $\varphi(x_0) = y_0$ et $\varphi'(x_0) = y'_0$.

Soit $f : \Omega \rightarrow \mathbb{R}$. Définissons une application $F : \Omega \rightarrow \mathbb{R}^2$ en posant $F(x, (y, z)) = (z, f(x, y, z))$ et considérons le système différentiel d'ordre 1 associé à F . En posant $U = \begin{pmatrix} y \\ z \end{pmatrix}$ ce système s'écrit

$$(S) \quad U' = F(x, U)$$

c'est à dire

$$\begin{cases} y' &= z \\ z' &= f(x, y, z) \end{cases}$$

Il est immédiat que si U est solution (resp. une solution maximale) de ce système sur un intervalle I , la première composante y de U est solution (resp. solution maximale) de (E) et inversement que si φ est une solution (resp. une solution maximale) de (E), la fonction $U = \begin{pmatrix} \varphi \\ \varphi' \end{pmatrix}$ est solution (resp. une solution maximale) de (S).

Le problème de Cauchy pour (E) associé à la donnée initiale (x_0, y_0, y'_0) correspond au problème de Cauchy pour (S) avec la donnée initiale $(x_0, (y_0, y'_0))$. Comme $F \in C^1 \Leftrightarrow f \in C^1$ on en déduit

THEOREME 17.3.1 (Existence et unicité d'une solution maximale)

Soient Ω un ouvert non vide de \mathbb{R}^3 , $f : \Omega \rightarrow \mathbb{R}$ une application de classe C^1 et $(x_0, y_0, y'_0) \in \Omega$.

1. Il existe une solution maximale (I, φ) de (E) telle que $\varphi(x_0) = y_0$, $\varphi'(x_0) = y'_0$.
2. Soit (J, ψ) une solution de (E) telle que $x_0 \in J$ et $\psi(x_0) = y_0$ et $\psi'(x_0) = y'_0$. On a $J \subset I$ et $\psi = \varphi|_J$. En particulier il n'y a qu'une solution maximale vérifiant $\varphi(x_0) = y_0$, $\varphi'(x_0) = y'_0$.
3. L'intervalle I de définition de la solution maximale est ouvert.

La proposition 17.1.1 devient

PROPOSITION 17.3.1

Soient Ω un ouvert de \mathbb{R}^3 , $f : \Omega \rightarrow \mathbb{R}$ de classe C^1 et $\psi : J \rightarrow \mathbb{R}$ une solution de (E) définie sur un intervalle ouvert $J =]\alpha, \beta[$ avec $\beta < +\infty$. Supposons que les fonctions φ et φ' aient des limites finies λ et λ' quand x tend vers β et que $(\beta, \lambda, \lambda') \in \Omega$. Alors ψ se prolonge en une solution sur $]\alpha, \beta]$ et par conséquent ψ n'est pas une solution maximale.

Notons que si φ' a une limite finie λ' quand x tend vers β , φ a automatiquement une limite finie quand x tend vers β .

17.4 Champs de vecteurs

17.4.1 Généralités

DEFINITION 17.4.1

Soit U un ouvert de \mathbb{R}^p . Un champ de vecteurs sur U est une application $V : U \rightarrow \mathbb{R}^p$.

Intuitivement, il faut penser à un champ de vecteurs comme à une application qui à un point $M \in U$, considéré comme sous ensemble de l'espace affine \mathbb{R}^p associe un vecteur $V(M)$ de l'espace vectoriel associé, que l'on représenterait comme un vecteur "lié" d'origine M .

Etant donné un tel champ, on peut chercher les courbes paramétrées $t \rightarrow \gamma(t)$ à valeurs dans U telles que pour chaque valeur de t , le vecteur vitesse $\gamma'(t)$ soit égal au champ de vecteur en ce point ce qui conduit à l'équation $\gamma'(t) = V(\gamma(t))$.

A un champ de vecteurs $V : U \rightarrow \mathbb{R}^p$ on associera donc l'ouvert $\Omega = \mathbb{R} \times U$ de $\mathbb{R} \times \mathbb{R}^p$, l'application $F : \Omega \rightarrow \mathbb{R}^p$ définie par $F(t, X) = V(X)$ et l'équation différentielle

$$(S) \quad X' = F(t, X) = V(X)$$

Une telle équation, où le second membre ne dépend pas de t est dite autonome. Si $\varphi : I \rightarrow \mathbb{R}^p$ est une solution, on a par définition $\varphi(t) \in U$ pour tout t . Notons que si $\varphi : I \rightarrow \mathbb{R}^p$ est une solution, pour tout réel a l'application $\varphi_a : I + a \rightarrow \mathbb{R}^p$ définie par $\varphi_a(t) = \varphi(t - a)$ est aussi solution. Si la première est une solution maximale, la seconde aussi. Il en résulte que l'on pourra dans l'étude du problème de Cauchy en (t_0, X_0) se limiter au cas où $t_0 = 0$.

Si le champ de vecteurs est de classe C^1 , le théorème de Cauchy-Lipschitz s'applique. On obtient donc

THEOREME 17.4.1

Soit $V : U \rightarrow \mathbb{R}^p$ un champ de vecteurs de classe C^1 et $X_0 \in U$. Il existe une solution maximale unique $\varphi : I \rightarrow U$ telle que $\varphi(0) = X_0$. L'intervalle de définition I de cette solution est un intervalle ouvert.

La solution maximale φ_a telle que $\varphi_a(a) = X_0$ est définie sur l'intervalle $I + a = \{s + a ; s \in I\}$ par $\varphi_a(t) = \varphi(t - a)$.

On appelle orbite du champ de vecteurs V le support d'une solution maximale considérée comme courbe paramétrée à valeurs dans U : si (I, φ) est une solution maximale, l'orbite correspondante est $C_\varphi = \{\varphi(t) ; t \in I\}$.

THEOREME 17.4.2

Soit $V : U \rightarrow \mathbb{R}^p$ un champ de vecteurs de classe C^1 . Deux orbites distinctes de V ont une intersection vide.

preuve

Soient (I, φ) et (J, ψ) deux solutions maximales telles que $C_\varphi \cap C_\psi \neq \emptyset$. Il existe donc $t_0 \in I$ et $s_0 \in J$ tels que $\varphi(t_0) = \psi(s_0)$. Soit $X_0 = \varphi(t_0) = \psi(s_0)$. Soient $J' = J + t_0 - s_0 = \{s + t_0 - s_0 ; s \in J\}$. Soit $\theta : J' \rightarrow U$ définie par $\forall x \in J' \theta(x) = \psi(x + s_0 - t_0)$. Ceci a un sens puisque $x + s_0 - t_0 \in J$ dès que $x \in J'$. D'une part, $t_0 \in J'$ et d'autre part θ est une solution maximale de (S) qui vérifie $\theta(t_0) = \psi(s_0) = \varphi(t_0)$. Par unicité des solutions maximales du problème de Cauchy pour la donnée initiale (t_0, X_0) on en déduit $J' = I$ et $\theta = \varphi$. De là on voit facilement que $C_\varphi = C_\psi$.

Comme tout point X_0 de U appartient à une orbite, celle de la solution maximale passant par X_0 au temps 0, les orbites de V forment une partition de U et la relation entre points de U "appartenir à la même orbite" est une relation d'équivalence.

Remarque

Soit $X_0 \in U$ tel que $V(X_0) = 0$. (On dit que X_0 est un zéro du champ de vecteurs V). La solution maximale passant par X_0 au temps 0 est la fonction constante $t \rightarrow X_0$, définie sur \mathbb{R} . L'orbite d'un zéro de V est donc un singleton.

Exemple

Prenons $U = \mathbb{R}^2$ euclidien et V défini par $V(x, y) = (-y, x)$. Le système associé est linéaire à coefficients constants. Il s'écrit $\begin{pmatrix} x' \\ y' \end{pmatrix} = J \begin{pmatrix} x \\ y \end{pmatrix}$ avec $J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. Les solutions maximales sont définies sur \mathbb{R} par $t \rightarrow \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \exp(tJ) \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$. En passant aux affixes complexes, on obtient $z(t) = e^{it} z_0$. Les orbites sont l'origine, zéro du champ, et les cercles de centre O parcourus dans le sens direct.

17.4.2 Exemple

L'exemple qui suit illustre les nombreuses techniques qui permettent l'étude qualitative des solutions d'équations différentielles.

Soit $a > 0$. On considère le champ de vecteurs défini sur \mathbb{R}^2 par $V(x, y) = (X(x, y), Y(x, y))$ avec $X(x, y) = y$, $Y(x, y) = x(a^2 - x^2)$. Le système associé est

$$(S) \quad \begin{cases} x' &= y \\ y' &= x(a^2 - x^2) \end{cases}$$

1. V est de classe C^∞ sur \mathbb{R}^2 . Le théorème de Cauchy Lipschitz s'applique en chaque point $m_0(x_0, y_0) \in \mathbb{R}^2$ et les solutions sont de classe C^∞ .

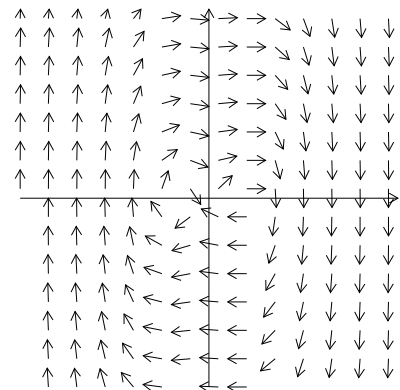
Les zéros du champ V sont les points $O(0, 0)$, $A(a, 0)$, $A'(-a, 0)$. Les applications constantes $\varphi_O : t \rightarrow (0, 0)$, $\varphi_A : t \rightarrow (a, 0)$ et $\varphi_{A'} : t \rightarrow (-a, 0)$ sont des solutions définies sur \mathbb{R} .

2. Soit $\varphi = (x, y) : I \rightarrow \mathbb{R}^2$ une solution maximale. On a $x'' = y' = x(a^2 - x^2)$ donc $2x'x'' - 2xx'(a^2 - x^2) = 0$ donc la fonction $t \rightarrow x'^2(t) - a^2x^2(t) + \frac{1}{2}x^4(t)$ est constante sur I .

Remarque

Ceci est une propriété très générale : considérons une équation différentielle de la forme $x'' = f(x)$ où f est de classe C^1 sur un ouvert de \mathbb{R} . Si F est une primitive de f , la fonction $\frac{1}{2}x'^2 - F(x)$ est constante pour toute solution y comme on le voit sans peine en dérivant.

Une telle équation est appelée équation de Newton. On la rencontre souvent en mécanique : soit W un ouvert de \mathbb{R}^d et $U : W \rightarrow \mathbb{R}$ une fonction de classe C^2 (potentiel). Le mouvement d'une particule de masse m soumise à la force $f = -\nabla U$ a pour équation $mx''(t) = -\nabla U(x(t))$. Elle décrit donc une trajectoire $t \rightarrow x(t)$ où x est solution de l'équation différentielle $mx'' = f(x)$. Le long de



cette trajectoire, la quantité $\frac{1}{2}mx'^2 + U(x)$ est constante. C'est l'énergie totale de la particule, le terme $U(x)$ étant l'énergie potentielle et le terme $\frac{1}{2}mx'^2$ l'énergie cinétique. Dans l'exemple, $d = 1$, $f(x) = x(a^2 - x^2)$ et $U = -F$ est une primitive de $-f$.

Soit $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}$ définie par $\Phi(x, y) = y^2 + \frac{1}{2}(x^2 - a^2)^2$. Pour toute solution $\varphi = (x, y)$, la fonction $t \rightarrow \Phi(x(t), y(t))$ est donc constante. Cette constante est nécessairement un nombre positif. Pour des raisons pratiques, nous mettrons cette constante sous la forme μa^4 avec $\mu > 0$. Nous noterons E_μ la ligne de niveau

$$E_\mu = \{(x, y) \in \mathbb{R}^2 \mid \Phi(x, y) = \mu a^4\} = \left\{ (x, y) \in \mathbb{R}^2 \mid \frac{1}{2}(x^2 - a^2)^2 + y^2 = \mu a^4 \right\}$$

L'ensemble E_0 est formé des trois zéros du champs de vecteurs.

3. Pour tout $\mu > 0$, l'ensemble E_μ est compact. En effet, il est fermé comme image réciproque de $\{\mu a^4\}$ par l'application continue Φ et manifestement borné.

Montrons que ceci implique que toutes les solutions maximales sont définies sur \mathbb{R} . En fait, on va montrer la proposition suivante qui améliore le résultat 17.1.1

PROPOSITION 17.4.1

Soient U un ouvert de \mathbb{R}^p , $V : U \rightarrow \mathbb{R}^p$ un champ de vecteurs de classe C^1 et K un compact contenu dans U . Soit $\varphi : I \rightarrow U$ une solution maximale de l'équation $X' = V(X)$ telle que $\forall t \in I \varphi(t) \in K$. Alors $I = \mathbb{R}$.

On résume cette proposition en disant que si une solution φ ne sort pas d'un compact K , alors elle est définie sur \mathbb{R} tout entier.

preuve

Soit $\varphi : I \rightarrow \mathbb{R}^p$ une solution maximale telle que $\varphi(t) \in K$ pour tout $t \in I$. Posons $I =]\alpha, \beta[$ et supposons par exemple $\beta < \infty$. Munissons \mathbb{R}^p d'une norme $\| \cdot \|$. Soit M tel que $\|V(x)\| \leq M$ pour tout $x \in K$. Un tel M existe car la fonction continue $x \rightarrow \|V(x)\|$ est majorée sur le compact K . Pour tout $t \in I$ on a $\|\varphi'(t)\| = \|V(\varphi(t))\| \leq M$ donc si $s, t \in I$ on a

$$\|\varphi(t) - \varphi(s)\| \leq M|t - s|$$

Il s'ensuit que pour toute suite (t_n) de points de I de limite β , la suite $\varphi(t_n)$ est de Cauchy donc converge. (C'est ici qu'on utilise l'hypothèse $\beta < \infty$: si (t_n) tend vers β la suite (t_n) est de Cauchy). On sait que ceci implique que $\lim_{t \rightarrow \beta} \varphi(t)$ existe. Si λ désigne cette limite, on a $\lambda \in K$ car K est fermé, donc a fortiori $\lambda \in W$. On peut alors appliquer la proposition 17.1.1 et conclure que φ se prolonge en une solution sur $]\alpha, \beta]$.

4. Soit $\varphi = (x, y) : \mathbb{R} \rightarrow \mathbb{R}^2$ une solution.

- Pour tout t_0 , $t \rightarrow \varphi(t_0 + t)$ est solution sur \mathbb{R} .
- $t \rightarrow -\varphi(t)$ est solution sur \mathbb{R} .
- ψ définie par $\psi(t) = (x(-t), -y(-t))$ est solution sur \mathbb{R} .

Montrons le pour la dernière : posons $x_1(t) = x(-t)$ et $y_1(t) = -y(-t)$. Il vient $x'_1(t) = -x'(-t) = -y(-t) = y_1(t)$ et $y'_1(t) = y'(-t) = x(-t)[a^2 - x(-t)^2] = x_1(t)[a^2 - x_1^2(t)]$ d'où la conclusion.

5. Allure des ensembles E_μ pour $\mu > 0$.

E_μ est la réunion des graphes des fonctions $x \rightarrow \pm \sqrt{\mu a^4 - \frac{1}{2}(x^2 - a^2)^2}$. L'étude est facile. L'ensemble E_μ admet les deux axes de coordonnées comme axe de symétrie. Il y a trois cas :

1. $0 < \mu < \frac{1}{2}$.

E_μ est réunion de deux courbes fermées symétriques par rapport à l'axe des y . Chacune de ces courbes est compacte. Nous les noterons C_μ^+ et C_μ^- . Dans ce cas, E_μ rencontre l'axe des x en quatre points, d'abscisses $\pm a\sqrt{1 \pm \sqrt{2\mu}}$.

2. $\mu = \frac{1}{2}$

E_μ a la forme d'un "huit". L'origine est un point double, les tangentes à l'origine sont les droites d'équation $y = \pm ax$. Nous noterons $\Gamma_+ = \{(x, y) \in E_{1/2} \mid x > 0\}$ et $\Gamma_- = \{(x, y) \in E_{1/2} \mid x < 0\}$ de sorte que $E_{1/2} = \Gamma_+ \cup \{O\} \cup \Gamma_-$. $E_{1/2}$ rencontre l'axe des x aux points d'abscisses $\pm a\sqrt{2}$ et en O .

3. $\mu > \frac{1}{2}$

E_μ est une courbe fermée. Elle rencontre l'axe des x aux points d'abscisses $\pm a\sqrt{1 + \sqrt{2\mu}}$. Voir plus loin l'allure de ces ensembles.

6. Orbites de V

Nous allons prouver le résultat suivant :

Les orbites de V sont

1. les singletons $\{O\}$, $\{A\}$, $\{A'\}$.
2. les courbes Γ_+ et Γ_- .
3. Toutes les E_μ pour $\mu > 1/2$
4. Toutes les courbes C_μ^+ et C_μ^- pour $0 < \mu < 1/2$.

De plus, dans les cas 3. et 4., les solutions sont périodiques.

Commençons par un résultat général

PROPOSITION 17.4.2

Soient U un ouvert de \mathbb{R}^p , $V : U \rightarrow \mathbb{R}^p$ un champ de vecteurs de classe C^1 et $\varphi : \mathbb{R} \rightarrow U$ une solution (nécessairement) maximale de l'équation $X' = V(X)$. Supposons que $\lim_{t \rightarrow +\infty} \varphi(t)$ existe, soit p et que $p \in U$. Alors p est un zéro de V .

preuve

Supposons $V(p) \neq 0$. Alors l'une des composantes de $V(p)$, par exemple $X_1(p)$ est non nulle. On a, en posant $\varphi = (\varphi_1, \dots, \varphi_p)$, $\varphi_1'(t) = X_1[\varphi_1(t)]$, donc lorsque t tend vers l'infini, $\varphi_1'(t)$ tend vers $X_1(p) \neq 0$ car $p \in U$. Mais c'est un résultat classique d'analyse que ceci implique $\lim_{t \rightarrow \infty} \varphi_1(t) = \pm\infty$ le signe étant celui de $X_1(p)$. Contradiction.

Soient $x_0 > 0$, $x_0 \neq a$, $m_0 = (x_0, 0)$. Soit $\varphi = (x, y) : \mathbb{R} \rightarrow \mathbb{R}^2$ la solution telle que $\varphi(0) = m_0$. Posons $\mu_0 = \frac{1}{2a^4}(x_0^2 - a^2)^2$ de sorte que $m_0 \in E_{\mu_0}$. On a alors $C_\varphi \subset E_{\mu_0}$.

1er cas : $x_0 > a\sqrt{2}$ i.e. $\mu_0 > 1/2$.

On a donc $x(0) = x_0$, $y(0) = 0$, $y'(0) = x_0(a^2 - x_0^2) < 0$. Par conséquent, $y(t)$ est strictement négative pour $t > 0$ voisin de 0 : $\exists \eta > 0$ $0 < t < \eta \Rightarrow y(t) < 0$.

• Il existe $t > 0$ tel que $y(t) = 0$. Supposons le contraire. Alors, $\forall t > 0$, $y(t)$ est de signe constant, donc strictement négatif d'après ce qui précède. De $x' = y$ on déduit que x est strictement décroissante. Mais $(x(t), y(t)) \in E_{\mu_0}$ compact, donc $x(t)$ est minoré et par conséquent a une limite finie λ quand t tend vers $+\infty$. Or $y(t) = -\sqrt{\mu_0 a^4 - \frac{1}{2}(x^2(t) - a^2)^2}$ donc $y(t)$ a aussi une limite finie λ' . $\forall t (x(t), y(t)) \in E_{\mu_0} \Rightarrow (\lambda, \lambda') \in \overline{E_{\mu_0}} = E_{\mu_0}$, D'après la proposition 17.4.2, (λ, λ') est un zéro de V ce qui est impossible car E_{μ_0} ne contient aucun zéro de V . D'où l'existence d'un t tel que $y(t) = 0$.

• Soit $T = \inf\{t > 0 \mid y(t) = 0\}$. Cet ensemble est fermé dans $[\eta, +\infty[$ donc dans \mathbb{R} donc $y(T) = 0$ et $T > 0$. Sur $]0, T[$ $x' = y$ est strictement négatif. La fonction x est strictement décroissante donc $x_1 = x(T) < x_0$. Le point $(x_1, 0)$ appartient à E_{μ_0} donc $x_1 = -x_0$. La fonction x induit une bijection strictement décroissante de $[0, T]$ sur $[x_1, x_0]$. Donc $\varphi([0, T]) = \{(x, y) \in E_{\mu_0} \mid y \leq 0\}$.

• Posons $\psi(t) = (x(2T - t), -y(2T - t))$. D'après le point 4., c'est une solution du système. On a $\psi(T) = (x(T), -y(T)) = (x_1, 0) = \varphi(T)$. ψ et φ sont deux solutions sur \mathbb{R} du système qui coïncident pour $t = T$. Par unicité, elles sont égales, donc $\forall t \in \mathbb{R} \psi(t) = \varphi(t)$ soit

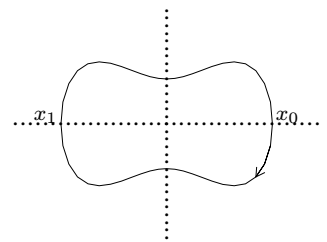
$$(I) \quad \forall t \in \mathbb{R} \quad x(2T - t) = x(t), \quad -y(2T - t) = y(t)$$

On en déduit d'abord que l'ensemble $\varphi([T, 2T])$ est le symétrique par rapport à l'axe des x de $\varphi([0, T])$. Donc

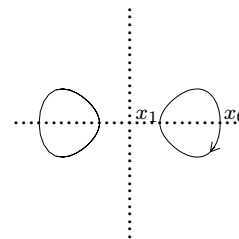
$$\varphi([0, 2T]) = E_{\mu_0} \quad \text{d'où} \quad C_\varphi = E_{\mu_0}$$

Ensuite, l'égalité (I) a lieu pour $t = 0$ ce qui donne $x(0) = x(2T)$ et $y(2T) = -y(0) = 0 = y(0)$. Les fonctions $t \rightarrow \varphi(t)$ et $t \rightarrow \varphi(t + 2T)$ sont donc deux solutions du système qui coïncident pour $t = 0$. Donc elles sont égales. La fonction φ est donc $2T$ -périodique. (N.B. Ce T dépend bien entendu de x_0).

2ème cas : $a < x_0 < a\sqrt{2}$ i.e. $\mu_0 < 1/2$.



Le raisonnement est exactement le même que dans le premier cas en remplaçant l'ensemble compact E_{μ_0} par l'ensemble compact $C_{\mu_0}^+$. Le point m_0 appartient à cet ensemble et $\varphi(\mathbb{R})$ est connexe et contenu dans E_{μ_0} donc est contenu dans $C_{\mu_0}^+$. On a toujours $y'(0) < 0$ et pour les mêmes raisons il existe $t > 0$ tel que $y(t) = 0$. On définit encore $T = \inf\{t > 0 \mid y(t) = 0\}$. On a alors $x(T) = x_1$ où cette fois $x_1 = a(1 - \sqrt{2\mu_0})$ (et $x_0 = a(1 + \sqrt{2\mu_0})$). On conclut comme ci dessus que $C_\varphi = C_{\mu_0}^+$ et que φ est périodique de période $2T$.

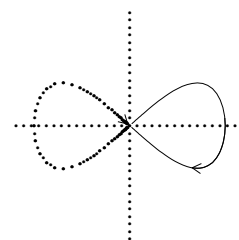


Troisième cas : $x_0 = a\sqrt{2}$ donc $\mu_0 = 1/2$.

• L'ensemble connexe $\varphi(\mathbb{R})$ est contenu dans $E_{1/2}$ et ne contient pas le point O (car O est une orbite). Donc $\varphi(\mathbb{R}) \subset \Gamma_+$.

• Montrons que $\forall t > 0 \ y(t) < 0$. On suppose le contraire et on définit $T = \inf\{t > 0 \mid y(t) = 0\}$. Comme dans les autres cas, x décroît strictement sur $[0, T]$ et $x_1 = x(T) < x_0$. Le point $(x_1, 0)$ est sur $\Gamma^+ \cap (Ox)$ et distinct de m_0 ce qui est impossible car cet ensemble est réduit à m_0 . D'où la conclusion.

• La fonction x décroît sur \mathbb{R}^+ et est minorée. Elle admet donc une limite finie $\lambda \geq 0$. Comme $y(t) = -\sqrt{a^2x^2 - x^4/2}$ on en déduit que $y(t)$ admet aussi une limite finie λ' quand t tend vers l'infini. D'après la proposition 17.4.2, (λ, λ') est un zéro de V . On a donc $(\lambda, \lambda') = (0, 0)$ où $(\lambda, \lambda') = (a, 0)$. Mais $(\lambda, \lambda') \in \overline{\Gamma_+}$ donc $(\lambda, \lambda') \neq (a, 0)$ et par conséquent $(\lambda, \lambda') = (0, 0)$.



• On peut alors conclure : on a $\varphi(\mathbb{R}^+) = \{(x, y) \in \Gamma_+ \mid y < 0\}$. La fonction $\psi(t) = (x(-t), -y(-t))$ est solution de (S) et $\psi(0) = m_0$. Donc $\psi = \varphi$ soit $\forall t \in \mathbb{R} \ x(-t) = x(t)$ et $y(-t) = -y(t)$. On en déduit que $\varphi(\mathbb{R}_-)$ est le symétrique par rapport à Ox de $\varphi(\mathbb{R}_+)$ donc que $\varphi(\mathbb{R}) = \Gamma_+$.

Enfin, si φ est solution, $-\varphi$ aussi donc les symétriques, par rapport à l'origine des ensembles précédemment trouvés sont aussi des orbites. Le symétrique de E_μ pour $\mu > 1/2$ est lui même. Le symétrique de C_μ^+ est C_μ^- et le symétrique de Γ_+ est Γ_- ce qui achève la preuve.

Remarque : On obtient du même coup les résultats concernant les solutions de l'équation différentielle $x'' = x(a^2 - x^2)$. En effet, une fonction x est solution de cette équation ssi c'est la première coordonnée d'une solution φ du système (S). On a donc

1. Toutes les solutions de cette équation sont définies sur \mathbb{R} .
 2. Les fonctions $t \rightarrow 0$, $t \rightarrow \pm a$ sont des solutions.
 3. Il existe une solution ψ , paire croissant de 0 à $a\sqrt{2}$ lorsque t varie de $-\infty$ à 0, puis décroissant jusqu'à 0 quand t varie de 0 à $+\infty$.
- La fonction $-\psi$ est aussi solutions.
4. Toutes les autres solutions sont périodiques.

La figure ci dessous donne l'allure des orbites du champ V .

