TOPOLOGIE ET CALCUL DIFFÉRENTIEL

Jean-Yves CHEMIN
Laboratoire J.-L. Lions, Case 187
adresse électronique: chemin@ann.jussieu.fr

Liste des questions de cours pour l'examen

Théorème 3.1.2 page 29 de Bolzano - Weierstrass (énoncé et démonstration)

Théorème 1.4.2 page 12 (énoncé et démonstration)

Théorème 1.3.3 page 10 (énoncé et démonstration)

Théorème 2.1.5 page 22 (énoncé et démonstration)

Théorème 2.2.1 page 23 (énoncé et démonstration)

Propoposition 3.1.4 page 30 (énoncé et démonstration)

Théorème 3.3.7 page 34 (énoncé et démonstration)

Théorème 4.1.4 page 40 (énoncé et démonstration)

Proposition 5.1.6 page 49 et Théorème 5.1.7 page 50 (énoncés et démonstrations)

Théorème 5.3.1 page 54 (énoncé et démonstration)

Théorème 5.5.1 page 57 (énoncé et démonstration)

Théorème 6.3.1 page 67 (énoncé et démonstration)

Théorème 7.3.2 page 79 (énoncé et démonstration)

Théorème 8.2.1 page 86 (énoncé et démonstration)

Théorème 9.3.1 page 95 (énoncé et démonstration)

Sommaire

Chapitre 1

Espaces métriques

1.1 Définition des espaces métriques et premiers exemples

Définition 1.1.1. *Soit X un ensemble, on appelle distance sur X toute application d de $X \times X$ dans $\mathbb{R}^+$ telle que*

$$\begin{aligned} d(x,y) &= 0 \Longleftrightarrow x = y \\ d(x,y) &= d(y,x) \\ d(x,y) &\leq d(x,z) + d(z,y) \quad \textit{(inégalité triangulaire)} \end{aligned}$$

pour tous éléments x, y, z de X. Le couple (X, d) est appelé un espace métrique.

Quelques exemples

- Prenons $X = \mathbb{R}$ et $d(x,y) = |x - y|$. Cela définit un espace métrique.
- Prenons $X = \mathbb{R}^N$ et choisissons les différentes distances suivantes:

$$\begin{aligned} d_e(x,y) &\stackrel{\text{déf}}{=} \left(\sum_{j=1}^{N} (x_j - y_j)^2\right)^{\frac{1}{2}} \\ d_\infty(x,y) &\stackrel{\text{déf}}{=} \max_{1 \leq j \leq N} |x_j - y_j| \\ d_1(x,y) &\stackrel{\text{déf}}{=} \sum_{j=1}^{N} |x_j - y_j|. \end{aligned}$$

Démontrons que ce sont bien des distances. On a, pour tout j,

$$|x_j - y_j| \leq |x_j - z_j| + |z_j - y_j|.$$

Par sommation, on obtient que d_1 est bien une distance. De plus, on a, pour tout j,

$$|x_j - y_j| \leq |x_j - z_j| + |z_j - y_j| \leq d_\infty(x,z) + d_\infty(z,y)$$

et on en déduit que d_∞ est bien une distance.

Démontrons que d_e est bien une distance. Les deux premières propriétés sont évidentes. Démontrons la troisième (l'inégalité triangulaire). Soit x, y et z trois points de $\mathbb{R}^N$. Écrivons que

$$(x_j - y_j)^2 = (x_j - z_j + z_j - y_j)^2 = (x_j - z_j)^2 + 2(x_j - z_j)(z_j - y_y) + (z_j - y_j)^2.$$

Par sommation par rapport à j, on trouve que

$$d_e(x,y)^2 = d_e(x,z)^2 + 2\sum_{j=1}^{N}(x_j - z_j)(z_j - y_y) + d_e(z,y)^2. \tag{1.1}$$

Pour la commodité du lecteur, nous allons redémontrer l'inégalité de Cauchy-Schwarz à savoir

$$\sum_{j=1}^{N} a_j b_j \leq \Big(\sum_{j=1}^{N} a_j^2\Big)^{\frac{1}{2}} \Big(\sum_{j=1}^{N} b_j^2\Big)^{\frac{1}{2}}. \tag{1.2}$$

Comme, pour tout réel λ,

$$\sum_{j=1}^{N}(a_j + \lambda b_j)^2 \geq 0,$$

on a, pour tout réel λ,

$$\sum_{j=1}^{N} a_j^2 + 2\lambda \sum_{j=1}^{N} a_j b_j^2 + \lambda^2 \sum_{j=1}^{N} b_j^2 \geq 0.$$

Le discrimant du polynôme en λ ci-dessus est donc négatif, d'où l'inégalité (1.2).

Appliquons cette inégalité à (1.1). Ceci donne

$$d_e(x,y)^2 \leq d_e(x,z)^2 + 2d_e(x,z)d_e(z,y) + d(z,y)^2 = \big(d_e(x,z) + d_e(z,y)\big)^2,$$

ce qui achève la démonstration du fait que d_e est une distance. Il s'agit de la distance dite "euclidienne".

- Prenons à nouveau $X = \mathbb{R}$, considérons une application injective φ de $\mathbb{R}$ dans $\mathbb{R}$ et définissons

  $$d_\varphi(x,y) = |\varphi(x) - \varphi(y)|.$$

  L'application d_φ est une distance sur X.

- Soit X un ensemble quelconque. On définit l'application suivante

  $$d \left\{ \begin{array}{lcl} X \times X & \longrightarrow & \mathbb{R}^+ \\ (x,y) & \longmapsto & 1 \text{ si } x \neq y\,,\ 0 \text{ sinon} \end{array} \right.$$

  C'est une distance sur X (Exercice: vérifiez-le!).

La notion de boule jouera un rôle crucial dans toute la suite.

Définition 1.1.2. *Soient (X,d) un espace métrique, x un point de X et α un réel strictement positif. On appelle boule ouverte de centre x et de rayon α, et l'on note $B(x,\alpha)$ l'ensemble des points y de X tels que $d(x,y) < \alpha$.*

La notion de distance permet de définir de manière abstraite, simple et générale, le concept de limite d'une suite et celui de fonction continue.

Définition 1.1.3 (convergence des suites)**.** *Soient $(x_n)_{n\in\mathbb{N}}$ une suite d'éléments d'un espace métrique (X,d) et ℓ un point de X. On dit que la suite $(x_n)_{n\in\mathbb{N}}$ converge vers ℓ si et seulement si*

$$\forall \varepsilon > 0\,,\ \exists\ n_0 \,/\, \forall n \geq n_0,\ x_n \in B(\ell, \varepsilon).$$

Remarquons que cette définition n'est qu'une écriture différente de la définition "classique" qui dit que la suite $(x_n)_{n\in\mathbb{N}}$ converge vers ℓ si et seulement si

$$\forall \varepsilon > 0\,,\ \exists n_0 \,/\, \forall n \geq n_0,\ d(x_n, \ell) < \varepsilon.$$

Définition 1.1.4 (continuité des fonctions)**.** *Soient (X,d) et (Y,δ) deux espaces métriques. On considère une fonction f de X dans Y et un point x_0 de X. La fonction f est continue en x_0 si et seulement si*

$$\forall \varepsilon > 0\,,\ \exists\ \alpha > 0 \,/\ f(B(x_0, \alpha)) \subset B(f(x_0), \varepsilon).$$

Remarquons que cette définition n'est qu'une écriture différente de la définition "classique" qui dit que la fonction f est continue en x_0 si et seulement si

$$\forall \varepsilon > 0\,,\ \exists\ \alpha > 0 \,/\, \forall x \in X,\ d(x, x_0) < \alpha \Longrightarrow \delta(f(x), f(x_0)) < \varepsilon.$$

Nous revisiterons ces notions dans les sections 1.3 et 1.4.

1.2 Les notions d'ouvert et de fermé

Définition 1.2.1. *Soit A une partie d'un espace métrique (X,d). On dit que A est un ouvert si et seulement si pour tout point x de A, il existe un réel strictement positif α tel que la boule ouverte de centre x de de rayon α soit incluse dans A. Par convention, l'ensemble vide est ouvert.*

La première classe d'exemples d'ouverts sont les boules ouvertes. En effet, on la la proposition suivante.

Proposition 1.2.2. *Soit (X,d) un espace métrique. Toute boule ouverte est un ouvert.*

Démonstration. Soit y un point de la boule ouverte $B(x,\alpha)$, considérons la boule ouverte de centre y et de rayon $\alpha - d(x,y)$ (qui est un nombre strictement positif). Soit z un point de cette boule. D'après l'inégalité triangulaire, on a

$$d(x,z) \leq d(x,y) + d(y,z) < d(x,y) + \alpha - d(x,y) = \alpha.$$

La proposition est démontrée. □

Proposition 1.2.3. *Toute réunion d'ouverts en est un. Toute intersection finie d'ouverts en est un.*

Démonstration. Soit $(U_\lambda)_{\lambda\in\Lambda}$ une famille quelconque d'ouverts et x un point de $U \stackrel{\text{déf}}{=} \bigcup_{\lambda\in\Lambda} U_\lambda$. Par définition, il existe un $\lambda \in \Lambda$ tel que $x \in U_\lambda$. Comme U_λ est ouvert (ce qui signifie égal à son intérieur),

$$\exists \alpha > 0 \,/\ B(x,\alpha) \subset U_\lambda \subset U$$

et donc U est ouvert.

Soit $U = \bigcap_{j=1}^{N}$ où les U_j sont des ouverts. Par définition, pour tout $j \in \{1, \cdots N\}$, il existe un réel strictement positif α_j tel que $B(x, \alpha) \subset U_j$. Soit $\alpha \stackrel{\text{déf}}{=} \min\{\alpha_j\ ,\ j \in \{1, \cdots N\}\}$. Pour tout j, la boule ouverte $B(x, \alpha)$ est incluse dans U_j et donc U est ouvert et la proposition est démontrée. □

Le lemme suivant décrit ce que sont les ouverts d'un espace métrique.

Lemme 1.2.4. *Soit (X, d) un espace métrique. Les ouverts de X sont les réunions de boules ouvertes.*

Démonstration. D'après la proposition 1.2.3 ci-dessus, une réunion de boules ouvertes est un ouvert. Soit U un ouvert de X, par définition, pour tout x de U, il existe un réel strictement positif α_x tel que $B(x, \alpha_x) \subset U$. On a donc

$$U \subset \bigcup_{x \in U} B(x, \alpha_x) \subset U.$$

D'où le lemme. □

Nous allons maintenant définir les ensembles fermés.

Définition 1.2.5. *Soit A une partie d'un espace métrique (X, d). On dit que A est un fermé si son complémentaire est ouvert, autrement dit si, pour tout x dans X,*

$$\Big(\forall \alpha \in \mathbb{R}^+\,, B(x, \alpha) \cap A \neq \emptyset\Big) \Longrightarrow x \in A.$$

Par convention, l'ensemble vide est fermé.

La proposition 1.2.3 implique immédiatement la proposition suivante.

Proposition 1.2.6. *Soit (X, d) un espace métrique. Toute intersection de fermés en est un. Toute réunion finie de fermés en est un.*

Définition 1.2.7. *Soit A une partie quelconque d'un espace métrique (X, d). L'*intérieur *de A est l'ensemble des points de A qui sont le centre d'une boule incluse dans A.*

*L'*adhérence *de A est l'ensemble des points de X tels que toute boule centrée en ce point rencontre A.*

La frontière *de A est l'ensemble des points de X tels que toute boule centrée en ce point rencontre à la fois A et son complémentaire.*

Ces ensembles sont notés respectivement $\overset{\circ}{A}, \overline{A}, \mathrm{Fr}A$.

Une partie A de X est dense *dans X si $\overline{A} = X$.*

La définition ci-dessus peut s'écrire de la manière suivante

$$x \in \overset{\circ}{A} \Longleftrightarrow (\exists \varepsilon > 0 \ \ B(x, \varepsilon) \subset A).$$

$$x \in \overline{A} \Longleftrightarrow (\forall \varepsilon > 0 \ \ B(x, \varepsilon) \cap A \neq \emptyset).$$

$$x \in \mathrm{Fr}A \Longleftrightarrow (\forall \varepsilon > 0 \ \ B(x, \varepsilon) \cap A \neq \emptyset \text{ et } B(x, \varepsilon) \cap (X \setminus A) \neq \emptyset.$$

D'après la définition, on a aussi $\mathrm{Fr}A = \overline{A} \cap \overline{A^c}$.

Proposition 1.2.8. *Soit A une partie d'un espace métrique (X,d). On a*

$$(\overset{\circ}{A})^c = \overline{A^c} \quad \text{et} \quad \overline{A}^c = (\overset{\circ}{A^c})\,.$$

Démonstration. Un point x de X appartient à $(\overset{\circ}{A})^c$ si et seulement si

$$\forall \alpha > 0\,,\ B(x,\alpha) \cap A^c \neq \emptyset,$$

ce qui signifie exactement que x appartient à l'adhérence du complémentaire de A. Un point x appartient à $(\overline{A})^c$ si et seulement si

$$\exists \alpha > 0\,,\ B(x,\alpha) \cap A = \emptyset,$$

c'est-à-dire $B(x,\alpha) \subset A^c$, ce qui signifie que $x \in (\overset{\circ}{A^c})$. La proposition est démontrée. □

Proposition 1.2.9. *L'intérieur de A est ouvert, et il contient tout ouvert inclus dans A. L'adhérence de A est fermée, et elle est incluse dans tout fermé contenant A.*

On résume ces propriétés en disant que l'intérieur de A est le plus grand ouvert inclus dans A, et l'adhérence de A est le plus petit fermé contenant A. En particulier, A est ouvert si et seulement si il est égal à son intérieur, et A est fermé si et seulement si il est égal à son adhérence.

Démonstration. Soit x un point de l'intérieur de A : il existe $\varepsilon > 0$ tel que la boule $B(x,\varepsilon)$ est incluse dans A. Pour voir que l'intérieur de A est ouvert, il s'agit de montrer que cette boule est en fait incluse dans l'intérieur de A. Nous avons vu que les boules ouvertes sont des ouverts (proposition 1.2.2) : pour tout $y \in B(x,\varepsilon)$, il existe $\alpha > 0$ tel que $B(y,\alpha) \subset B(x,\varepsilon) \subset A$, ce qui prouve que y est dans l'intérieur de A, comme voulu.

Pour montrer que l'intérieur contient tout ouvert O inclus dans A, il suffit de partir d'un point de O et d'écrire la définition des ouverts.

Les propriétés de l'adhérence peuvent s'en déduire par passage au complémentaire à l'aide de la proposition 1.2.8. □

Proposition 1.2.10 (critère séquentiel)**.** *Soit A une partie d'un espace métrique (X,d). Un point x de X appartient à $\overline{A}$ si et seulement si il existe une suite $(a_n)_{n\in\mathbb{N}}$ d'éléments de A tellle que $\lim_{n\to\infty} a_n = x$. La partie A est fermé si et seulement si pour toute suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de A, supposée convergeante, la limite de la suite $(x_n)_{n\in\mathbb{N}}$ appartient à A.*

Démonstration. Supposons que x soit limite d'une suite $(a_n)_{n\in\mathbb{N}}$ d'éléments de A. Pour tout réel strictement positif α, il existe un entier n_0 tel que a_{n_0} appartienne à $B(x,\alpha)$ ce qui implique que $B(x,\alpha) \cap A \neq \emptyset$. Donc $x \in \overline{A}$.

Réciproquement, supposons que $x \in \overline{A}$. Alors, pour tout entier strictement positif n, il existe un élément a_n de X tel que

$$a_n \in A \cap B(x, \frac{1}{n})$$

Soit $(a_n)_{n\in\mathbb{N}}$ la suite ainsi définie. La suite $(a_n)_{n\in\mathbb{N}}$ vérifie

$$d(a_n, x) \leq \frac{1}{n}\,\cdot$$

La suite $(a_n)_{n\in\mathbb{N}}$ converge donc vers x et la première partie de la proposition est démontrée.

La seconde partie découle de la première, et du fait que A est fermé si et seulement si $A = \overline{A}$. □

1.3 La continuité des fonctions décrite en termes d'ouverts

Commençons par redémonter quelques propriétés bien connues dans le cas usuel des fonctions réelles d'une variable réelle.

Proposition 1.3.1 (composition des fonctions continues)**.** *Soient (X, d), (Y, δ) et (Z, ρ) trois espaces métriques, f et g deux fonctions respectivement de X dans Y et de Y dans Z. Soit x_0 un point de X tel que f soit continue en x_0 et g le soit en $f(x_0)$. Alors la fonction $g \circ f$ est continue en x_0.*

Démonstration. Soit $\varepsilon > 0$. La fonction g étant continue en $y_0 = f(x_0)$, il existe $\alpha > 0$ tel que

$$g(B(y_0, \alpha) \subset B((g(y_0), \varepsilon).$$

La fonction f étant continue en x_0, il existe $\beta > 0$ tel que

$$f(B(x_0, \beta)) \subset B(f(x_0), \alpha).$$

On en déduit alors que

$$(g \circ f)(B(x_0, \beta) \subset B((g \circ f)(x_0), \varepsilon)$$

(on a utilisé l'implication $A \subset B \Rightarrow g(A) \subset g(B)$). Ceci conclut la démonstration. □

Proposition 1.3.2. *Considérons deux espaces métriques (X, d) et (Y, δ) et f une fonction de X dans Y et $(x_n)_{n\in\mathbb{N}}$ une suite d'éléments de X. On suppose que $(x_n)_{n\in\mathbb{N}}$ converge vers ℓ et que la fonction f est continue en ℓ. Alors la suite d'éléments de Y définie par $(f(x_n))_{n\in\mathbb{N}}$ converge vers $f(\ell)$.*

La démonstration de cette proposition est analogue à celle de la proposition précédente et est laissée en exercice au lecteur.

Nous allons maintenant donner une caractérisation des fonctions continues en tout point en terme d'ouverts (et de fermés).

Théorème 1.3.3 (Caractérisation des applications continues)**.** *Soit f une application entre deux espaces métriques (X, d) et (Y, δ). Les trois assertions suivantes sont équivalentes.*

- *L'application f est continue en tout point de X,*
- *l'image réciproque d'un ouvert est un ouvert*
- *l'image réciproque d'un fermé est un fermé.*

Démonstration. Remarquons tout d'abord que l'équivalence entre les points deux et trois résultent du fait que les complémentaires des ouverts sont les fermés (et réciproquement) et que

$$\begin{aligned} f^{-1}(U) &= \{x \in X \;/\; f(x) \in U\} \\ &= \{x \in X \;/\; f(x) \in U^c\}^c \\ &= (f^{-1}(U^c))^c. \end{aligned}$$

Démontrons maintenant l'équivalence entre les points un et deux. Tout d'abord, faisons une petite digression sur l'image réciproque d'un ensemble.

Lemme 1.3.4. *Soit f une application d'un ensemble X dans un ensemble Y. Soit A une partie de X et B une partie de Y. On définit*

$$f^{-1}(B) = \{x \in X \,/\, f(x) \in B\} \quad \textit{et} \quad f(A) = \{y \in Y \,/\, \exists a \in A \,/\, f(a) = y\}.$$

Alors on a

$$f(f^{-1}(B)) = B \cap f(X) \quad \textit{et} \quad f^{-1}(f(A)) = \{x \in X \,/\, \exists a \in A \,/\, f(x) = f(a)\}.$$

En particulier, $f(f^{-1}(B)) \subset B$ et $A \subset f^{-1}(f(A))$.

Démonstration. Soit y un élément de $B \cap f(X)$. Il existe un x de X tel que $y = f(X)$. Par définition de l'image réciproque d'un ensemble, $x \in f^{-1}(B)$ et donc $y \in f(f^{-1}(B))$ et donc $B \cap f(X) \subset f(f^{-1}(B))$. L'inclusion réciproque est évidente. Donc $f(f^{-1}(B)) = B \cap f(X)$.

De plus, par définition de l'image réciproque, x appartient à $f^{-1}(f(A))$, si et seulement si il existe un élément b de $f(A)$ tel que $f(x) = b$, c'est-à-dire qu'il existe un élément a de A tel que $f(x) = f(a)$. D'où le lemme. □

Suite de la démonstration du Théorème 1.3.3 Supposons f continue en tout point de X et considérons un ouvert U de (Y, δ) et un point x de $f^{-1}(U)$. L'ensemble U étant un ouvert qui contient $f(x)$, il existe par définition un réel strictement positif ε_0 tel que $B(f(x), \varepsilon_0) \subset U$. La fonction f étant continue en x, il existe $\alpha > 0$ tel que

$$f(B(x, \alpha)) \subset B(f(x), \varepsilon_0).$$

Ainsi donc

$$B(x, \alpha) \subset f^{-1}\Big(f(B(x), \alpha)\Big) \subset f^{-1}(B(f(x), \varepsilon_0)) \subset f^{-1}(U)$$

ce qui montre que U est ouvert.

Réciproquement, supposons que l'image réciproque de tout ouvert est un ouvert, et montrons que f est continue. Soit x_0 un point de X, et $\varepsilon > 0$. La boule $B(f(x_0), \varepsilon)$ est un ouvert de (Y, δ). Par hypothèse, $f^{-1}(B(f(x_0), \varepsilon))$ est un ouvert de (X, d). Donc il existe $\alpha > 0$ tel que $B(x_0, \alpha)$ soit inclus dans $f^{-1}(B(f(x_0), \varepsilon))$ et on a

$$f(B(x_0, \alpha)) \subset f\Big(f^{-1}(B(f(x_0), \varepsilon))\Big) \subset B(f(x_0), \varepsilon)$$

ce qui conclut la démonstration. □

1.4 Notion de valeur d'adhérence d'une suite

Le concept suivant sera très important pour la suite du cours.

Définition 1.4.1. *Soit $(x_n)_{n\in\mathbb{N}}$ une suite d'éléments d'un espace métrique (X, d). On définit l'ensemble (éventuellement vide) des valeurs d'adhérence de $(x_n)_{n\in\mathbb{N}}$ que l'on note $\mathrm{Adh}(x_n)$ par*

$$\mathrm{Adh}(x_n) \stackrel{\text{déf}}{=} \bigcap_{n\in\mathbb{N}} \overline{A}_n \quad \textit{avec} \quad A_n \stackrel{\text{déf}}{=} \{x_m\,, m \geq n\}.$$

Autrement dit, x est une valeur d'adhérence de la suite si et seulement si il existe des termes arbitrairement grands de la suite qui sont arbitrairement proches de x :

$$\forall \varepsilon > 0, \forall n_0 \geq 0,\ \exists n \geq n_0 \ \,/ x_n \in B(x, \varepsilon).$$

Exemples Dans $\mathbb{R}$ muni de la distance $|x-y|$, la suite $(x_n)_{n\in\mathbb{N}}$ définie par $x_n = n$ est telle que $\mathrm{Adh}(x_n) = \emptyset$ et la suite $(y_n)_{n\in\mathbb{N}}$ définie par $y_n = (-1)^{n+1}$ est telle que $\mathrm{Adh}(x_n) = \{-1, 1\}$.

Théorème 1.4.2. *Soit $(x_n)_{n\in\mathbb{N}}$ une suite d'éléments d'un espace métrique (X, d). Un point ℓ de X appartient à $\mathrm{Adh}(x_n)$ si et seulement si il existe une fonction strictement croissante ϕ de $\mathbb{N}$ dans $\mathbb{N}$ telle que*

$$\lim_{n\to\infty} x_{\phi(n)} = \ell.$$

Démonstration. Avant d'entamer la démonstration proprement dite, remarquons qu'une fonction strictement croissante de $\mathbb{N}$ dans $\mathbb{N}$ vérifie

$$\forall n \in \mathbb{N}\,,\ \phi(n) \geq n. \tag{1.3}$$

Démontrons cette propriété par récurrence. Elle est bien sûr vraie pour $n = 0$. Supposons la réalisée pour n. Comme ϕ est strictement croissante, on a $\phi(n+1) > \phi(n) \geq n$. Ceci implique que $\phi(n+1) \geq n+1$. En particulier, la suite $(\phi(n))_{n\in\mathbb{N}}$ tend vers l'infini avec n.

Revenons à notre démonstration. Si ℓ appartient à $\mathrm{Adh}(x_n)$, on définit par récurrence la fonction de la manière suivante: on choisit $\phi(0) = 0$ et puis, on définit $\phi(n+1)$ à partir de $\phi(n)$ comme

$$\phi(n+1) \stackrel{\text{déf}}{=} \min\Big\{m > \phi(n) \,/ d(x_m, \ell) < \frac{1}{n+1}\Big\}.$$

(Le fait que ℓ soit une valeur d'adhérence de la suite garantit que l'ensemble intervenant dans cette définition est non vide.) Par construction, la fonction ϕ est strictement croissante et nous avons, pour tout $n \geq 1$,

$$d(x_{\phi(n)}, \ell) \leq \frac{1}{n} \quad \text{et donc} \quad \lim_{n\to\infty} x_{\phi(n)} = \ell.$$

Réciproquement, s'il existe une fonction strictement croissante ϕ telle que $\lim\limits_{n\to\infty} x_{\phi(n)} = \ell$, pour tout $\varepsilon > 0$, il existe n_0 tel que

$$\forall n \geq n_0\,,\ d(x(\phi(n)), \ell) < \varepsilon.$$

Comme, d'après (1.3), $\phi(n) \geq n$, on en déduit que

$$\forall \varepsilon > 0\,, \forall n \in \mathbb{N}\,,\ \exists m \geq n \,/\ d(x_m, \ell) < \varepsilon,$$

ce qui signifie exactement que ℓ appartient à $\mathrm{Adh}(x_n)$. □

Proposition 1.4.3. *Soit $(x_n)_{n\in\mathbb{N}}$ une suite d'un espace métrique (X, d) que l'on suppose convergeant vers un point ℓ de X. Alors $\mathrm{Adh}(x_n) = \{\ell\}$.*

Démonstration. Par définition de la limite d'une suite, pour tout $\varepsilon > 0$, il existe n_0 tel que l'ensemble A_{n_0} de la définition ci-dessus soit inclus dans la boule ouverte $B(\ell, \varepsilon)$. Donc soit x un point de X distinct de ℓ, en prenant ε strictement inférieur à $d(\ell, x)$, on trouve que $x \notin \overline{A}_{n_0}$ et donc que $x \notin \overline{A}_n$ pour tout $n \geq n_0$. D'où la proposition. □

Remarque Il est possible qu'une suite $(x_n)_{n\in\mathbb{N}}$ soit telle que $\mathrm{Adh}(x_n) = \{\ell\}$ et ne converge pas. Considérons par exemple la suite de nombres réels définie par

$$x_{2n} = 2n \quad \text{et} \quad x_{2n+1} = \frac{1}{2n+1}\,\cdot$$

Bien que cette suite ne converge pas, on a $\mathrm{Adh}(x_n) = \{0\}$.

1.5 Notion de distance induite et de sous espace métrique

Définition 1.5.1. *Soit (X, d) un espace métrique et A une partie de X. On appelle sous-espace métrique associé à la partie A le couple $(A, d_{A\times A})$).*

Il est trivial que la restriction de la distance d à $A \times A$ définit une distance sur A. Il est maintenant nécessaire de comprendre ce que sont les ouverts du sous espace métrique A. Ceci est décrit par la proposition suivante.

Proposition 1.5.2. *Soient (X, d) un espace métrique et A une partie de X; les ouverts de $(A, d_{A\times A})$ sont les ensembles qui s'écrivent comme intersection de A avec un ouvert de X.*

Démonstration. Soit U un ouvert de X, démontrons que $U \cap A$ est un ouvert de $(A, d_{A\times A})$. C'est évident si $U \cap A = \emptyset$. Soit a un point de $U \cap A$. Comme U est un ouvert de X, il existe un réel strictement positif α tel que $B_X(a, \alpha) \subset U$ (où $B_X(a, \alpha)$ désigne l'ensemble des points x de X tels que $d(a, x) < \alpha$). Mais $B_X(a, \alpha) \cap A$ est exactement la boule ouverte de centre a et de rayon α au sens de $(A, d_{A\times A})$. Donc $U \cap A$ est un ouvert de l'espace métrique $(A, d_{A\times A})$.

Démontrons maintenant la réciproque. Soit U un ouvert de $(A, d_{A\times A})$, le lemme 1.2.4 dit que U est une réunion de boules ouvertes de $(A, d_{A\times A})$, plus précisement que

$$U = \bigcup_{x\in U} B_A(x, \alpha_x).$$

Or $B_A(x, \alpha_x) = B_X(x, \alpha_x) \cap A$. Soit $V \stackrel{\text{déf}}{=} \bigcup\limits_{x\in U} B_X(x, \alpha_x)$. C'est un ouvert de X car réunion de boules ouvertes (de (X, d)) et $U = V \cap A$ et la proposition est ainsi démontrée. $\square$

ATTENTION! Lorsque l'on parle d'ouverts et de fermés dans le cas de sous espaces métriques, il faut être prudent et réfléchir pour éviter les erreurs. Remarquons tout de suite que A est fermé et ouvert au sens de l 'espace métrique $(A, d_{A\times A})$ sans que cela préjuge en rien de ces propriétés en tant que partie de l'espace métrique (X, d).

Prenons un exemple un peu moins trivial. Soit $X = \mathbb{R}$ muni de la distance $d(x, y) = |x-y|$. Considérons quatre nombres réels $a < b < c < d$ et posons $A = [a, b] \cup [c, d]$. Nous allons démontrer que $[a, b]$ est un ouvert et un fermé de $(A, d_{A\times A})$.

Démontrons que $[a, b]$ est fermé dans A. Soit $(x_n)_{n\in\mathbb{N}}$ une suite déléments de $[a, b]$ convergeant (au sens de $(A, d_{A\times A})$ vers un élément ℓ de A. Cette suite converge au sens de (X, d) vers ℓ; Comme $[a, b]$ est un fermé de X, ℓ appartient à $[a, b]$. Donc $[a, b]$ est fermé **pour l'espace métrique** $(A, d_{A\times A})$. Démontrons qu'il est ouvert **pour l'espace métrique** $(A, d_{A\times A})$. Soit $x \in]a, b[$, si $\varepsilon < \min\{b - x, x - a\}$, alors l'intervalle ouvert (c'est-à-dire la boule ouverte) de centre x et de rayon ε est inclus dans $]a, b[$. Démontrons maintenant que a et b sont des points intérieurs de $[a, b]$ **pour l'espace métrique** $(A, d_{A\times A})$. Soit $\varepsilon < b - a$, la boule ouverte de A de centre a et de rayon ε est

$$]a - \varepsilon, a + \varepsilon[\cap A = [a, a + \varepsilon[$$

et donc incluse dans $[a, b]$. Donc a est un point intérieur de $[a, b]$ **pour l'espace métrique** $(A, d_{A\times A})$.

Démontrons que b est aussi un point intérieur de $[a, b]$ **pour l'espace métrique** $(A, d_{A\times A})$. Soit ε un rél strictement positif strictement inferieur à $\min\{b - a, c - b <\}$. La boule ouverte de A de centre b et de rayon ε est

$$]b - \varepsilon, b + \varepsilon[\cap A =]b - \varepsilon, b]$$

et donc incluse dans A. Donc le point a est un point intérieur de $[a, b]$ **pour l'espace métrique** $(A, d_{A\times A})$.

Cet exemple est à méditer.

1.6 Produit d'espaces métriques

Définition 1.6.1. *On se donne N espaces métriques $(X_1, d_1), \dots, (X_N, d_N)$. On peut munir l'ensemble produit*

$$X_1 \times \cdots \times X_N = \{(x_1, \dots, x_N) \mid x_1 \in X_1, \dots, x_N \in X_N\}$$

de la distance d_∞ définie par

$$d_\infty(\underbrace{(x_1, \dots, x_N)}_{\in X_1\times\cdots\times X_N}, \underbrace{(x'_1, \dots, x'_N)}_{\in X_1\times\cdots\times X_N}) := \max\big(d_1(x_1, x'_1), \dots, d_N(x_N, x'_N)\big)$$

dont le lecteur vérifiera qu'il s'agit bien d'une distance. On obtient ainsi un espace métrique appelé produit des N espaces $(X_1, d_1), \dots, (X_N, d_N)$.

Remarquons qu'une suite $(x_1^{(n)}, \dots, x_N^{(n)})_{n\in\mathbb{N}}$ dans cet espace converge vers une limite $(\ell_i, \dots, \ell_N)$ si et seulement si chacune des suites $(x_i^{(n)})_{n\in\mathbb{N}}$ converge vers ℓ_i.

La formule ci-dessus n'est pas la seule possible, on peut aussi définir, par analogie avec $\mathbb{R}^N$, les métriques d_1 et d_2 :

$$d_1((x_1, \dots, x_N), (x'_1, \dots, x'_N)) \stackrel{\text{déf}}{=} \sum_{i=1}^{N} d_i(x_i, x'_i)$$

$$d_2((x_1, \dots, x_N), (x'_1, \dots, x'_N)) \stackrel{\text{déf}}{=} \left(\sum_{i=1}^{N} d_i(x_i, x'_i)^2\right)^{\frac{1}{2}}.$$

Exemple Lorsqu'on applique cette construction avec $X_1 = \cdots = X_N = \mathbb{R}$ et en prenant pour chacune des distances d_i la distance usuelle sur $\mathbb{R}$, on retrouve l'espace métrique $\mathbb{R}^N$ munit des métriques définies au début du chapitre.

Définition 1.6.2. *Une distance d est* équivalente *à une autre distance δ définie sur le* ***même*** *ensemble X s'il existe une constante $k \geq 1$ telle que les inégalités*

$$d(x, y) \leq k\delta(x, y), \quad \delta(x, y) \leq kd(x, y)$$

sont vérifiées pour tout x, y dans X. On dira que les distances induisent la même topologie *si l'ensemble des parties ouvertes définies à l'aide de d est le même que l'ensemble des parties ouvertes définies à l'aide de δ.*

L'équivalence de distance est une *relation d'équivalence*, ce qui signifie qu'elle est *réflexive* (toute distance est équivalente à elle-même), *symétrique* (si d est équivalente à δ alors δ est équivalente à d), et *transitive* (si d est équivalente à δ, et si δ est équivalente à ρ, alors d est équivalente à ρ). Toutes ces propriétés sont très faciles à montrer. Le fait d'induire la même topologie est également une relation d'équivalence (c'est immédiat). La proposition suivante montre que les trois distances d_1, d_2, d_∞ que l'on vient de définir sur le produit de N ensemble induisent la même topologie : en, particulier les notions de suites convergeantes, ou d'applications continues, sont les mêmes pour les trois distances.

Proposition 1.6.3.

- *Les trois distances d_1, d_2, d_∞ définies sur l'espace produit $X_1 \times X_2$ sont équivalentes.*
- *Deux distances d, d' qui sont équivalentes sur un ensemble X induisent la même topologie.*

Démonstration. On donne la preuve seulement dans le cas du produit de deux espaces métriques, à charge pour le lecteur d'écrire les détails dans le cas d'un produit de N espaces avec N quelconque. Considérons d'abord deux nombres positifs a, b, et les quantités

$$A_1 = a + b, \quad A_2 = (a^2 + b^2)^{\frac{1}{2}}, \quad A_\infty = \max(a, b).$$

on a clairement $A_1 \leq 2A_\infty$ et $A_\infty \leq A_1$. En appliquant ceci avec $a = d_1(x_1, x_1')$ et $b = d_2(x_2, x_2')$, on obtient l'équivalence entre d_1 et d_∞. De même, on a $A_2 \leq \sqrt{2}A_\infty$ et $A_\infty \leq A_2$, on en déduit que d_∞ et d_2 sont équivalentes. Finalement, par transitivité, les trois distances sont équivalentes.

Soient maintenant d, δ deux distances équivalentes sur un ensemble X. Pour tout $x, y \in X$ et $\varepsilon > 0$, on a

$$\delta(x, y) < \frac{\varepsilon}{k} \Rightarrow d(x, y) < \varepsilon$$

ce qui signifie que la boule $B_d(x, \varepsilon)$ pour la première distance contient alors la boule $B_\delta(x, \varepsilon/k)$ pour la deuxième distance. Maintenant si O est un ouvert pour la distance d, et x un point de O, il existe $\varepsilon > 0$ tel que la boule $B_d(x, \varepsilon)$ soit incluse dans O, mais alors la boule $B_\delta(x, \varepsilon/k)$ est également incluse dans O, ce qui veut dire que O est également ouvert pour δ. La réciproque est analogue. □

La notion de distance équivalente peut être interprétée en termes de fonctions lipschitziennes.

Définition 1.6.4. *Soient (X, d) et $Y, \delta)$ deux espaces métriques, f une fonction de X dans Y et k un réel strictement positif. On dit que f est lipschitzienne de rapport k si et seulement si*

$$\forall (x, x') \in X^2\,,\ \delta(f(x), f(x')) \leq kd(x, x').$$

Remarquons qu'une fonction lipschitzienne est continue. (Exercice: vérifiez le!). Donnons quelques exemples de fonctions lischitziennes.

Proposition 1.6.5. *Soient (X, d) un espace métrique, x_0 un point de X et A une partie de X. La fonction $x \mapsto d(x, x_0)$ est 1-lipschitzienne. Définissons*

$$d(x, A) \stackrel{\text{déf}}{=} \inf_{a \in A} d(x, a).$$

L'application $x \mapsto d(x, A)$ est 1-lipschitzienne.

Démonstration. L'inégalité triangulaire dit que

$$d(x, x_0) \leq d(x, x') + d(x', x_0)$$

ce qui peut s'écrire

$$d(x, x_0) - d(x', x_0) \leq d(x, x').$$

Par symétrie, on en déduit que

$$\big|d(x, x_0) - d(x', x_0)\big| \leq d(x, x')$$

ce qui assure le premier résultat.

Soit a un point quelconque de A. D'après l'inégalité triangulaire, on a

$$d(x,a) \leq d(x,x') + d(x',a)$$

La borne inférieure est un minorant. On a donc, pour tout a de A que $d(x,A) \leq d(x,a)$ et donc que, pour tout a de A

$$d(x,A) \leq d(x,x') + d(x',a)$$

ce qui s'écrit

$$d(x,A) - d(x,x') \leq d(x',a).$$

Ceci signifie que $d(x,A) - d(x,x')$ est un minorant de $\{d(x',a)\,,\ a \in A\}$. comme la borne inférieure est le plus grand des minorants, on a

$$d(x,A) - d(x,x') \leq d(x',A)$$

ce qui s'écrit

$$d(x,A) - d(x',A) \leq d(x,x').$$

Par symétrie, on en déduit que

$$\big|d(x,A) - d(x',A)\big| \leq d(x,x').$$

La proposition est ainsi démontrée. $\square$

Grâce à la distance que nous venons de définir sur les espaces produit, nous pouvons énoncer la continuité de l'application distance.

Proposition 1.6.6. *Soit (X,d) un espace métrique. L'application $d : X \times X \to \mathbb{R}$ est continue (et même lipschitzienne).*

D'après la proposition 1.3.2, on en déduit que lorsqu'on a deux suites $(x_n)_{n\in\mathbb{N}}$ et $(y_n)_{n\in\mathbb{N}}$ convergeant respectivement vers x et y, la suite des distances $(d(x_n,y_n))_{n\in\mathbb{N}}$ converge vers la distance $d(x,y)$. Ceci nous permettra, en particulier au chapitre suivant, de passer à la limite dans les inégalités.

Démonstration. Etant donné quatres points x_0, y_0, x, y dans X, nous voulons voir que, lorsque x est proche de x_0 et y proche de y_0, la distance de x à y est proche de la distance de x_0 à y_0. L'inégalité triangulaire entraîne tout d'abord

$$d(x,y) \leq d(x_0,y_0) + d(x_0,x) + d(y_0,y).$$

Ceci s'écrit aussi

$$d(x,y) - d(x_0,y_0) \leq d(x_0,x) + d(y_0,y) = d_1((x,y),(x_0,y_0)).$$

Par symétrie, on a

$$\big|d(x,y) - d(x_0,y_0)\big| \leq= d_1((x,y),(x_0,y_0))$$

ce qui assure le résultat. $\square$

1.7 Notion d'espace topologique

Définition 1.7.1. *Soit X un ensemble. On munit X d'une structure d'espace topologique en se donnant un sous-ensemble Θ de $\mathcal{P}(X)$ l'ensemble des parties de X qui vérifie:*

- *L'ensemble X et l'ensemble vide appartiennent à Θ,*
- *Toute réunion d'éléments de Θ appartient à Θ,*
- *Toute intersection finie d'éléments de Θ appartient à Θ.*

La proposition 1.2.3 dit exactement que l'ensemble des ouverts définis par la définition 1.2.1 munissent l'espace métrique X d'une structure d'espace topologique.

Un exemple d'espace topologique non métrique Soit X un ensemble ayant au moins deux éléments. On considère $\Theta = \{\emptyset, X\}$. Il n'existe pas de distance d sur X telle que les ouverts associés à la distance d coïncide avec Θ. En effet, soit d une distance sur X et x_1 et x_2 deux points distincs de X. Les boules ouvertes $\overset{\circ}{B}(x_1, r)$ et $\overset{\circ}{B}(x_2, r)$ sont d'intersection vide dès que r est inférieur à la moitié de la distance des deux points. En effet s'il existe x appartenant à $\overset{\circ}{B}(x_1, r) \cap \overset{\circ}{B}(x_2, r)$ alors l'inégalité triangulaire implique que

$$d(x_1, x_2) \leq d(x_1, x) + d(x, x_2) \leq 2r.$$

Pour deux points dinstincts x_1 et x_2, il existe donc deux ouverts U_1 et U_2 contenant contenant respectivement x_1 et x_2 d'intersection vide (c'est la propriété dite de séparation). Or, dans l'exemple considéré ici, il n'existe qu'un seul ouvert non vide, c'est X, donc si X contient au moins deux points distincts, il ne peut pas exister deux ouverts disjoints.

Définition 1.7.2. *Soient (X, d) et (Y, δ) deux espaces métriques et f une bijection de X dans Y. On dit que f est un homéomorphisme si et seulement si la bijection f et la bijection réciproque f^{-1} sont continues en tout point.*

Remarquons que c'est une notion topologique (on pourrait la définir indépendamment de la distance en utilisant smplement une famille d'ouverts). D'après le théorème 1.3.3, cela signifie que les ouverts de Y sont exactement les images par f des ouverts de X. Intuitivement, il faut comprendre cela comme voulant dire X et Y ont la même forme.

Définition 1.7.3. *Soit (X, d) un espace métrique et x un point de X; on appelle voisinage de x tout ensemble qui contient un ouvert contenant x. Remarquons aussi que, comme toute intersection finie d'ouvers est un ouvert, toute intersection finie de voisinage de x est un voisinage de x.*

Remarques

- On peut donner une définition équivalente en disant qu'un voisinage de x est un ensemble qui admet x comme point intérieur.
- Un ouvert apparait comme un ensemble qui est un voisinage de chacun de ses points.

La définition d'une suite convergente et d'une fonction continues ainsi que de l'adhérence et de l'intérieur peut se faire dans ce cadre; ce sont des notions "topologiques".

Définition 1.7.4. *Soit (X, Θ) un espace topologique. Soient $(x_n)_{n\in\mathbb{N}}$ une suite d'éléments de X et ℓ un point de X. On dit que la suite $(x_n)_{n\in\mathbb{N}}$ converge vers ℓ si et seulement si pour tout voisinage V de ℓ, il existe un entier n_0 tel que*

$$\forall n \geq n_0\,,\ x_n \in V.$$

Observons que, lorsque X est muni d'une distance d que la famille Θ est la famille des ouverts au sens de la définition 1.2.1, cette définition coïncide avec la définition 1.1.3. En effet, soit V un voisinage de ℓ, il contient un ouvert contenant ℓ, et donc une boule ouverte $\overset{\circ}{B}(\ell, \alpha)$. La définition 1.1.3 implique l'existence d'un entier n_0 tel que

$$\forall n \geq n_0\,,\ x_n \in B(\ell, \alpha) \subset V.$$

Réciproquement, si la suite $(x_n)_{n\in\mathbb{N}}$ converge vers ℓ au sens de la définition ci-dessus, elle converge au sens de la définition 1.1.3; il suffit en effet d'appliquer la définition ci-dessus en choisissant $V = \overset{\circ}{B}(x_0, r)$.

Définition 1.7.5. *Soient (X, Θ) et $(Y, \widetilde{\Theta})$ deux espaces tologiques et f une fonction de X dans Y. Considérons un point x_0 de X. On dit que f est continue au point x_0 si et seulement si pour tout voisinage V de $f(x_0)$, il existe un voisinage U de x_0 tel que $f(U) \subset V$.*

Le lecteur vérifiera de même que lorsque X est muni d'une distance d et que la famille Θ est la famille des ouverts au sens de la définition 1.2.1, cette définition coïncide avec la définition 1.1.4.

De même, si (X, Θ) est un espace topologique et A une partie de X, on peut munir A d'une structure d'espace topologique en définissant Θ_A par

$$\Theta_A = \Big\{ U\,, /\ \exists V \in \Theta \,/\ U = V \cap A \Big\}.$$

Le lecteur vérifiera que Θ_A satisfait aux axiomes de la définition 1.7.1. On parle alors de topologie induite. La proposition 1.5.2 assure la cohérence de cette définition avec celle issue de l'approche métrique.

Enfin, la proposition suivante explicite la seconde partie de la définition 1.6.2.

Proposition 1.7.6. *Soient X un ensemble et d et δ deux distances sur X. Les distances d et δ induisent la même topologie si et seulement si*

$$\forall x\,, \forall \varepsilon\,, \exists \eta \,/\quad B_d(x, \eta) \subset B_{d'}(x, \varepsilon) \quad \text{et} \quad B_d(x, \eta) \subset B_{d'}(x, \varepsilon). \tag{1.4}$$

La démonstration de cette proposition n'est qu'une application immédiate des définitions; elle est laissée au lecteur.

L'assertion (1.4) implique que les ouverts (et donc les fermés) associés aux distances d et d' sont les mêmes.

1.8 Compléments et exercices

Ceci n'a pas été traité en amphi.

Proposition 1.8.1 (Séparation de deux fermés disjoints)**.** *Soient A et B deux parties fermées disjointes d'un espace métrique (X, d); il existe une fonction continue de (X, d) dans $[0, 1]$ telle que $f_{|A} = 0$ et $f_{|B} = 1$.*

Pour démontrer cela, il suffit de vérifier que convient la fonction f définie par

$$f(x) \stackrel{\text{déf}}{=} \frac{d(x,A)}{d(x,A)+d(x,B)}.$$

Cette fonction distance à un ensemble permet de construire des partitions de l'unité. Voici ce dont il s'agit.

Théorème 1.8.2. *Soient (X,d) un espace métrique et $(U_n)_{n\in\mathbb{N}}$ une suite d'ouverts localement finie (ce qui signifie que, pour tout point x de X, il existe un réel $\alpha > 0$ tel que $U_n \cap B(x,\alpha) = \emptyset$ sauf pour un nombre fini d'indice n) tel que*

$$\bigcup_{n\in\mathbb{N}} U_n = X.$$

Alors il existe une suite de fonctions continues $(f_n)_{n\in\mathbb{N}}$ de (X,d) dans l'intervalle $[0,1]$ telle que

$$\sum_{n\in\mathbb{N}} f_n(x) = 1 \quad \text{et} \quad f_{n|X\setminus U_n} = 0.$$

Démonstration. Posons

$$F_n \stackrel{\text{déf}}{=} U_n^c \quad \text{et} \quad \varphi_n(x) \stackrel{\text{déf}}{=} d(x,F_n).$$

Remarquons que, grâce à l'exercice 1.8.4, $\varphi_n(x) = 0$ si et seulement si $x \in F_n$. Donc φ_n est strictement positive sur U_n et nulle sur F_n. Comme la famille $(U_n)_{n\in\mathbb{N}}$ est localement finie, pour tout x de X, il existe un α strictement positif tel qu'il n'existe qu'un nombre fini d'indice n tels que φ_n ne soit pas identiquement nulle sur $B(x,\alpha)$. Donc, sur la boule $B(x,\alpha)$, la fonction

$$S(x) \stackrel{\text{déf}}{=} \sum_{n\in\mathbb{N}} \varphi_n(x)$$

est une somme finie. La fonction S est donc continue sur $B(x,\alpha)$. La formule ci-dessus définit donc une fonction continue sur X. De plus, la fonction S est strictement positive sur chaque ouvert U_n, donc partout puisque la famille d'ouverts $(U_n)_{n\in\mathbb{N}}$ recouvre X. On pose alors

$$f_n \stackrel{\text{déf}}{=} \frac{\varphi_n(x)}{S(x)}$$

et le théorème est démontré.

Exercice 1.8.3. *Soit A une partie finie d'un espace métrique (X,d). Alors $\overline{A} = A$.*

Exercice 1.8.4. *Démontrez que $\overline{A}$ est l'ensemble des points x de X tels que $d(x,A) = 0$.*

Exercice 1.8.5. *Soit A une partie non vide et bornée d'un espace métrique (X,d). Démontrez qu'il existe un plus petit réel δ tel que*

$$\forall (x,y) \in A\times A\,,\ d(x,y) \leq \delta.$$

On note $\delta(A)$ ce réel et on l'appelle le diamètre de A.

Chapitre 2

Espaces complets

2.1 Définition et exemples

Définition 2.1.1. *Soit (X,d) un espace métrique, on appelle suite de Cauchy de X toute suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de X telle que*

$$\forall \varepsilon > 0\,,\ \exists n_0 \in \mathbb{N} \ /\ \forall n \geq n_0\,,\ \forall m \geq n_0\,,\ d(x_n, x_m) < \varepsilon.$$

Comme l'indique la proposition suivante, la notion de suite de Cauchy est plus générale que celle de suite convergente.

Proposition 2.1.2. *Dans un espace métrique (X,d), toute suite convergente est de Cauchy et toute suite de Cauchy ayant une valeur d'adhérence ℓ converge vers ℓ.*

Démonstration. Considérons une suite $(x_n)_{n\in\mathbb{N}}$ convergente vers une limite ℓ. Pour tout ε strictement positif, il existe un entier n_0 tel que l'on ait

$$\forall n \geq n_0\,,\ d(x_n, \ell) < \frac{\varepsilon}{2}\cdot$$

Donc, si n et m sont deux entiers plus grands que n_0, on a

$$\begin{aligned} d(x_n, x_m) &\leq d(x_n, \ell) + d(\ell, x_m) \\ &\leq \varepsilon \end{aligned}$$

ce qui prouve qu'il s'agit d'une suite de Cauchy. □

Définition 2.1.3. *Soit (X,d) un espace métrique, on dit que cet espace est complet si et seulement si toute suite de Cauchy est convergente.*

Cete notion est extrêmement importante. Elle permet notamment de démontrer qu'une suite est convergente sans connaître a priori la limite. Elle permet ainsi de "créer", c'est-à-dire de démontrer l'existence d'objets et notamment de solutions de problèmes tels que les équations différentielles.

Donnons maintenant quelques exemples d'espaces complets. Un exemple fondamental d'espace complet: c'est l'espace $\mathbb{R}$ muni de la distance $d(x,y) = |x-y|$. Nous allons maintenant examiner comment fabriquer des espaces complets à partir d'espaces complets déjà connus. Dit autrement, cela signifie que l'on va rechercher les opérations sur les ensembles qui laissent stable la propriété d'espace complet.

Proposition 2.1.4. *Soit $(X_1, d_1), \cdots, (X_N, d_N)$ une famille de N espaces métriques complets. Alors l'espace métrique produit $(X_1 \times \cdots \times X_N, d_\infty)$ est un espace complet.*

Démonstration. Soit $(x^{(n)})_{n\in\mathbb{N}}$ une suite de Cauchy de X. Par définition de d_∞,

$$\forall \varepsilon > 0\,,\ \exists n_0 \,/\, \forall j \in \{1, \cdots N\}\,,\ \forall n \geq n_0\,,\ \forall p\,,\ d_j(x_j^{(n+p)}, y_j^{(n)}) < \varepsilon. \tag{2.1}$$

C'est-à-dire que, pour tout j, la suite $(x_j^{(n)})_{n\in\mathbb{N}}$ est une suite de Cauchy de l'espace métrique complet (X_j, d_j). Donc elle converge, notons x_j sa limite. La passage à la limite lorsque p tend vers l'infini dans (2.1) assure le résultat (on utilise ici la continuité de l'application distance, proposition 1.6.6). □

Une application $f : X \to Y$ à valeur dans un espace métrique (Y, δ) est dite *bornée* si son image $f(X)$ est bornée, c'est-à-dire s'il existe une constante $C > 0$ telle que, pour tout $x, x' \in X$, $\delta(f(x), f(x')) \leq C$. Comme le montre la proposition suivante, les espaces de fonctions à valeurs dans des espaces complets forment une classe d'exemples importants d'espaces complets.

Théorème 2.1.5. *Soient (X, d) et (Y, δ) deux espaces métriques, désignons par $\mathcal{B}(X, Y)$ (resp. $\mathcal{C}(X, Y)$) l'espace des fonctions bornées (resp. bornées et continues) de X dans Y et définissons une distance D sur $\mathcal{B}(X, Y)$ par*

$$D(f, g) = \sup_{x\in X} \delta(f(x), g(x)).$$

Si l'espace (Y, δ) est complet, alors les espaces $(\mathcal{B}(X, Y), D)$ et $(\mathcal{C}(X, Y), D)$ le sont aussi.

Démonstration. Soit $(f_n)_{n\in\mathbb{N}}$ une suite de Cauchy de $(\mathcal{B}(X, Y), D)$. Par définition, on a

$$\forall \varepsilon\,,\ \exists n_0 \,/\, \forall n \geq n_0\,,\ \forall p\,,\ \forall x \in X\,,\ \delta(f_n(x), f_{n+p}(x)) < \varepsilon. \tag{2.2}$$

En particulier, pour tout x, la suite $(f_n(x))_{n\in\mathbb{N}}$ est une suite de Cauchy de Y. Comme cet espace est complet, cette suite est convergente. Donc, pour tout x de X, il existe un élément de Y, noté $f(x)$, tel que l'on ait

$$\lim_{n\to\infty} f_n(x) = f(x).$$

Vérifions que $f \in \mathcal{B}(X, Y)$. D'après l'inégalité (2.2) appliquée par exemple avec $\varepsilon = 1$, on a

$$\forall p\,,\ \forall x \in X\,,\ \delta(f_{n_0}(x), f_{n_0+p}(x)) < 1.$$

Par passage à la limite lorsque p tend vers l'infini, on obtient

$$\forall x \in X\,,\ \delta(f_{n_0}(x), f(x)) \leq 1.$$

La fonction f_{n_0} étant bornée, la fonction f l'est également. En effet, on a

$$\begin{aligned} \delta(f(x), f(x')) &\leq \delta(f(x), f_{n_0}(x) + \delta(f_{n_0}(x), f_{n_0}(x')) + \delta(f_{n_0}(x'), f(x')) \\ &\leq 2 + \delta(f_{n_0}(x), f_{n_0}(x')). \end{aligned}$$

Il faut maintenant vérifier que la suite $(f_n)_{n\in\mathbb{N}}$ converge vers f dans l'espace $(\mathcal{B}(X, Y), D)$. Pour ce faire, passons à la limite lorsque p tend vers l'infini dans l'inégalité (2.2), ce qui donne

$$\forall \varepsilon > 0\,,\ \exists n_0 \in \mathbb{N} \,/\, \forall n \geq n_0\,,\ \forall x \in X\,,\ \delta(f_n(x), f(x)) \leq \varepsilon,$$

ce qui démontre que $(\mathcal{B}(X, Y), D)$ est complet. Pour démontrer que $(\mathcal{C}(X, Y), D)$ est complet, il suffit que démontrer qu'il est fermé grâce à la proposition suivante.

Proposition 2.1.6. *Soient (X,d) un espace complet et A une partie de X. Alors l'espace métrique $(A, d_{A\times A})$ est complet si et seulement si A est une partie fermée de X.*

Démonstration. Supposons que A soit fermée. Soit $(a_n)_{n\in\mathbb{N}}$ une suite de Cauchy d'éléments de A. C'est une suite de Cauchy d'éléments de X. Donc, puisque X est complet, elle converge vers un élément de X désigné par a_∞. Cet élément de X est limite d'une suite d'éléments de A qui est un fermé de X. D'après le critère séquentiel, a_∞ appartient à A, ce qui prouve que A est complet.

Réciproquement, supposons que A soit complet. Considérons un élément x de X tel qu'il existe une suite $(a_n)_{n\in\mathbb{N}}$ d'éléments de A dont la limite soit x. La suite $(a_n)_{n\in\mathbb{N}}$ est une suite convergente (dans X), c'est donc une suite de Cauchy dans X, donc aussi dans A. Comme A est complet, la suite $(a_n)_{n\in\mathbb{N}}$ converge dans A. L'unicité de la limite dans l'espace X entraîne que $x \in A$. D'où la proposition. □

Poursuite de la démonstration du théorème 2.1.5 Démontrons que, si une suite $(f_n)_{n\in\mathbb{N}}$ de $\mathcal{C}(X,Y)$ converge vers $f \in \mathcal{B}(X,Y)$ au sens de D, alors $f \in \mathcal{C}(X,Y)$.[1] Par définition, pour tout $\varepsilon > 0$, il existe un entier n_0 tel que

$$\forall x \in X\,,\ \ \delta(f_{n_0}(x), f(x)) < \frac{\varepsilon}{4}\,\cdot \tag{2.3}$$

L'utilisation répétée de l'inégalité triangulaire et l'inégalité (2.3) permettent d'écrire, pour tous x_0, x dans X,

$$\begin{aligned}\delta(f(x), f(x_0)) &\leq \delta(f(x), f_{n_0}(x)) + \delta(f_{n_0}(x), f_{n_0}(x_0)) + \delta(f_{n_0}(x_0), f(x_0))\\ &\leq \frac{\varepsilon}{2} + \delta(f_{n_0}(x), f_{n_0}(x_0)).\end{aligned}$$

La fonction f_{n_0} étant continue, pour tout x_0 de X, il existe $\alpha > 0$ tel que

$$\forall x \in X,\ d(x,x_0) < \alpha \Longrightarrow \delta(f_{n_0}(x), f_{n_0}(x_0)) < \frac{\varepsilon}{2}\,\cdot$$

On en déduit que

$$\forall x \in X,\ d(x,x_0) < \alpha \Longrightarrow \delta(f(x), f(x_0)) < \varepsilon,$$

ce qui achève la démonstration du théorème. □

2.2 Deux théorèmes sur les espaces complets

Lorsqu'un espace (X,d) est complet, cela permet de démontrer l'existence de certains objets. Le théorème suivant en est l'illustration la plus spectaculaire.

Théorème 2.2.1 (de point fixe de Picard)**.** *Soit f une application d'un espace métrique complet (X,d) dans lui-même telle qu'il existe un réel k de l'intervalle* $]0,1[$ *vérifiant*

$$\forall x, y \in X,\ d(f(x), f(y)) \leq k d(x,y).$$

Il existe alors un unique point z tel que $f(z) = z$.

[1] Le lecteur aura reconnu qu'il s'agit de démontrer qu'une limite uniforme de fonctions continues est une fonction continue.

Remarques

- Un point z vérifiant $f(z) = z$ comme dans la conclusion est appelé *point fixe* de f.
- Une application f vérifiant l'hypothèse "$\forall x, y \in X,\ d(f(x), f(y)) \leq kd(x, y)$" comme dans le théorème est dite *k-lipschitzienne*. Le lecteur vérifie facilement que toute application k-lipschitzienne est continue.
- La preuve ci-dessous montre de plus que toute suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de X définie par la donnée de x_0 et la relation de récurrence $x_{n+1} = f(x_n)$ converge vers l'unique point fixe de f, et donne même une estimation de la vitesse de convergence.

Démonstration. Étant donné un élément x_0 de X, on considère la suite $(x_n)_{n\in\mathbb{N}}$ définie par

$$x_{n+1} = f(x_n).$$

On peut écrire que

$$\begin{aligned} d(x_{n+1}, x_n) &= d(f(x_n), f(x_{n-1})) \\ &\leq kd(x_n, x_{n-1}). \end{aligned}$$

Par itération multiplicative, on trouve que

$$d(x_{n+1}, x_n) \leq k^n d(x_1, x_0).$$

Ainsi, pour tout couple d'entiers (n, p) , on a

$$\begin{aligned} d(x_{n+p}, x_n) &\leq \sum_{m=1}^{p} d(x_{n+m}, x_{n+m-1}) \\ &\leq d(x_1, x_0) \sum_{m=1}^{p} k^{n+m-1} \\ &\leq \frac{k^n}{1-k} d(x_1, x_0). \end{aligned}$$

Puisque $k < 1$, la suite $(k^n)_{n\in\mathbb{N}}$ tend vers 0, on en déduit que la suite $(x_n)_{n\in\mathbb{N}}$ est de Cauchy. Soit z sa limite. Comme la fonction f est lipschitzienne, donc continue, on obtient, en passant à la limite dans la relation de définition de la suite $(x_n)_{n\in\mathbb{N}}$ que $z = f(z)$.

Il nous reste à démontrer l'unicité. Soient z_1 et z_2 deux solutions de $z = f(z)$. On a, d'après l'hypothèse faite sur f que

$$d(z_1, z_2) \leq kd(z_1, z_2).$$

Le fait que k soit strictement inférieur à 1 assure que $d(z_1, z_2) = 0$, donc que $z_1 = z_2$. Le théorème est ainsi démontré. □

Remarques Ce théorème est à la base d'innombrables théorèmes d'existence et d'unicité. En voici un exemple.

Théorème 2.2.2 (de Cauchy-Lipschitz)**.** *Soit f une fonction continue sur un ouvert de U de $\mathbb{R}\times\mathbb{R}^N$ à valeurs dans $\mathbb{R}^N$. On suppose que pour un point (t_0, x_0) de U, il existe un ouvert V contenant (t_0, x_0) et une constante K tel que*

$$\forall\big((t, x), (t', y)\big) \in V^2\,,\ \|f(t, x) - f(t, y)\| \leq K\|x - y\|.\ ^2$$

[2] La fonction est dite localement lipschitzienne en (t_0, x_0)

Il existe un réel $\alpha > 0$ tel qu'il existe une unique fonction continue de $[t_0 - \alpha, t_0 + \alpha]$ à valeurs dans $\mathbb{R}^N$ telle que, pour tout t de l'intervalle $[t_0 - \alpha, t_0 + \alpha]$, et

$$\forall t \in [t_0 - \alpha, t_0 + \alpha]\,,\ x(t) = x_0 + \int_{t_0}^t f(t', x(t'))dt'.$$

Démonstration. La fonction f étant continue au point (t_0, x_0), il existe deux réels strictement positifs α et r_0 tel que

$$M(\alpha, r) \stackrel{\text{déf}}{=} \sup_{\substack{t\in[t_0-\alpha,t_0+\alpha]\\ x\in B(x_0,r_0)}} \|f(t,x)\|$$

soit fini. En remarquant que $M(\alpha, r)$ est une fonction décroissante de α (pour r fixé), on peut supposer que

$$\alpha M(\alpha, r) \leq r. \tag{2.4}$$

Soit $F_{\alpha,r}$ l'ensemble des fonctions continues de $[t_0 - \alpha, t_0 + \alpha]$ à valeurs dans $B(x_0, r)$. Cet ensemble est un fermé de l'ensemble des fonctions continues bornées à valeurs dans $\mathbb{R}^N$ pour la distance de la convergence uniforme

$$d(f, g) \stackrel{\text{déf}}{=} \sup_{t\in[t_0-\alpha,t_0+\alpha]} \|f(t) - g(t)\|.$$

(Exercice: démontrez-le!). Pour un réel strictement positif λ, on considère la distance

$$d_\lambda(f, g) \stackrel{\text{déf}}{=} \sup_{t\in[t_0-\alpha,t_0+\alpha]} e^{-\lambda|t-t_0|}\|f(t) - g(t)\|.$$

Comme on a

$$e^{-\lambda\alpha}d(f, g) \leq d_\lambda(f, g) \leq d(f, g),$$

les deux distance sont équivalentes. D'après le théorème 2.1.5 et la proposition 2.1.6, $(F_{\alpha,r}, d_\lambda)$ est un espace métrique complet. Si x appartient à $F_{\alpha,r}$, alors on a

$$\Big\|\int_{t_0}^t f(t', x(t'))dt'\Big\| \leq \Big|\int_{t_0}^t M dt'\Big| \leq \alpha M(\alpha, r).$$

D'après la condition (2.4), il en résulte que la formule

$$F(x)(t) \stackrel{\text{déf}}{=} x_0 + \int_{t_0}^t f(t', x(t'))dt'$$

définit une application de $F(\alpha, r)$ dans lui même. De plus, si x et y sont deux fonctions de $F(\alpha, r)$, on a, pour tout t de l'intervalle $[t_0 - \alpha, t_0 + \alpha]$,

$$\begin{aligned}\|F(x)(t) - F(y)(t)\| &\leq \Big|\int_{t_0}^t f(t', x(t'))dt' - \int_{t_0}^t f(t', y(t'))dt'\Big| \\ &\leq K\Big|\int_{t_0}^t |x(t') - y(t')|dt'\Big|.\end{aligned}$$

Par définition de d_λ, ceci implique que

$$d_\lambda(F(x), F(y)) \leq A\alpha \sup_{t\in[t_0-\alpha,t_0+\alpha]} \Big|\int_{t_0}^t e^{-\lambda[t'-t_0|}dt'\Big|\, d_\lambda(x, y).$$

Un calcul simple, en distinguant le cas où t est plus grand que t_0 de celui où t est plus petit que t_0, assure que

$$d_\lambda(F(x), F(y)) \leq \frac{A}{\lambda} d_\lambda(x, y).$$

En choisissant λ tel que $A\lambda < 1$, on peut appliquer le théorème de point fixe de Picard pour obtenir le résultat. $\square$

Nous allons maintenant y démontrer un théorème profond qui implique des propriétés que l'on serait souvent bien en peine de démontrer autrement.

Théorème 2.2.3 (de Baire)**.** *Soit (X, d) un espace métrique complet, on considère une suite $(U_n)_{n\in\mathbb{N}}$ d'ouverts denses dans X. Alors $\displaystyle\bigcap_{n\in\mathbb{N}} U_n$ est dense.*

Démonstration. Elle repose sur le lemme suivant qui a son intérêt propre.

Lemme 2.2.4. *Soit (X, d) un espace métrique complet et $(F_n)_{n\in\mathbb{N}}$ une suite décroissante de fermés non vides de X dont le diamètre tend vers 0. Il existe alors un élément x de X tel que*

$$\bigcap_{n\in\mathbb{N}} F_n = \{x\}.$$

Démonstration. On choisit dans chaque F_n un élément que l'on désignera par x_n. Comme la suite $(F_n)_{n\in\mathbb{N}}$ est décroissante, pour tout entier p, on sait que x_{n+p} appartient à F_n, donc que $d(x_n, x_{n+p}) \leq \delta(F_n)$. Donc la suite $(x_n)_{n\in\mathbb{N}}$ converge vers un élément x de X qui appartient bien sûr à l'intersection des F_n. Soit y un élément de cette intersection. On a alors

$$\forall n\,,\ d(x, y) \leq \delta(F_n)$$

et donc $d(x, y) = 0$; d'où le lemme. $\square$

Suite de la démonstration du théorème 2.2.3 Soit V un ouvert quelconque de X, nous allons démontrer que $\displaystyle\bigcap_n U_n \cap V \neq \emptyset$, ce qui assurera le théorème.

L'ouvert U_0 est dense, donc $U_0 \cap V$ est un ouvert non vide. Donc il existe un réel strictement positif α_0 (que l'on peut supposer inférieur à 1) et un point x_0 de X tels que

$$B_f(x_0, \alpha_0) \subset U_0 \cap V. \tag{2.5}$$

L'ouvert U_1 est dense, donc l'ensemble $U_1 \cap B(x_0, \alpha_0)$ est un ouvert non vide. Donc il existe un réel strictement positif α_1 (que l'on peut supposer inférieur à $1/2$) et un point x_1 de X tels que

$$B_f(x_1, \alpha_1) \subset U_1 \cap B(x_0, \alpha_0).$$

Nous allons procéder par récurrence et supposer construite une suite $(x_j)_{0\leq j\leq n}$ d'éléments de X et une suite $(\alpha_j)_{0\leq j\leq n}$ telles que, pour tout $j \leq n$, on ait

$$\alpha_j \leq \frac{1}{j+1} \quad \text{et} \quad B_f(x_j, \alpha_j) \subset U_j \cap B(x_{j-1}, \alpha_{j-1}). \tag{2.6}$$

L'ouvert U_{n+1} est dense, donc l'ouvert $U_{n+1} \cap B(x_n, \alpha_n)$ est non vide. Il existe donc un réel strictement positif α_{n+1} (que l'on peut supposer inférieur à $1/(n+2)$) et un point x_{n+1} de X telles que

$$B_f(x_{n+1}, \alpha_{n+1}) \subset B(x_n, \alpha_n) \cap U_{n+1}.$$

On a ainsi construit une suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de X et une suite $(\alpha_n)_{n\in\mathbb{N}}$ de réels strictement positifs qui tend vers 0 et telle que

$$\forall m \geq n\,,\ x_m \in B(x_n, \alpha_n). \tag{2.7}$$

Comme la suite $(\alpha_n)_{n\in\mathbb{N}}$ tend vers 0, la suite $(x_n)_{n\in\mathbb{N}}$ est de Cauchy. Donc elle converge dans X qui est complet. Appelons x sa limite. D'après la relation (2.7), le point x appartient à $B_f(x_n, \alpha_n)$ pour tout entier n. Vu les relations (2.5) et (2.6), on a

$$\forall n \in \mathbb{N}\,,\ x \in V \cap \bigcap_{j\leq n} U_j.$$

Le théorème de Baire est ainsi démontré. □

On utilise souvent l'énoncé suivant qui n'est rien d'autre que le théorème de Baire "passé au complémentaire".

Théorème 2.2.5. *Soit (X,d) un espace métrique complet, on considère une suite $(F_n)_{n\in\mathbb{N}}$ de fermés d'intérieur vide dans X. Alors $\bigcup_{n\in\mathbb{N}} F_n$ est d'intérieur vide.*

Nous utiliserons souvent le théorème de Baire sous la forme de ce corollaire, dont la démonstration, très facile, est laissée en exercice.

Corollaire 2.2.6. *Soit $(F_n)_{n\in\mathbb{N}}$ une suite de fermés d'un espace métrique complet (X,d) dont la réunion est X. Alors il existe un entier n_0 tel que $\overset{\circ}{F}_{n_0} \neq \emptyset$.*

Dit autrement, un espace métrique complet n'est pas réunion dénombrable de fermés d'intérieur vide, ce qui est le cas d'un espace notoirement non complet, l'espace $\mathbb{Q}$.

Proposition 2.2.7. *Toute suite de Cauchy d'un espace métrique est bornée.*

Démonstration. En effet, si la suite $(x_n)_{n\in\mathbb{N}}$ est de Cauchy, il existe un entier n_0 tel que, pour tout $n \geq n_0$ et tout $m \geq n_0$, on ait $d(x_n, x_m) \leq 1$. Ainsi donc, on a

$$\sup_{(n,m)\in\mathbb{N}^2} d(x_n, x_m) \leq 1 + \max_{(n,m)\in\{0,\dots,n_0\}^2} d(x_n, x_m).$$

D'où la proposition. □

Pour conclure cette section sur les espaces complets, nous allons examiner l'invariance de la propriété d'être complet par changement de distance.

Proposition 2.2.8. *Soient d et δ deux distances équivalentes sur X. Toute suite de Cauchy pour l'une est une suite de Cauchy pour l'autre.*

Démonstration. Par définition il existe $k \geq 1$ tel que les deux inégalités

$$(1)\quad d(x,y) \leq k\delta(x,y), \quad (2)\quad \delta(x,y) \leq kd(x,y)$$

pour tout x, y dans X. Considérons une suite $(x_n)_{n\in\mathbb{N}}$ de Cauchy pour la distance d. Pour tout réel strictement positif ε, il existe un entier n_0 tel que

$$\forall n \geq n_0\,,\ \forall m \geq n_0\,,\ d(x_n, x_m) < \frac{\varepsilon}{k}\,;$$

on en déduit à l'aide de la deuxième inégalité que

$$\forall n \geq n_0\,,\ \forall m \geq n_0\,,\ \delta(x_n, x_m) < \varepsilon$$

autrement dit la suite est de Cauchy pour δ. □

Chapitre 3

La notion de compacité

3.1 Définition et propriétés de base

Définition 3.1.1. *Soit (X,d) un espace métrique. On dit que (X,d) est un espace compact si et seulement si toute suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de X possède une valeur d'adhérence.*

Un exemple fondamental d'espaces compacts est donné par le théorème suivant.

Théorème 3.1.2 (de Bolzano-Weierstrass)**.** *Les intervalles $[a,b]$ de $\mathbb{R}$ munis de la distance usuelle sont des espaces compacts.*

Démonstration. Soit $(x_n)_{n\in\mathbb{N}}$ une suite de l'intervalle $[a,b]$. L'ensemble $A_n \stackrel{\text{déf}}{=} \{x_m\,,\ m\geq n\}$ est une partie majorée de réels qui admet donc une borne supérieure que l'on note L_n. Comme la suite A_n est une suite décroissante au sens de l'inclusion, la suite $(L_n)_{n\in\mathbb{N}}$ est une suite décroissante qui est minorée par a. La suite $(L_n)_{n\in\mathbb{N}}$ est donc une suite convergente. Soit L sa limite. Démontrons que L est une valeur d'adhérence de la suite $(x_n)_{n\in\mathbb{N}}$.

Soit ε un réel strictement positif et m un entier positif. Il existe un entier n_0 tel que

$$\forall \varepsilon>0/\,,\forall n\geq n_0\,,\ 0\leq L_n-L<\frac{\varepsilon}{2}\cdot$$

Soitn $n_1 \stackrel{\text{déf}}{=} \max\{m,n_0\}$. Comme L_{n_1} et le plus petit des majorants de A_{n_1}, il existe un entier n tel que

$$|x_n-L_{n_1}|<\frac{\varepsilon}{2}\cdot$$

Donc, pour tout réel strictement positif ε et tout entier positif m, il existe un entier n supérieur ou égal à m tel que

$$|x_n-L|<\varepsilon.$$

D'où le théorème. □

L'idée intuitive qui soustend la définition d'un espace compact est qu'il n'est pas trop "gros". Le lemme suivant qui sera utile dans la suite illustre cette idée.

Lemme 3.1.3. *Soit (X,d) un espace métrique compact. Alors, pour tout réel strictement positif α, on peut recouvrir X par un nombre fini de boules de rayon α.*

Démonstration. Procédons par contraposition. Supposons qu'il existe un réel strictement positif α tel que l'on ne puisse recouvrir X par un nombre fini de boules de rayon α. Soit x_0 un élément quelconque de X. Il existe un élément x_1 de X n'appartenant pas à $B(x_0,\alpha)$.

Soit $(x_0, \cdots, x_p)$ un p-uplet d'éléments de X tel que, pour tout $m \neq n$, on ait $d(x_m, x_n) \geq \alpha$. Par hypothèse,

$$\bigcup_{n=1}^{p} B(x_n, \alpha) \neq X.$$

Donc, il existe un point x_{p+1} de X qui n'appartient pas à la réunion de boules ci-dessus. Par récurrence, nous construisons ainsi une suite $(x_n)_{n\in\mathbb{N}}$ telle que

$$m \neq n \Rightarrow d(x_m, x_n) \geq \alpha.$$

Une telle suite n'a bien sûr pas de valeur d'adhérence (exercice, vérifiez-le!). D'où le lemme. □

La proposition suivante permet de trouver beaucoup d'espaces compacts.

Proposition 3.1.4. *Soit $(X_1, d_1), \cdots, (X_N, d_N)$ une famille de N espaces métriques compacts. Posons*

$$X = X_1 \times \cdots \times X_N \quad \textit{et} \quad d_\infty((x_1, \cdots, x_N), (y_1, \cdots, y_N)) = \max_{1\leq j\leq N} d_j(x_j, y_j).$$

Alors l'espace métrique (X, d_∞) est compact.

Démonstration. Soit $(x_n)_{n\in\mathbb{N}}$ une suite d'éléments de X. Par définition, il existe pour chaque j une suite $(x_n^j)_{n\in\mathbb{N}}$ de X_j telle que $x_n = (x_n^1, \cdots, x_n^N)$. Par définition d'une partie compacte, il existe une fonction d'extraction ϕ_1 et un point a_1 de X_1 tel que

$$\lim x^1_{\phi_1(n)} = a_1.$$

De même, il existe une fonction d'extraction ϕ_2 et un point a_2 de X_2 tel que

$$\lim x^2_{\phi_1\circ\phi_2(n)} = a_2.$$

En itérant N fois le processus, on construit une fonction d'extraction ϕ en posant

$$\phi \stackrel{\text{déf}}{=} \phi_1 \circ \cdots \circ \phi_N$$

telle que, pour tout $j \in \{1, \cdots, N\}$, on ait

$$\lim_{n\to\infty} x^j_{\phi(n)} = a_j.$$

Par définition de la distance sur X, cela implique que la suite $x_{\phi(n)}$ converge vers $(a_1, \cdots, a_N)$. D'où le résultat. □

Proposition 3.1.5. *Soit (X, d) un espace métrique compact, il est alors complet.*

Démonstration. En effet, soit $(x_n)_{n\in\mathbb{N}}$ une suite de Cauchy de X. Comme l'espace X est compact, cette suite admet une valeur d'adhérence ℓ. Démontrons maintenant qu'une telle suite est convergente.

La suite étant de Cauchy, il existe un entier n_0 tel que

$$\forall n \geq n_0\,,\ \forall m \geq n_0\,,\ d(x_n, x_m) \leq \frac{\varepsilon}{2}\cdot$$

La suite ayant ℓ pour valeur d'adhérence, il existe un entier $n_1 \geq n_0$ tel que

$$d(x_{n_1}, \ell) < \frac{\varepsilon}{2}\cdot$$

On en déduit alors que, pour tout entier $n \geq n_0$, on a

$$d(x_n, \ell) \leq d(x_n, x_{n_1}) + d(x_{n_1}, \ell) < \varepsilon.$$

D'où la proposition. □

La réciproque de cette proposition est bien sûr fausse. Par exemple, $\mathbb{R}$ muni de la distance $d(x,y) = |x-y|$ est un espace métrique complet qui n'est pas compact. Cependant, dans le cas des espaces métriques complets, on peut trouver une caractérisation des ensembles compacts uniquement en termes de boules.

3.2 Caractérisation des espaces compacts en termes d'ouverts et de fermés

Le théorème suivant donne une caractérisation de la compacité en terme de recouvrements par des ouverts. Bien qu'il soit dans le cadre métrique, le lemme 3.1.3 peut être compris comme un premier pas dans cette direction. Il est d'ailleurs utile pour démontrer le théorème ci-dessous.

Théorème 3.2.1. *Soit (X,d) un espace métrique, les trois propriétés suivantes sont équivalentes.*

i) Pour toute famille $(U_\lambda)_{\lambda\in\Lambda}$ d'ouverts de X recouvrant X, c'est-à-dire telle que

$$X = \bigcup_{\lambda\in\Lambda} U_\lambda,$$

on peut extraire un sous-recouvrement fini, c'est-à-dire qu'il existe une famille finie $(\lambda_j)_{1\leq j\leq N}$ telle que

$$X = \bigcup_{j=1}^{N} U_{\lambda_j}.$$

ii) Pour toute famille de fermés $(F_\lambda)_{\lambda\in\Lambda}$ on a

$$\forall N\,,\ \forall(\lambda_1\cdots,\lambda_N)\in\Lambda^N \,/\, \bigcap_{j=1}^{N} F_{\lambda_j} \neq \emptyset \Longrightarrow \bigcap_{\lambda\in\Lambda} F_\lambda \neq \emptyset.$$

iii) Toute suite d'éléments de X admet une valeur d'adhérence.

Démonstration. L'équivalence entre les points i) et ii) se démontrent par contraposition. En effet, par contraposition le point ii) est équivalent à :

Pour toute famille de fermés $(F_\lambda)_{\lambda\in\Lambda}$,

$$\bigcap_{\lambda\in\Lambda} F_\lambda = \emptyset \Longrightarrow \exists(\lambda_1\cdots,\lambda_N)\in\Lambda^N \,/\, \bigcap_{j=1}^{N} F_{\lambda_j} = \emptyset.$$

Par passage au complémentaire, ceci est équivalent au point i).

Démontrons que ii) implique iii). Par définition de l'ensemble des valeurs d'adhérence, on a

$$\mathrm{Adh}(x_n) = \bigcap_n \overline{A}_n \quad \text{avec} \quad A_n \stackrel{\text{déf}}{=} \{x_m\,, m\geq n\}.$$

La suite $\overline{A}_n$ est une suite décroissante de fermés non vides. Donc, d'après la propriéte ii), l'intersection de tous les $\overline{A}_n$ (qui est l'ensemble des valeurs d'adhérence) est non vide. D'où le point iii).

Démontrons maintenant que *iii*) implique *i*) ce qui constitue le point délicat de la démonstration. Il s'agit de déduire d'une propriété sur les suites, une propriété sur les recouvrements d'ouverts. Le lemme suivant est crucial, et peut aussi être utile utile pour démontrer d'autres propriétés des espaces compacts.

Lemme 3.2.2 (de Lebesgue)**.** *Soit (X,d) un espace métrique compact. Pour toute famille $(U_\lambda)_{\lambda\in\Lambda}$ d'ouverts recouvrant X, il existe un nombre réel α strictement positif tel que*

$$\forall x \in X\,,\ \exists \lambda \in \Lambda \,/\, B(x,\alpha) \subset U_\lambda.$$

Démonstration. Pour démontrer ce lemme, considérons la fonction δ définie par

$$\delta \left\{ \begin{array}{ccl} X & \to & \mathbb{R}_+^\star \\ x & \mapsto & \sup\{\beta \,/\, \exists \lambda \,/\, B(x,\beta) \subset U_\lambda\}. \end{array} \right.$$

Démontrer que la fonction δ possède un minimum implique le lemme. Pour démontrer cela, établissons tout d'abord l'assertion suivante.

$$\forall \varepsilon\,,\ \forall x \in X\,,\ \exists \beta \,/\, d(x,y) < \beta \Longrightarrow \delta(y) > \delta(x) - \varepsilon. \tag{3.1}$$

Soient x un point de X et ε un réel strictement positif inférieur à $\delta(x)$. Supposons que y soit un point de X tel que $d(x,y) < \varepsilon/2$. L'inégalité triangulaire implique alors que

$$B(y, \delta(x) - \varepsilon) \subset B(x, \delta(x) - \varepsilon/2) \subset U_\lambda.$$

Et donc l'assertion (3.1) est démontrée.

Considérons maintenant une suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de X telle que $\lim\limits_{n\to\infty} \delta(x_n) = \inf\limits_{x\in X} \delta(x)$. Par hypothèse, on peut extraire de la suite $(x_n)_{n\in\mathbb{N}}$ une sous-suite convergeant vers un point x_∞ de X. On note toujours $(x_n)_{n\in\mathbb{N}}$ cette suite extraite. D'après l'assertion (3.1), pour tout $\varepsilon > 0$, il existe un entier n_0 tel que

$$n \geq n_0 \Longrightarrow \delta(x_n) > \delta(x_\infty) - \varepsilon.$$

Par passage à la limite, on en déduit que

$$\inf_{x\in X} \delta(x) \geq \delta(x_\infty) - \varepsilon,$$

et ce pour tout réel strictement positif ε. Il en résulte que $\delta(x_\infty) = \inf\limits_{x\in X} \delta(x)$ et donc que $\inf\limits_{x\in X} \delta(x)$ est strictement positif. Ceci conclut la démonstration du lemme. □

Conclusion de la démonstration du théorème 3.2.1

Pour montrer que iii) implique i), nous supposons que X est compact (de toute suite on peut extraire une suite convergente), et nous considérons une famille $(U_\lambda)_{\lambda\in\Lambda}$ d'ouverts de X recouvrant X, dont il s'agit d'extraire un sous-recouvrement fini. Soit $\alpha > 0$ donné par le lemme de Lebesgue. D'après le lemme 3.1.3, il existe un nombre fini de boules $B_1, \ldots, B_N$ de rayon α qui recouvrent X. D'après la propriété de α donnée par le lemme de Lebesgue, chacune de ces boules B_j est incluse dans l'un des ouvert U_{λ_j} de la famille, et par conséquent

$$X \subset \bigcup_{j=1}^N B_j \subset \bigcup_{j=1}^N U_{\lambda_j},$$

ce que l'on voulait. □

Nous allons donner une réciproque du lemme 3.1.3 dans le cadre des espaces complets.

Théorème 3.2.3. *Soit (X, d) un espace métrique complet. Si, pour tout réel strictement positif α, on peut recouvrir X par un nombre fini de boules ouvertes de rayon α, alors (X, d) est un espace compact.*

Démonstration. Considérons une suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de X. Par hypothèse on peut recouvrir X par un nombre fini de boules de rayon 1. D'après le principe des tiroirs, l'une de ces boule contient une infinité de termes de la suite ; notons B_1 cette boule, et P_1 l'ensemble infini des entiers n tels que $x_n \in B_1$. On peut également recouvrir X par un nombre fini de boules de rayon $1/2$. À nouveau, l'une de ces boules, notons-là B_2, contient une infinité de termes x_n parmi les n appartenant à P_1 ; notons P_2 l'ensemble infini des éléments n de P_1 tels que $x_n \in B_2$.

De cette façon on construit des ensembles infini d'entiers $P_1 \subset P_2 \subset ...$ tels que, pour tout i, les termes x_n avec $n \in P_i$ appartiennent à une boule de rayon $1/i$. Puisque ces ensembles sont tous infinis, on peut choisir une extraction ϕ telle que, pour tout n, $\phi(n) \in P_n$. Montrons que la suite extraite $(x_{\phi(n)})_{n\in\mathbb{N}}$ converge. Etant donné $\varepsilon > 0$ on choisit un entier $n_0 > 2\varepsilon^{-1}$. Pour tout $n, m \geq n_0$, les termes $x_{\phi(n)}, x_{\phi(m)}$ appartiennent à une même boule de rayon $1/n_0$, on a donc

$$d(x_{\phi(n)}), x_{\phi(m)})) < \frac{2}{n_0} < \varepsilon.$$

La suite extraite est de Cauchy, elle converge puisque X est complet. Ceci prouve que X est compact. □

Remarque La complétude de l'espace métrique (X, d) est essentielle. Considérons $X = [0, 1] \cup \mathbb{Q}$ muni de la distance usuelle. Il est clair que X peut être recouvert par un nombre fini de boules de rayon arbitraire mais que X n'est pas compact.

3.3 Parties compactes d'un espace métrique

Définition 3.3.1. *Soit (X, d) un espace métrique et A une partie de X. On dit que A est une partie compacte de X si et seulement si le sous-espace métrique $(A, d_{A\times A})$ est un espace compact.*

Contrairement à la propriété d'être ouvert ou fermé dans X, la compacité de A est une propriété intrinsèque, elle ne dépend que de la restriction de d à $A \times A$.

Proposition 3.3.2. *Soit A une partie d'un espace métrique (X, d). Si A est compacte, alors A est fermée.*

Démonstration. Soit x un point de $\overline{A}$. D'après la proposition 1.2.10, il existe une suite $(a_n)_{n\in\mathbb{N}}$ d'éléments de A qui converge vers x. Comme A est supposée compacte, il existe une fonction d'extraction ϕ et un point ℓ de A tel que $\lim_{n\to\infty} a_{\phi(n)} = \ell$. Comme la suite $(a_{\phi(n)})_{n\in\mathbb{N}}$ converge vers x, l'unicité de la limite implique que $\ell = x$ et donc que $x \in A$. Donc A est une partie fermée de X. □

Exercice 3.3.3. *Soit A une partie d'un espace métrique (X, d). Le fermé $\overline{A}$ est compact si et seulement si toute suite d'éléments de A possède une valeur d'adhérence dans X.*

Proposition 3.3.4. *Soit A une partie fermée d'un espace métrique compact (X, d). Alors A est aussi compact.*

Démonstration. Soit $(a_n)_{n\in\mathbb{N}}$ une suite d'éléments de A. Comme X est compact, il existe une fonction d'extraction ϕ telle que $(a_{\phi(n)})_{n\in\mathbb{N}}$ converge vers $\ell \in X$. Comme A est un fermé de X, $\ell \in A$. Donc A est compacte. $\square$

Le théorème 3.1.2 peut s'énoncer ainsi.

Théorème 3.3.5 (de Bolzano-Weierstrass)**.** *Les intervalles fermés* $[a, b]$ *sont des parties compactes de* $(\mathbb{R}, |x - y|)$.

On en déduit, grâce au théoreme 3.1.2 et à la proposition 3.1.4, le théorème suivant.

Théorème 3.3.6. *Les parties compactes de* $(\mathbb{R}^N, d_\infty)$ *sont les parties fermées de* $\mathbb{R}^N$ *incluses dans un pavé* $[a, b]^N$.

Nous allons maintenant étudier l'action des fonctions continues sur les parties compactes d'un espace métrique. Soient (X, d) et (Y, δ) deux espaces métriques. Une application $f : X \to Y$ est *uniformément continue* si

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

Théorème 3.3.7 (de Heine)**.** *Soient* (X, d) *et* (Y, δ) *deux espaces métriques et* f *une application continue de* X *dans* Y*. Alors, pour toute partie compacte* A *de* X, $f(A)$ *est une partie compacte de* Y*. De plus, si* X *est compact, alors* f *est une application uniformément continue de* X *dans* Y*.*

Démonstration. Soit $(y_n)_{n\in\mathbb{N}}$ une suite d'éléments de $f(A)$. Par définition, il existe une suite $(a_n)_{n\in\mathbb{N}}$ d'éléments de A telle que $y_n = f(a_n)$. Comme A est une partie compacte de X, il existe un point a' de A et une fonction d'extraction ϕ telle que

$$\lim_{n\to\infty} a_{\phi(n)} = a'$$

La fonction f étant continue, on a

$$\lim_{n\to\infty} y_{\phi(n)} = f(a') \in f(A).$$

Ceci démontre le premier point du théorème.

Pour démontrer le second point, observons que, par définition d'une fonction continue, pour tout $\varepsilon > 0$, pour tout $x \in X$, il existe α_x telle que l'on ait

$$d(x, x') < \alpha_x \Rightarrow \delta(f(x), f(x')) < \frac{\varepsilon}{2}\cdot$$

L'ensemble des boules $B(x, \alpha_x/2)$ forme un recouvrement de X par des ouverts. Par compacité (théorème 3.2.1), on en extrait un sous-recouvrement fini $(B(x_j, \alpha_{x_j}/2))_{1\leq j\leq N}$. On pose alors

$$\alpha = \inf_{1\leq j\leq N} \alpha_{x_j}.$$

Soit $(x, x') \in X^2$ tel que $d(x, x') < \alpha/2$. Comme la famille $(B(x_j, \alpha_{x_j}/2))_{1\leq j\leq N}$ recouvre X, il existe un indice $j \in \{1, \cdots, N\}$ tel que x appartienne à la boule $B(x_j, \alpha_{x_j}/2)$, et l'inégalité triangulaire antraîne $B(x, \alpha/2) \subset B(x_j, \alpha_{x_j})$. Mais alors, le point x' appartient à $B(x_j, \alpha_{x_j})$ et l'on a ainsi

$$\begin{aligned}\delta(f(x), f(x')) &\leq \delta(f(x), f(x_j)) + \delta(f(x_j), f(x'))\\ &< \varepsilon.\end{aligned}$$

D'où le théorème. $\square$

Exercice 3.3.8. *Démontrez ce théorème en utilisant le lemme 3.2.2 de Lebesgue.*

Des théorèmes 3.3.6 et 3.3.7, on déduit aisément le corollaire suivant.

Corollaire 3.3.9. *Soit f une fonction continue de (X,d) dans $(\mathbb{R},|x-y|)$. Si (X,d) est compact, alors la fonction f admet un minimum et un maximum, c'est-à-dire qu'il existe x_m et x_M dans X tels que*

$$\forall x \in X\,,\ f(x_m) \leq f(x) \leq f(x_M).$$

3.4 Compacité des espaces de fonctions

Cette section est hors programme. On s'y intéresse à décrire les parties compacts de l'espces des fonctions continues d'un espace métrique compact (X,d) à valeurs dans un espace mt́rique complet (Y,δ).

Tout d'abord, observons que la suite de fonctions $(f_n)_{n\in\mathbb{N}}$ de $[0,1]$ dans $[0,1]$ définies par $f_n(x)=x^n$ n'admet pas de valeurs d'adhérence. En effet, la seule limite possible de toute suite extraite de la suite $(f_n)_{n\in\mathbb{N}}$ est la fonction qui vaut 0 sur $[0,1[$ et 1 en 1 qui n'est pas continue.

Le théorème qui décrit les compacts est le suivant.

Théorème 3.4.1 (d'Ascoli)**.** *Soient (X,d) et (Y,δ) deux espaces métriques. Supposons que (X,d) soit compact et (Y,δ) complet. On considère une partie A de $\mathcal{C}(X,Y)$ l'ensemble des fonctions continues de X dans Y muni de la distance*

$$D(f,g) = \sup_{x\in X} \delta(f(x),g(x)).$$

Faisons les deux hypothèses suivantes :

i) La partie A est équicontinue i.e.

$$\forall x \in X\,,\ \forall \varepsilon > 0\,,\ \exists \alpha > 0 \,/\, d(x,x') < \alpha \Rightarrow \forall f \in A\,,\ \delta(f(x),f(x')) < \varepsilon\,;$$

ii) pour tout $x \in X$, l'ensemble $\{f(x)\,,\ f\in A\}$ est d'adhérence compacte.

Alors A est d'adhérence compacte.

Démonstration. Pour démontrer ce théorème, nous allons tout d'abord observer que la première hypothèse se traduit par

$$\forall \varepsilon > 0\,,\ \exists \alpha > 0 \,/\, \forall x \in X\,,\ \forall f \in A\,,\ d(x,x') < \alpha \Rightarrow \delta(f(x),f(x')) < \varepsilon\,; \tag{3.2}$$

La démonstration de cette assertion est exactement la même que celle du théorème 3.3.7 ci-dessus (exercice : écrivez-la!). Nous allons maintenant énoncer un lemme montrant que les parties A qui sont uniformément équicontinues (c'est-à-dire qui vérifient l'assertion (3.2) ci-dessus) ont une propriété très spéciale vis-à-vis de la convergence.

Lemme 3.4.2. *Soient A une partie uniformément équicontinue de $\mathcal{C}(X,Y)$ et $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions de A. On a l'équivalence suivante*

$$\forall x \in X\,,\ \lim_{n\to\infty} f_n(x) = f(x) \Longleftrightarrow f \in \mathcal{C}(X,Y) \quad \text{et} \quad \lim_{n\to\infty} D(f_n,f) = 0.$$

Démonstration. Il n'y a quelque chose à démontrer que de gauche à droite. Soit ε un réel strictement positif arbitraire. Par hypothèse, il existe un réel α strictement positif tel que

$$d(x,x') < \alpha \Rightarrow \forall n\,,\ \delta(f_n(x),f_n(x')) < \frac{\varepsilon}{4}\cdot$$

Par passage à la limite, on obtient immédiatement que la fonction f est uniformément continue de X dans Y. On recouvre le compact X par une famille finie de boules $(B(x_j,\alpha))_{1\leq j\leq N_\varepsilon}$. Ainsi donc, on a

$$\begin{aligned} \delta(f_n(x),f(x)) &\leq \delta(f_n(x),f_n(x_j)) + \delta(f_n(x_j),f(x_j)) + \delta(f(x_j),f(x)) \\ &\leq \frac{\varepsilon}{2} + \max_{1\leq j\leq N_\varepsilon} \delta(f_n(x_j),f(x_j)). \end{aligned}$$

Par hypothèse, il existe n_0 tel que

$$\forall n \geq n_0\,,\ \max_{1\leq j\leq N_\varepsilon} \delta(f_n(x_j), f(x_j)) < \frac{\varepsilon}{2}\cdot$$

D'où le lemme. □

Suite de la démonstration du théorème 3.4.1 Considérons une suite $(f_n)_{n\in\mathbb{N}}$ d'éléments de A. Nous allons en extraire une sous-suite convergente.

Nous démontrerons très bientôt que dans tout espace compact, il existe une partie dénombrable dense (prendre un recouvrement fini de boules de rayon n^{-1} et l'ensemble de tous les centres forment une partie dense) que l'on note $\mathcal{X} \stackrel{\text{déf}}{=} \{x_p, p \in \mathbb{N}\}$. Par hypothèse, l'ensemble

$$\{f(x_0)\,,\ f \in A\}$$

est d'adhérence compacte. Donc il existe une fonction φ_0 strictement croissante de $\mathbb{N}$ dans $\mathbb{N}$ telle que

$$\lim_{n\to\infty} f_{\varphi_0(n)}(x_0) = \ell(x_0).$$

De même, il existe un point $\ell(x_1)$ de Y et une fonction φ_1 strictement croissante de $\mathbb{N}$ dans $\mathbb{N}$ telle que

$$\lim_{n\to\infty} f_{\varphi_0\circ\varphi_1(n)}(x_1) = \ell(x_1).$$

On définit ainsi par récurrence une suite $(\varphi_p)_{p\in\mathbb{N}}$ de fonctions strictement croissantes de $\mathbb{N}$ dans $\mathbb{N}$ telle que

$$\forall p \in \mathbb{N}\,,\ \forall k \leq p\,,\ \lim_{n\to\infty} f_{\varphi_0\circ\cdots\circ\varphi_p(n)}(x_k) = \ell(x_k).$$

On définit alors la fonction ψ de $\mathbb{N}$ dans $\mathbb{N}$ par

$$\psi(n) = \varphi_0 \circ \cdots \circ \varphi_n(n).$$

C'est une fonction strictement croissante de $\mathbb{N}$ dans $\mathbb{N}$. En effet, observons tout d'abord que toute fonction μ strictement croissante de $\mathbb{N}$ dans $\mathbb{N}$ vérifie $\mu(n) \geq n$. Donc, si $n < m$, on a

$$\psi(n) = \varphi_0 \circ \cdots \circ \varphi_n(n) < \varphi_0 \circ \cdots \circ \varphi_n \circ \varphi_{n+1} \circ \varphi_m(m) = \psi(m).$$

Par construction, la suite extraite $(f_{\psi(n)})_{n\in\mathbb{N}}$ vérifie

$$\forall p \in \mathbb{N}\,,\ \lim_{n\to\infty} f_{\psi(n)}(x_p) = \ell(x_p). \tag{3.3}$$

Démontrons maintenant que la fonction ℓ est uniformément continue sur $\mathcal{X}$. Soit ε un réel strictement positif, on considère un réel α vérifiant l'assertion (3.2). Pour tout couple d'entiers (p,q) tel que $d(x_p, x_q) < \alpha$, on a

$$\begin{aligned} \delta(\ell(x_p), \ell(x_q)) &\leq \delta(\ell(x_p), f_{\psi(n)}(x_p)) + \delta(f_{\psi(n)}(x_p), f_{\psi(n)}(x_q)) + \delta(f_{\psi(n)}(x_q), \ell(x_q)) \\ &\leq \varepsilon + \delta(\ell(x_p), f_{\psi(n)}(x_p)) + \delta(f_{\psi(n)}(x_q), \ell(x_q)). \end{aligned}$$

L'inégalité ci-dessus étant vraie pour tout n, on obtient, en passant à la limite $n \to \infty$,

$$d(x_p, x_q) < \alpha \Rightarrow d(\ell(x_p), \ell(x_q)) \leq \varepsilon.$$

La fonction ℓ étant uniformément continue sur la partie dense $\mathcal{X}$. Nous avons besoin du résultat suivant qui affirme que l'on peut prolonger ℓ en une fonction uniformément continue sur X tout entier.

Théorème 3.4.3 (de prolongement des fonctions uniformément continues)**.** *Soient (X,d) et (Y,δ) deux espaces métriques, A une partie dense de X, et f une application uniformément continue de $(A, d_{A\times A})$ dans (Y,δ). Si Y est complet, alors il existe une unique application uniformément continue $\widetilde{f}$ de (X,d) dans (Y,δ) telle que $\widetilde{f}_{|A} = f$.*

Démonstration. Considérons un élément x de X et essayons de définir $\widetilde{f}(x)$. L'ensemble A étant dense, la proposition 1.2.10 nous assure qu'il existe une suite $(a_n)_{n\in\mathbb{N}}$ d'éléments de A convergeant vers x. La fonction f est uniformément continue, ce qui signifie que

$$\forall\varepsilon>0\,,\ \exists\alpha\,/\,\forall(a,b)\in A^2\,,\ d(a,b)<\alpha\Rightarrow\delta(f(a),f(b))<\varepsilon.$$

Mais la suite $(a_n)_{n\in\mathbb{N}}$ est de Cauchy car convergente. Donc il existe un entier n_0 tel que

$$\forall n\geq n_0\,,\ \forall p\,,\ d(a_n,a_{n+p})<\alpha.$$

Donc la suite $(f(a_n))_{n\in\mathbb{N}}$ est une suite de Cauchy de Y, donc elle converge vers une limite y.

Une première chose à vérifier: cette limite y est indépendante de la suite $(a_n)_{n\in\mathbb{N}}$ choisie. En effet, soient $(a_n)_{n\in\mathbb{N}}$ et $(b_n)_{n\in\mathbb{N}}$ deux suites convergeant vers x. Par un raisonnement analogue au précédent, l'uniforme continuité de la fonction f sur A assure que

$$\lim_{n\to\infty}\delta(f(a_n),f(b_n))=0.$$

Donc la limite y est bien indépendante du choix de la suite $(a_n)_{n\in\mathbb{N}}$. On peut alors définir la fonction $\widetilde{f}$ par

$$\widetilde{f}(x)=\lim_{n\to\infty}f(a_n)\text{ pour toute suite }(a_n)_{n\in\mathbb{N}}\in A^{\mathbb{N}}\,/\,\lim_{n\to\infty}a_n=x.$$

Comme la fonction f est continue sur A, il est clair que $\widetilde{f}_{|A}=f$. Vérifions maintenant que la fonction $\widetilde{f}$ est uniformément continue sur X. Le fait que la fonction f soit uniformément continue sur A se traduit par

$$\forall\varepsilon\,,\ \exists\alpha\,/\,\forall(a,b)\in A^2\,,\ d(a,b)<\alpha\Rightarrow\delta(f(a),f(b))<\varepsilon. \tag{3.4}$$

Considérons maintenant un couple (x,y) d'éléments de X tels que $d(x,y)<\alpha$. Il existe alors deux suites $(a_n)_{n\in\mathbb{N}}$ et $(b_n)_{n\in\mathbb{N}}$ d'éléments de A telles que

$$\lim_{n\to\infty}a_n=x\quad\text{et}\quad\lim_{n\to\infty}b_n=y.$$

Il existe donc un entier n_0 tel que

$$n\geq n_0\Rightarrow d(a_n,b_n)<\alpha.$$

Donc, d'après la relation (3.4), on a

$$\delta(f(a_n),f(b_n))<\varepsilon.$$

Par passage à la limite, on obtient que

$$d(x,y)<\alpha\Rightarrow\delta(\widetilde{f}(x),\widetilde{f}(y))\leq\varepsilon,$$

ce qui achève la démonstration du théorème. □

Suite de la démonstration du théorème 3.4.1 Cette fonction est définie par

$$\ell(x)=\lim_{p\to\infty}\ell(y_p)\,,\ (y_p)_{p\in\mathbb{N}}\in\mathcal{X}^{\mathbb{N}}\,/\,\lim_{p\to\infty}y_p=x. \tag{3.5}$$

D'après le lemme 3.4.2, il suffit de démontrer que

$$\forall x\in X\,,\ \lim_{n\to\infty}f_{\psi(n)}(x)=\ell(x).$$

La réunion de A et de $\{\ell\}$ est naturellement une partie uniformément équicontinue. Soit ε un réel strictement positif arbitraire et x un élément de X. Il existe un réel strictement positif α tel que

$$\forall n\in\mathbb{N}\,,\ d(x,x')<\alpha\Rightarrow\sup_{n\in\mathbb{N}}\delta(f_{\psi(n)}(x),f_{\psi(n)}(x'))+\delta(\ell(x),\ell(x'))<\frac{\varepsilon}{2}.$$

Il existe un entier p tel que

$$d(x,x_p)<\alpha.$$

Il en résulte que

$$\begin{aligned} \delta(f_{\psi(n)}(x),\ell(x)) &\leq \delta(f_{\psi(n)}(x),f_{\psi(n)}(x_p))+\delta(f_{\psi(n)}(x_p),\ell(x_p))+\delta(\ell(x_p),\ell(x)) \\ &\leq \frac{\varepsilon}{2}+\delta(f_{\psi(n)}(x_p),\ell(x_p)). \end{aligned}$$

La relation (3.3) permet de conclure la démonstration du théorème d'Ascoli. □

Exercice 3.4.4. *Démontrez la réciproque du théorème d'Ascoli, à savoir que si une partie A de $\mathcal{C}(X,Y)$ est compacte, alors elle vérifie les conditions $i)$ et $ii)$ de l'énoncé du théorème d'Ascoli.*

Exercice 3.4.5. *Soient (X,d) et (Y,δ) deux espaces métriques compacts. Démontrez que l'ensemble des fonctions k-lipschitziennes de (X,d) dans (Y,δ) est compact dans $\mathcal{C}(X,Y)$.*

3.5 Compléments et exercices

Exercice 3.5.1. *Soit (X,d) un espace métrique. On considère l'ensemble $X^{\mathbb{N}}$ des suites d'éléments de X.*

1) On définit l'application $d_{\mathbb{N}}$ de $X^{\mathbb{N}}\times X^{\mathbb{N}}$ dans $\mathbb{R}^+$ par

$$d_{\mathbb{N}}((x(n))_{\mathbb{N}},(y(n))_{\mathbb{N}})=\sum_{n=0}^{\infty}\frac{1}{2^n}\min\{d(x(n),y(n)),1\}.$$

Démontrer que $d_{\mathbb{N}}$ est une distance sur $X^{\mathbb{N}}$.

2) Soit $(x_p)_{p\in\mathbb{N}}$ une suite d'éléments de $X^{\mathbb{N}}$ et x un élément de $X^{\mathbb{N}}$. Démontrez que

$$\lim_{p\to\infty}x_p=x \quad \text{si et seulement si} \quad \forall n\in\mathbb{N}\,,\ \lim_{p\to\infty}x_p(n)=x(n).$$

3) Démontrez que si (X,d) est un espace métrique compact, alors $(X^{\mathbb{N}},d_{\mathbb{N}})$ est aussi un espace métrique compact.

4) Démontrez que U est un ouvert de $(X^{\mathbb{N}},d_{\mathbb{N}})$ si et seulement si pour tout $x\in U$, il existe une partie finie J de $\mathbb{N}$ et un réel strictement positif α tel que

$$\forall y\in X^{\mathbb{N}}\,,\ \forall j\in J\,,\ d(y_j,x_j)<\alpha\Rightarrow y\in U.$$

Exercice 3.5.2. *Exhiber une suite de fonctions $(f_n)_{n\in\mathbb{N}}$ de $[0,1]$ dans $\mathbb{R}$ qui soit une suite bornée de l'espace des fonctions Lipschitziennes, telle que la suite $(f_n)_{n\in\mathbb{N}}$ tend vers 0 dans l'espace $\mathcal{C}([0,1],\mathbb{R})$ mais $(f_n)_{n\in\mathbb{N}}$ ne tende pas vers 0 dans l'espace des fonctions lipschiztiennes de $[0,1]$ dans $\mathbb{R}$.*

Chapitre 4

Connexité

Dans tout ce chapitre, on peut remplacer (X, d) espace métrique par (X, Θ) espace topologique.

4.1 Notion de connexité par arcs

Définition 4.1.1. *Soit A une partie d'un espace métrique (X, d). On dit que A est connexe par arcs si et seulement si pour tout couple (x_0, x_1) de points de X, il existe une application continue γ de $[0, 1]$ dans A telle que $\gamma(j) = x_j$. Si la partie X est connexe par arcs, on dit alors que l'espace métrique (X, d) est connexe par arcs.*

Donnons quelques exemples.

- Tout d'abord si x est un point d'un espace métrique (X, d), alors la partie $\{x\}$ est connexe par arcs (prendre $\gamma(t) \equiv x$.

- L'espace $\mathbb{R}^N$ est connexe par arcs. En effet, soit deux points x_0 et x_1 de $\mathbb{R}^N$, l'application

$$\gamma(t) \stackrel{\text{déf}}{=} (1 - t)x_0 + tx_1$$

  est une apllication continue de $[0, 1]$ dans $\mathbb{R}^N$ telle que $\gamma(j) = x_j$.

Proposition 4.1.2. *Les parties connexes par arcs de $(\mathbb{R}, |\cdot|)$ sont les intervalles.*

Démonstration. Soit I un intervalle de $\mathbb{R}$ et (a, b) un couple de points de I. Posons $\gamma(t) = a + t(b - a)$. Par définition d'un intervalle $\gamma([0, 1])$ est inclus dans I. Donc I est connexe par arcs.

Considérons maintenant une partie A de $\mathbb{R}$ qui n'est pas un intervalle. Ceci signife qu'il existe un triplet (x_0, x_1, α) de $A^2 \times A^c$ tel que $x_0 < c < x_1$. Soit γ une fonction continue de $[0, 1]$ dans $\mathbb{R}$ telle que $\gamma(j) = x_j$. Le théorème des valeurs intermédiares implique l'existence de point t_0 de l'intervalle $[0, 1]$ tel que $\gamma(t_0) = \alpha$. Donc $\gamma([0, 1)]$ n'est pas inclus dans A et donc A n'est pas connexe par arcs. □

Un autre exemple d'espaces connexes par arcs est donné par le théorème suivant.

Proposition 4.1.3. *Soit x un point de $\mathbb{R}^2$ et y_0 et y_1 deux points de $\mathbb{R}^2 \setminus \{x\}$. Il existe une application continue γ de $[0, 1]$ de $\mathbb{R}^2 \setminus \{x\}$ telle que $\gamma(0) = y_0$ et $\gamma(1) = y_1$.*

Démonstration. Si x n'appartient pas au segment $[y_0, y_1]$, alors la fonction γ définie sur $[0, 1]$ par

$$\gamma(t) \stackrel{\text{déf}}{=} (1-t)y_0 + ty_1$$

convient. Si x appartient au segment $]y_0, y_1[$, alors il existe t_0 dans l'intervalle $[0, 1]$ tel que

$$x = (1-t_0)y_0 + t_0y_1.$$

Soit ε_0 un réel strictement positif tel que ε_0 soit strictement inférieur à $\min\{t_0, 1-t_0\}$ et $\overrightarrow{N}$ un vecteur de $\mathbb{R}^2$ non colinéaire à $y_1 - y_0$. On définit alors la fonction γ par

$$\begin{aligned}
\gamma(t) &\stackrel{\text{déf}}{=} (1-t)y_0 + ty_1 \quad \text{pour} \quad t \in\ [0,1]\setminus]t_0-\varepsilon_0, t_0+\varepsilon_0[\quad \text{et sinon}\\
\gamma(t) &\stackrel{\text{déf}}{=} (1-t_0)y_0 + t_0y_1\ - \varepsilon_0(y_1-y_0)\cos\Big(\pi\frac{t-t_0+\varepsilon_0}{2\varepsilon_0}\Big) + \varepsilon_0\overrightarrow{N}\sin\Big(\pi\frac{t-t_0+\varepsilon_0}{2\varepsilon_0}\Big).
\end{aligned}$$

D'où la proposition. □

Théorème 4.1.4. *Soient (X, d) et (Y, δ) deux espaces métriques, f une application continue de X dans Y. Si A est connexe par arcs, alors $f(A)$ est connexe par arcs.*

Démonstration. Soit y_0 et y_1 deux points de $f(A)$. Par définition, il existe un couple de points (x_0, x_1) de A tel que $f(x_j) = y_j$. Comme A est connexe par arcs, il existe une application continue γ de $[0, 1]$ dans A tel que $\gamma(j) = x_j$. D'après la proposition 1.3.1, l'application $\widetilde{g} \stackrel{\text{déf}}{=} f \circ \gamma$ est une application continue. De tout évidence $\widetilde{\gamma}([0, 1])$ est inclus dans $f(A)$ et $\widetilde{\gamma}(j) = y_j$. D'où le théorème. □

Des trois résultats ci-dessus, on peut déduire le corollaire suivant

Corollaire 4.1.5. *Il n'existe pas d'homéomorphismes (c'est-à-dire de bijection continue d'inverse continue) de $\mathbb{R}^2$ dans $\mathbb{R}$.*

Démonstration. Soit φ une application bijective de $\mathbb{R}^2$ dans $\mathbb{R}$. C'est une application bijective de $\mathbb{R}^2 \setminus \{0\}$ sur $\mathbb{R}\setminus\{\varphi(0)\}$. Comme $\mathbb{R}^2 \setminus \{0\}$ est connexe par arcs et pas $\mathbb{R}\setminus\{\varphi(0)\}$. Donc φ n'est pas continue. □

Théorème 4.1.6. *Soit $(A_\lambda)_{\lambda\in\Lambda}$ une famille de parties connexes par arcs d'un espace métrique (X, d). On a*

$$\bigcap_{\lambda\in\Lambda} A_\lambda \neq \emptyset \Longrightarrow A \stackrel{\text{déf}}{=} \bigcup_{\lambda\in\Lambda} A_\lambda \text{ est connexe par arcs.}$$

Démonstration. Soient x_0 et x_1 deux poins de A. Par définition de A, il existe λ_0 et λ_1 tels que x_j soit dans A_{λ_j}. Par hypothèse, il existe un point z appartenant à l'intersection de A_{λ_0} et de A_{λ_1}. Ces deux ensembles sont connexes par arcs; il existe donc deux applications continues γ_0 et γ_1 telles que

$$\gamma_j(0) = x_j \quad \text{et} \quad \gamma_j(1) = z.$$

On définit alors l'application γ par

$$\gamma(t) = \gamma_0(2t) \quad \text{si } t \in \Big[0, \frac{1}{2}\Big] \quad \text{et} \quad \gamma(t) = \gamma_1(2(1-t)) \quad \text{si } t \in \Big[\frac{1}{2}, 1\Big].$$

L'application γ est continue (exercice:vérifiez-le!) et donc A est connexe par arcs. □

Grâce à cette proposition, nous allons pouvoir définir la notion de composante connexe par arcs d'une partie d'une espace métrique.

Définition 4.1.7. *Soit x un point d'un espace métrique (X, d). On appelle composante connnexe par arcs de x (et l'on note Π^a_x) la réunion de toutes les parties connexes par arcs de X contenant x.*

Remarques

- Il existe toujours une partie connexe par arcs contenant x, c'est $\{x\}$.

- D'après le théorème 4.1.6, c'est la plus grand (au sens de l'inclusion) connexe par arcs de X contenant x.

- Toujours d'après le théorème 4.2.6, si x et y sont deux points distincts de X, ou bien $\Pi^a_x = \Pi^a_y$ ou bien $\Pi^a_x \cap \Pi^a_y = \emptyset$. Ceci implique que l'on a une partition de l'espace X en parties connexes par arcs.

Nous allons démontrer la proposition suivante qui donne une condition nécessaire en terme d'ouverts pour qu'une partie soit connexe par arcs.

Proposition 4.1.8. *Soit A une partie connexe par arcs d'un espace métrique (X, d). On considère deux ouverts U_0 et U_1 de X tels que A soit inclus dans $U_0 \cup U_1$. Si de plus $A \cap U_0$ et $A \cap U_1$ sont non vides, alors $A \cap U_0 \cap U_1$ est non vide.*

Démonstration. Remarquons tout d'abord que si $A \cap U_0 = A \cap U_1$, alors $A \cap U_0 \cap U_1 = A \cap U_0$ et donc $A \cap U_0 \cap U_1$ est non vide. Si $A \cap U_0$ est différent de $A \cap U_1$, il existe alors (x_0, x_1) dans $(A \cap U_0) \times A \cap U_1 \setminus U_0$. Comme A est supposé connexe par arcs, il existe une application continue de $[0, 1]$ dans A telle que $\gamma(j) = x_j$. Introduisons l'ensemble

$$I \stackrel{\text{déf}}{=} \bigl\{t \in [0, 1] \,/\, \forall t' \leq t \;\; \gamma(t') \in U_0\bigr\}.$$

Soit t_0 borne supérieure de I. Comme γ est continue, il existe un réel strictement positif α tel que $\gamma([0, \alpha[)$ soit inclus dans $A \cap U_0$. Donc t_0 est strictement positif. De plus, si t est un élément de I, $\gamma(t)$ appartient à U_0. Il existe donc un réel strictement positif β tel que $\gamma([t, t + \beta[)$ soit inclus dans $A \cap U_0$. Donc t n'est pas un majorant de I. Donc $\gamma(t_0)$ n'est pas dans $A \cap U_0$. Ainsi $\gamma(t_0)$ appartient à $A \cap U_1$. De par la continuité de γ, il existe un réel strictement positif β' tel que $\gamma(]t_0 - \beta', t_0])$ est inclus dans $A \cap U_1$. Comme $\gamma([0, t_0[)$ est inclus dans $A \cap U_0$, l'ensemble $A \cap U_0 \cap U_1$ contient $\gamma(]t_0 - \beta', t_0[)$ et est donc non vide. □.

On peut déduire de ce résultat le corollaire suivant.

Corollaire 4.1.9. *Soient x_0 et x_1 deux points distincts d'un espace métrique (X, d). Alors l'ensemble $\{x_0, x_1\}$ n'est pas connexe par arcs.*

Démonstration. Soit $U_j = B(x_j, \alpha)$ où α est un réel strictement positif strictement inférieur à $d(x_0, x_1)$. Il est clair que $\{x_0, x_1\}$ est inclus dans $U_0 \cup U_1$, que $A \cap U_j = \{x_j\}$ et que $A \cap U_0 \cap U_1 = \emptyset$. D'où le corollaire. □

4.2 Notion de connexité

Définition 4.2.1. *Soit (X, d) un espace métrique. On dit qu'une partie A de X est connexe si et seulement si, pour tout couple d'ouverts U_1 et U_2 de X, tels que $A \subset U_0 \cup U_1$, on a*

$$(A \cap U_0 \neq \emptyset \quad et \quad A \cap U_1 \neq \emptyset) \Longrightarrow A \cap U_0 \cap U_1 \neq \emptyset.$$

Lorsque la propriété ci-dessus est vérifiée pour $A = X$, on dit que l'espace X est connexe.

Premiers exemples

- Soit (X, d) un espace métrique, pour tout x de X, la partie $\{x\}$ est connexe. En effet, si deux parties U_0 et U_1 de X sont telles x soit dans U_0 et dans U_1, alors x est dans $U_0 \cap U_1$ qui est donc non vide.

- Soit (X, d) un espace métrique contenant au moins deux points distincts x_0 et x_1. La partie $\{x_0, x_1\}$ n'est pas connexe. En effet, on considère les deux boules de centre x_j et de rayon r strictement inférieur à la moitié de la distance de x_0 à x_1. On a alors

  $$\{x_0, x_1\} \subset B(x_0, r) \cup B(x_1, r) \quad \text{et} \{x_0, x_1\} \cap B(x_j, r) = \{x_j\}.$$

  alors que $\{x_0, x_1\} \cap B(x_0, r) \cap B(x_1, r) = \emptyset$.

- La dernière proposition de la section précédente montre que toute partie connexe par arcs est connexe.

Théorème 4.2.2. *Les parties connexes de $(\mathbb{R}, |\cdot|)$ sont les intervalles.*

Démonstration. Soit A une partie de $\mathbb{R}$ qui n'est pas un intervalle. Ceci signifie qu'il existe un triplet (a, b, c) de $A^2 \times A^c$ tel que $a < c < b$. Posons

$$U_1 \stackrel{\text{déf}}{=}]-\infty, c[\quad \text{et} \quad U_2 \stackrel{\text{déf}}{=}]c, \infty[.$$

Ces deux ensembles sont des ouverts de $\mathbb{R}$ tels que

$$A \subset U_1 \cup U_2\,,\; U_1 \cap U_2 \cap A = \emptyset\,,\; a \in U_1 \quad \text{et} \quad b \in U_2.$$

Donc A n'est pas connexe.

Réciproquement, soit I un intervalle de $\mathbb{R}$ (non réduit à un point). Nous allons démontrer par contraposition que I est connexe. Soit U_1 et U_2 deux ouverts de $\mathbb{R}$ tels que I soit inclus dans $U_1 \cup U_2$ et tels que $I \cap U_1$ et $I \cap U_2$ soient non vides. Nous allons démontrer qu'ils sont alors d'intersection non vide, c'est-à-dire que $I \cap U_1 \cap U_2$ est non vide.

Pour j valant 1 ou 2, on considère x_j un point de $I \cap U_j$. On peut, dans perte de généralité, supposer que $x_1 < x_2$. Comme U_1 est un ouvert, il existe $x > x_1$ tel que $[x_1, x[$ soit inclus dans U_1. On peut prendre x strictement inférieur à x_2. Comme I est un intervalle, x est dans I. Soit y dans $]x_1, x_2[$ (donc dans I) tel que $[x_1, y[$ soit inclus dans U_1. Comme U_1 est ouvert, il existe un z dans $]y, x_2[$ tel que $[y, z[$ soit inclus dans U_1. Donc si M_1 désigne la borne supérieure des $x < x_2$ tels que $[x_1, x[$ soit inclus dans $I \cap U_1$, alors M_1 n'appartient pas à U_1. Mais il appartient à I car M_1 est entre x_1 et x_2 et I est un intervalle. Donc M_1 appartient à $I \cap U_2$. Donc, il existe un $x_3 < x_2$ tel que $]x_3, x_2]$ soit inclus dans $I \cap U_2$. Mais, pour tout x dans $[x_1, M_1[$, x appartient à $I \cap U_1$. Donc $I \cap U_1 \cap U_2 \neq \emptyset$. □

Théorème 4.2.3. *Soient (X, d) et (Y, δ) deux espaces métriques, f une application continue de X dans Y. Alors, si A est connexe, $f(A)$ est connexe.*

Démonstration. Nous allons à nouveau procéder par contraposition. Soit A une partie de X telle que $f(A)$ ne soit pas connexe. Il existe alors deux ouverts V_1 et V_2 de X tels que

$$f(A) \subset V_1 \cup V_2\,,\ f(A) \cap V_1 \cap V_2 = \emptyset\,,\ f(A) \cap U_1 \neq \emptyset \quad \text{et} \quad f(A) \cap U_1 \neq \emptyset.$$

Comme la fonction f est continue, les deux ensembles $U_j \overset{\text{déf}}{=} f^{-1}(V_j)$ sont ouverts. De plus, on a

$$A \subset f^{-1}(f(A)) \subset f^{-1}(V_1 \cup V_2) \subset f^{-1}(V_1) \cup f^{-1}(V_2) = U_1 \cup U_2. \tag{4.1}$$

Si x est un point de $A \cap U_1 \cap U_2$, alors $f(x)$ appartient à $f(A) \cap V_1 \cap V_2$ qui est l'ensemble vide. Donc $A \cap U_1 \cap U_2 = \emptyset$. Or cet ensemble est supposé être vide. Donc

$$A \cap U_1 \cap U_2 = \emptyset. \tag{4.2}$$

De plus $f(A) \cup V_j$ est supposé non vide. Ceci implique l'existence d'un point x_j de A tel que $f(x_j)$ soit dans V_j ce qui signifie exactement que $A \cap f^{-1}(V_j)$ est non vide. Avec (4.1) et (4.2), ceci signifie que A n'est pas connexe. □

Proposition 4.2.4. *Soit A une partie d'un espace métrique (X,d). Alors si A est connexe, alors $\overline{A}$ est connexe.*

Démonstration. Nous allons procéder par contraposition. Si $\overline{A}$ n'est pas connexe, alors il existe deux ouverts de X U_1 et U_2 tels que

$$\overline{A} \subset U_1 \cup U_2\,,\ \overline{A} \cap U_1 \cap U_2 = \emptyset\,,\ \overline{A} \cap U_1 \neq \emptyset \quad \text{et} \quad \overline{A} \cap U_2 \neq \emptyset.$$

Comme $A \subset \overline{A}$, la seule chose à démontrer pour démontrer que A n'est pas connexe est que $A \cap U_j$ est non vide pour $j = 1$ ou $j = 2$. Soit x_j un point de $\overline{A} \cap U_j$. Par définition de l'adhérence, cela implique que $A \subset U_j$ n'est pas vide. D'où la proposition. □

Remarques Il est par contre faux en général que l'intérieur d'une partie connexe soit connexe, même s'il se trouve que c'est vrai sur $\mathbb{R}$. En effet, d'après le théorème 4.2.2, l'intérieur d'un intervalle non réduit à un point est un intervalle.

Il est faux que l'adhérence d'une partie connexe par arcs le soit. En effet, considérons la partie A de $\mathbb{R}^2$ définie par

$$A \overset{\text{déf}}{=} \Big\{ \Big(t, \sin\frac{1}{t}\Big),\ t \in]0,\pi] \Big\}.$$

C'est une partie connexe par arcs (exercice:vérifiez-le!). L'adhérence de A est $A \cup \{0\} \times [-1,1]$ qui n'est pas connexe par arcs. (exercice pas si facile: démontrez-le). Ainsi $\overline{A}$ est connexe et par connexe par arcs.

Proposition 4.2.5. *Soit A une partie d'un espace métrique (X,d). Cette partie est connexe si et seulement si pour toute partie B de A,*

$$\Big(B \ \textit{ouverte et fermée } \underline{\textit{de } A}\Big) \Longrightarrow \Big(B = A \textit{ ou } B = \emptyset\Big).$$

Démonstration. Soit B une partie ouverte et fermée de A. Comme B est une partie ouverte de A, il existe un ouvert U_1 de X tel que $B = U_1 \cap A$. Comme B est une partie fermée de A, B^c est une partie ouverte de A et donc il existe un ouvert U_2 de X tel que $B^c = U_2 \cap A$. On a donc

$$A \subset U_1 \cup U_2 \quad \text{et} \quad U_1 \cap U_2 \cap A = \emptyset.$$

L'espace X étant connexe, on a

$$A \subset U_1 \quad \text{ou} \quad A \subset U_2,$$

ce qui signifie que $B = A$ ou $B = \emptyset$.

Réciproquement, soient U_1 et U_2 deux ouverts tels que

$$A \subset U_1 \cup U_2 \quad \text{et} \quad U_1 \cap U_2 \cap A.$$

Par définition des ouverts de A, $B \stackrel{\text{déf}}{=} A \cap U_1$ est un ouvert de A dont le complémentaire dans A est $U_2 \cap A$ qui est aussi un ouvert. Donc $A \cap U_1 = A$ ou $A \cap U_2 = \emptyset$, ce qui signifie, comme A est inclus dans $U_1 \cup U_2$, que A est inclus dans U_1 ou dans U_2. D'où la proposition. □

Théorème 4.2.6. *Soit $(A_\lambda)_{\lambda \in \Lambda}$ une famille de parties connexes d'un espace métrique (X, d). On a*

$$\bigcap_{\lambda \in \Lambda} A_\lambda \neq \emptyset \Longrightarrow \bigcup_{\lambda \in \Lambda} A_\lambda \text{ est connexe.}$$

Démonstration. Nous allons procéder par contraposition. Supposons que la réunion des A_λ (notée A dans la suite de la démonstration) ne soit pas connexe. Cela implique l'existence de deux ouverts de X tels que

$$A \subset U_1 \cup U_2\,,\ A \cap U_1 \cap U_2 = \emptyset\,,\ A \cap U_1 \neq \emptyset \quad \text{et} \quad A \cap U_2 \neq \emptyset.$$

Comme chaque parties A_λ est connexe, et que

$$A_\lambda \subset U_1 \cup U_2 \quad \text{et} \quad A \cap U_1 \cap U_2 = \emptyset,$$

la partie A_λ est ou bien incluse dans U_1 ou bien incluse dans U_2. Soit Λ_1 et Λ_2 les deux ensembles d'indice défini par

$$\Lambda_j \stackrel{\text{déf}}{=} \{\lambda \in \Lambda \,/\, A_\lambda \subset U_j\}.$$

Comme, pour chaque j, l'ensemble $A \cap U_j$ est non vide, les deux ensembles d'indice Λ_1 et Λ_2 sont non vides. Ainsi donc, on peut écrire

$$\begin{aligned} \bigcap_{\lambda \in \Lambda} A_\lambda &= \Big(\bigcap_{\lambda \in \Lambda_1} A_\lambda\Big) \cap \Big(\bigcap_{\lambda \in \Lambda_2} A_\lambda\Big) \\ &\subset U_1 \cap U_2. \end{aligned}$$

Comme A (qui est la réunion des A_λ) est d'intersection vide avec $U_1 \cap U_2$, on a

$$\bigcap_{\lambda \in \Lambda} A_\lambda \subset A \cap U_1 \cap U_2 = \emptyset$$

ce qui démontre le théorème. □

Grâce à cette proposition, nous allons pouvoir, comme dans la section précédente, définir la notion de composante connexe d'une partie d'une espace métrique.

Définition 4.2.7. *Soit x un point d'un espace métrique (X, d). On appelle composante connnexe de x (et l'on note Π_x) la réunion de toutes les parties connexes de X contenant x.*

Remarques

- Il existe toujours une partie connexe contenant x, c'est $\{x\}$.
- D'après le théorème 4.2.6, c'est la plus grand (au sens de l'inclusion) connexe de X contenant x.
- Toujours d'après le théorème 4.2.6, si x et y sont deux points distincts de X, ou bien $\Pi_x = \Pi_y$ ou bien $\Pi_x \cap \Pi_y = \emptyset$.

Chapitre 5

Espaces normés et espaces de Banach

Dans toute la suite du cours, $\mathbb{K}$ désignera $\mathbb{R}$ ou $\mathbb{C}$.

5.1 Définition des espaces normés et de Banach

Définition 5.1.1. *Soit E un espace vectoriel sur $\mathbb{K}$. On dit qu'une application N de E dans $\mathbb{R}^+$ est une norme sur E si et seulement si les trois conditions suivantes sont satisfaites:*

- $N(x) = 0 \Longleftrightarrow x = 0$,
- *Pour tout λ appartenant à $\mathbb{K}$, on a $N(\lambda x) = |\lambda| N(x)$.*
- $N(x + y) \leq N(x) + N(y)$,

Le couple (E, N) est appelé un espace normé

Notations Très souvent, on désigne une norme par $\|\cdot\|_E$ ou bien par $\|\cdot\|$.

Exemples

- On considère sur $\mathbb{R}^N$ les différentes normes suivantes:

$$\begin{aligned}
\|x\|_e &\stackrel{\text{déf}}{=} \left(\sum_{j=1}^{N} x_j|^2\right)^{\frac{1}{2}} \\
\|x\|_\infty &\stackrel{\text{déf}}{=} \max_{1\leq j\leq N} |x_j| \\
\|x\|_1 &\stackrel{\text{déf}}{=} \sum_{j=1}^{N} |x_j|.
\end{aligned}$$

- Soit $\mathcal{M}_{m,n}(\mathbb{K})$ l'ensemble des matrices à m lignes et n colonnes. On pose, pour une matrice $A = (a^i_j)$

$$\|A\| \stackrel{\text{déf}}{=} (\mathrm{tr}(A^t \overline{A}))^{\frac{1}{2}} \stackrel{\text{déf}}{=} \Big(\sum_{i,j} |a^i_j|^2\Big)^{\frac{1}{2}}.$$

- Soit X un ensemble, on considère l'espace des fonctions bornées de X dans $\mathbb{K}$. On définit alors l'application
$$f \longmapsto \sup_{x\in X} |f(x)|\,;$$
c'est une norme au sens ci-dessus.

Si $X = \mathbb{N}$, cet espace est l'espace des suites bornées à valeurs dans $\mathbb{K}$. On note cet espace $\ell^\infty(\mathbb{N};\mathbb{K})$ et l'on note un élément x de cet espace $(x(n))_{n\in\mathbb{N}}$. Les espaces de suites constitue les exemples les plus simples à étudier d'espaces vectoriels de dimension infinie. Voici deux autres exemples d'espaces de suites qui sont très intéressants.

Définition 5.1.2. *On appelle $\ell^1(\mathbb{N};\mathbb{K})$ l'ensemble des suites à valeurs dans $\mathbb{K}$ telles que la série de terme général $x(n)$ soit absolument convergente.*

On appelle $\ell^2(\mathbb{N};\mathbb{K})$ l'ensemble des suites à valeurs dans $\mathbb{K}$ telles que la série de terme général $|x(n)|^2$ soit convergente.

Proposition 5.1.3. *Les espaces $\ell^1(\mathbb{N};\mathbb{K})$ et $\ell^2(\mathbb{N},\mathbb{K})$ sont des espaces vectoriels et les applications*
$$\|x\|_1 \overset{\text{déf}}{=} \sum_{n=0}^{\infty} |x(n)| \quad et \quad \|x\|_2 \overset{\text{déf}}{=} \Big(\sum_{n=0}^{\infty} |x(n)|^2\Big)^{\frac12}$$
définissent des normes sur $\ell^1(\mathbb{N};\mathbb{K})$ et $\ell^2(\mathbb{N},\mathbb{K})$ respectivement.

Démonstration. Si $\|x\|_1 = 0$, alors on a, pour tout n, $x(n) = 0$ et donc $x \equiv 0$. Si la série de terme général $|x(n)|$ converge, celle de terme général $|\lambda x(n)|$ aussi et l'on a
$$\sum_{n=0}^{\infty} |\lambda x(n)| = |\lambda| \sum_{n=0}^{\infty} |x(n)|.$$
Enfin, pour tout entier N, on a
$$\begin{aligned} \sum_{n=0}^{N} |x(n)+y(n)| &\leq \sum_{n=0}^{N} |x(n)| + \sum_{n=0}^{N} |y(n)| \\ &\leq \sum_{n=0}^{\infty} |x(n)| + \sum_{n=0}^{\infty} |y(n)| \end{aligned}$$
Ainsi donc on a $\|x+y\|_1 \leq \|x\|_1 + \|y\|_1$. Pour l'espace $\ell^2(\mathbb{N};\mathbb{K})$, la démonstration est la même une fois observé que
$$\Big(\sum_{n=0}^{N} |x(n)+y(n)|^2\Big)^{\frac12} \leq \Big(\sum_{n=0}^{N} |x(n)|^2\Big)^{\frac12} + \Big(\sum_{n=0}^{N} |y(n)|^2\Big)^{\frac12}$$
□

Remarque On a $\ell^1(\mathbb{N};\mathbb{K}) \subset \ell^2(\mathbb{N};K) \subset \ell^\infty(\mathbb{N};\mathbb{K})$.

La proposition suivante va munir naturellement un espace normé d'une structure d'espace métrique.

Proposition 5.1.4. *Soit $(E, \|\cdot\|)$ un espace normé, alors l'application définie par*
$$d \left\{ \begin{array}{ccc} E\times E & \to & \mathbb{R}^+ \\ (x,y) & \mapsto & \|x-y\| \end{array} \right.$$
est une distance sur E, appelée distance associée à la norme $\|\cdot\|$.

Démonstration. Tout d'abord on a,

$$d(x,y) = 0 \iff \|x-y\| = 0 \iff x-y = 0.$$

De plus, en appliquant le deuxième propriété de la définition des normes avec $\lambda = -1$,

$$d(x,y) = \|x-y\| = \|y-x\| = d(y,x).$$

Ensuite, on a

$$d(x,y) = \|x-y\| = \|x-z+z-y\| \leq \|x-z\| + \|z-y\| = d(x,z) + d(z,y).$$

Définition 5.1.5. *Soit $(E, \|\cdot\|)$ un espace normé. On dit que $(E, \|\cdot\|)$ est un espace de Banach si et seulement si l'espace métrique (E,d) où d est la distance associée à la norme $\|\cdot\|$ (i.e. $d(x,y) = \|x-y\|$) est un espace complet.*

Exemples L'espace $\mathbb{R}^N$ muni de la norme euclidienne est un espace complet. Plus généralement, tout sous-espace vectoriel de dimension finie d'un espace vectoriel normé est un espace de Banach. Soient E un espace de Banach et X un ensemble; l'ensemble $\mathcal{B}(X,E)$ des fonctions bornées de X dans E muni de la norme

$$\|f\|_{\mathcal{B}(X,E)} \overset{\text{déf}}{=} \sup_{x\in X} \|f(x)\|_E$$

est un espace de Banach (voir Proposition 2.1.5 page 22). En particulier l'espace $\ell^\infty(\mathbb{N};\mathbb{K})$ est un espace de Banach. Pour démontrer que les espaces $\ell^1(\mathbb{N};\mathbb{K})$ et $\ell^2(\mathbb{N};K)$ sont de Banach, il est nécessaire de mieux comprendre le comportement des séries dans les espaces de Banach. Dans un espace de Banach, les séries absolument convergentes sont convergentes, ce qui se traduit par la proposition suivante.

Proposition 5.1.6. *Soit E un espace de Banach et $(x_n)_{n\in\mathbb{N}}$ une suite d'éléments de E. Alors, si la série $(\|x_n\|_E)_{n\in\mathbb{N}}$ est convergente, alors la suite*

$$S_N = \sum_{n=0}^{N} x_n \quad \textit{est convergente.}$$

Démonstration. On a

$$\begin{aligned} \|S_{N+P} - S_N\| &= \left\| \sum_{n=N+1}^{N+P} x_n \right\| \\ &\leq \sum_{n=N+1}^{N+P} \|x_n\|. \end{aligned}$$

L'hypothèse de sommabilité implique que, pour tout ε strictement positif, il existe un entier N_0 tel que

$$\forall N \geq N_0\,,\ \forall P\,,\ \sum_{n=N+1}^{N+P} \|x_n\| < \varepsilon.$$

La suite $(S_N)_{N\in\mathbb{N}}$ est donc de Cauchy, ce qui, vu que l'espace E est de Banach, assure la proposition. □

Pour démontrer que les espaces $\ell^1(\mathbb{N};\mathbb{K})$ et $\ell^2(\mathbb{N};K)$ sont de Banach, on utilise le résultat suivant, qui est une réciproque de la proposition ci-dessus.

Théorème 5.1.7. *Si E est un espace normé tel que pour toute suite $(x_n)_{n\in\mathbb{N}}$, on ait*

$$\sum_{n\in\mathbb{N}} \|x_n\|_E < \infty \Rightarrow S_N \stackrel{\text{déf}}{=} \sum_{n=0}^{N} x_n \quad \textit{converge,}$$

alors l'espace E est de Banach.

Démonstration. Soit $(a_n)_{n\in\mathbb{N}}$ une suite de Cauchy de E, nous allons extraire de $(a_n)_{n\in\mathbb{N}}$ une suite convergente, ce qui, d'après la proposition 2.1.2, assurera le résultat.

On définit l'application ϕ de $\mathbb{N}$ dans $\mathbb{N}$ par $\phi(0) = 0$ et

$$\phi(n+1) \stackrel{\text{déf}}{=} \min\left\{ m \geq \phi(n)+1\,,\ \sup_{p\geq 0} \|a_m - a_{m+p}\| \leq \frac{1}{(2+n)^2} \right\}.$$

Ainsi donc, on a, pour tout entier n,

$$\|a_{\phi(n+1)} - a_{\phi(n)}\| \leq \frac{1}{(1+n)^2}.$$

Posons $x_n \stackrel{\text{déf}}{=} a_{\phi(n+1)} - a_{\phi(n)}$. On a

$$\sum_{n\in\mathbb{N}} \|x_n\| \leq \sum_{n\in\mathbb{N}} \frac{1}{(1+n)^2} < \infty.$$

Donc par hypothèse, la suite

$$S_N = \sum_{n=0}^{N} x_n = a_{\phi(N)} - a_0$$

converge; d'où le théorème. □

Nous allons maintenant appliquer ce théorème pour démontrer le résultat suivant.

Théorème 5.1.8. *L'espace normé $(\ell^1(\mathbb{N};\mathbb{K}), \|\cdot\|_1)$ est un espace de Banach.*

Démonstration. Soit $(x_p)_{p\in\mathbb{N}}$ une suite d'éléments de $\ell^1(\mathbb{N};\mathbb{K})$ telle que

$$\sum_{p=0}^{\infty} \|x_p\|_1 < \infty.$$

Ceci signifie qu'il existe un réel strictement positif M tel que

$$\forall (P,N) \in \mathbb{N}^2\,,\ \sum_{p=0}^{P} \sum_{n=0}^{N} |x_p(n)| \leq M. \tag{5.1}$$

Ceci implique que, pour tout n fixé, la série de terme général $x_p(n)$ est absolument convergente dans $\mathbb{K}$ (qui est $\mathbb{R}$ ou $\mathbb{C}$). Elle converge donc vers $S(n)$. D'après (5.1), on a

$$\forall N \in \mathbb{N}\,,\ \sum_{n=0}^{N} |S(n)| \leq M.$$

Donc S appartient à $\ell^1(\mathbb{N};\mathbb{K})$. Il s'agit maintenant d'estimer $\Big\|S-\sum_{p=0}^{P}x_p\Big\|_1$ pour un entier P donné. Pour tout entier N, on a

$$\sum_{n=0}^{N}\Big|S(n)-\sum_{p=0}^{P}x_p(n)\Big| = \sum_{n=0}^{N}\Big|\lim_{P'\to\infty}\sum_{p=P+1}^{P'}x_p(n)\Big|.$$

La valeur absolue de la limite est la limite de la valeur absolue et la limite d'une somme (finie) est la somme des limites. Ainsi donc, on a

$$\sum_{n=0}^{N}\Big|S(n)-\sum_{p=0}^{P}x_p(n)\Big| = \lim_{P'\to\infty}\sum_{n=0}^{N}\Big|\sum_{p=P+1}^{P'}x_p(n)\Big|.$$

L'inégalité triangulaire implique que

$$\Big|\sum_{p=P+1}^{P'}x_p(n)\Big| \leq \sum_{p=P+1}^{P'}|x_p(n)|$$

D'où, en intervertissant les sommes (finies en n et p),

$$\sum_{n=0}^{N}\Big|S(n)-\sum_{p=0}^{P}x_p(n)\Big| \leq \limsup_{P'\to\infty}\sum_{p=P+1}^{P'}\sum_{n=0}^{N}|x_p(n)|.$$

Par définition de la norme $\|\cdot\|_1$, on a

$$\sum_{n=0}^{N}\Big|S(n)-\sum_{p=0}^{P}x_p(n)\Big| \leq \sum_{p=P+1}^{\infty}\|x_p\|_1.$$

Cette inégalité étant vraie pour tout N, on en déduit que

$$\Big\|S-\sum_{p=0}^{P}x_p\Big\|_1 \leq \sum_{p=P+1}^{\infty}\|x_p\|_1.$$

Par hypothèse, la série de terme général $\|x_p\|_1$ est convergente. Ceci implique que, pour tout $\varepsilon>0$, il existe un entier P_0 tel que

$$\sum_{p=P_0}^{\infty}\|x_p\|_1<\varepsilon.$$

D'après le théorème 5.1.7, le théorème est démontré. □

Théorème 5.1.9. *L'espace normé* $(\ell^2(\mathbb{N};\mathbb{K}),\|\cdot\|_2)$ *est un espace de Banach.*

Démonstration. Soit $(x_p)_{p\in\mathbb{N}}$ une suite d/'elément sde $\ell^2(\mathbb{N})$ telle que

$$\sum_{p\in\mathbb{N}}\|x_p\|_2<\infty.$$

Comme, pour tout n, on a $|x_p(n)| \leq \|x_p\|_2$, on en déduit que, pour tout n,

$$\sum_{p\in\mathbb{N}} |x_p(n)| < \infty.$$

D'après la théorie des séries à termes réelles ou complexes, on en déduit que, pour chaque n, la série (en p) de terme général $x_p(n)$ est absolument convergente et donc converge vers une limite que nous désignerons par $S(n)$. Démontrons que S appartient à $\ell^2(\mathbb{N})$ puis que la série des $(x_p)_{p\in\mathbb{N}}$ converge vers S dans $\ell^2(\mathbb{N})$.

Pour ce faire, nous allons considérer uniquement des valeurs des suites $x_p(n)$ et $S(n)$ pour les entiers n inférieurs à un entier N. Ainsi, la norme $\|\cdot\|_2$ devient la norme euclidienne sur $\mathbb{R}^N$. Notons S_N et $x_{p,N}$ les restrictions de S et de x_p à l'ensemble $\{0,\ldots,N\}$. Observons que l'inégalité triangualire implique, pour tout N, on a

$$\|S_N\|_2 \leq \sum_{p=0}^{\infty} \|x_{p,N}\|_2.$$

Comme $\|x_{p,N}\|_2 \leq \|x_p\|_2$, on en déduit que, pour tout N, $\|S_N\|_2$ est majoré par

$$\sum_{p=0}^{\infty} \|x_p\|_2$$

qui est indépendant de N. Ceci signifie que, pour tout N,

$$\sum_{n=0}^{N} |S(n)|^2 \leq \Big(\sum_{p=0}^{\infty} \|x_p\|_2\Big)^2$$

c'est-à-dire que S appartient à $\ell^2(N)$.

Démontrons maintenant la convergence dans $\ell^2(\mathbb{N})$. Pour tout N, on a

$$\begin{aligned}\Big\|S_N - \sum_{p=0}^{P} x_{p,N}\Big\|_2 &= \lim_{Q\to\infty}\Big\|\sum_{p=P}^{Q} x_{p,N}\Big\|_2 \\ &\leq \lim_{Q\to\infty}\sum_{p=P}^{Q} \|x_{p,N}\|_2.\end{aligned}$$

Comme $\|x_{p,N}\|_2 \leq \|x_p\|_2$, qui est le terme général d'une série convergente, on en déduit que, pour tout N,

$$\Big\|S_N - \sum_{p=0}^{P} x_{p,N}\Big\|_2 \leq \sum_{p=P}^{\infty} \|x_p\|_2.$$

Ceci se traduit par le fait que

$$\sum_{n=0}^{N}\Big(S(n) - \sum_{p=0}^{P} x_p\Big)^2 \leq \Big(\sum_{p=P}^{\infty} \|x_p\|_2\Big)^2.$$

En passant à la limite lorsque N tend vers l'infini et en prenant la racone carrée de l'inégalité ainsi obtenue, on en déduit que

$$\Big\|S - \sum_{p=0}^{P} x_p\|_2 \leq \sum_{p=P}^{\infty} \|x_p\|_2$$

ce qui conclut la démonstration du théorème. $\square$

5.2 Le cas des espaces de dimension finie

Le fait que la dimension de l'espace soit finie ou non induit de grandes différences sur la topologie comme on peut le voir au travers de l'énoncé suivant.

Théorème 5.2.1 (de Riesz)**.** *Soit E un espace vectoriel normé; la dimension de E est finie si et seulement si la boule unité fermée de E est compacte.*

Démonstration. Soit E un espace vectoriel normé de dimension finie N. Nous allons démontrer qu'il existe une bijection linéaire continue d'inverse continue de $\mathbb{R}^N$ muni de la norme $\|\cdot\|_\infty$ sur E. Considérons une base $(\vec{e}_j)_{1\leq j\leq N}$ de E et l'application linéaire bijective I définie par

$$I\left\{\begin{array}{ccc} \mathbb{R}^N & \longrightarrow & E \\ x=(x_j)_{1\leq j\leq N} & \longmapsto & \displaystyle\sum_{j=1}^N x_j\vec{e}_j. \end{array}\right.$$

L'application $I(x)$ est une bijection linéaire. Montrons qu'elle est continue. On a

$$\begin{aligned} \|I(x)\|_E &\leq \sum_{j=1}^N |x_j|\,\|\vec{e}_j\|_E \\ &\leq M\|x\|_\infty \quad \text{avec} \quad M \stackrel{\text{déf}}{=} \max_{1\leq j\leq N}\|\vec{e}_j\|_E. \end{aligned}$$

L'application I étant linéaire, on en déduit que

$$\|I(x)-I(y)\|_E \leq M\|x-y\|_\infty.$$

L'application

$$x \longmapsto \|I(x)\|_E$$

est donc continue de $\mathbb{R}^N$ dans $\mathbb{R}^+$. D'après le corollaire 3.3.6 page 34, la sphère $\mathbb{S}^{N-1}$, (i.e. l'ensemble des x de $\mathbb{R}^N$ tels que $\|x\|_\infty=1$) est compacte. De plus, I etant bijective, elle ne s'annule pas sur $\mathbb{S}^{N-1}$. D'après le théorème 3.3.9 page 35, il existe un réel strictement positif m tel que

$$\forall x\in\mathbb{R}^N\setminus\{0\}\,,\ m\leq \Big\|I\Big(\frac{x}{\|x\|_\infty}\Big)\Big\|_E\leq M.$$

L'application I étant linéaire, on a

$$\forall x\in\mathbb{R}^N\,,\ m\|x\|_\infty\leq\|I(x)\|_E\leq M\|x\|_\infty. \tag{5.2}$$

Donc I et I^{-1} sont des bijections linéaires continues. Donc les ouverts (resp. les fermés) (resp. les compacts) de E sont exactement les images des ouverts (resp. des fermés) (resp. des compacts) de $\mathbb{R}^N$ par I. Donc la boule unité ferme de E est un compact car c'est un fermé inclus dans l'image par I dans la boule fermé de centre 0 et de rayon m^{-1}. Donc la boule unité de E est compact.

Nous admettons la réciproque qui se démontre en prouvant que si l'espace vectoriel E n'est pas de dimension finie, alors la boule unité n'est pas compact. □

Définition 5.2.2. *Soit E un espace vectoriel et N_1 et N_2 deux normes sur E. Elles sont dites équivalentes si et seulement si il existe un réel strictement positif C tel que, pour tout x de E, on ait*

$$C^{-1}N_1(x)\leq N_2(x)\leq CN_1(x).$$

Remarques Cela signifie exactement que l'application Id (qui est bien sûr linéaire) est continue de (E, N_1) dans (E, N_2) et de (E, N_2) dans (E, N_1) (voir la section suivante). Dans ce cas, les distances associées sont équivalentes (voir la définition 1.6.2).

Proposition 5.2.3. *Soit E un espace vectoriel de dimension finie, alors toutes les normes sont équivalentes entre elles.*

Démonstration. Soient N_1 et N_2 deux normes sur E. D'aprè s la relation (5.2), pour $j \in \{1, 2\}$, on a

$$\forall (x, y) \in \mathbb{R}^N \times \mathbb{R}^N\,, \; m_j \|x - y\|_\infty \leq N_j(I(x - y)) \leq M_j \|x - y\|_\infty.$$

Ainsi donc, comme I est une bijection, on a pour tout y de E

$$\frac{m_2}{M_1} \leq N_2(x) \leq \frac{M_2}{m_1} N_1(x)$$

D'où le résultat en prenant

$$C = \max\left\{\frac{M_2}{m_1}, \frac{M_1}{m_2}\right\}.$$

□

5.3 Continuité des applications linéaires

La continuité des applications linéaires est décrite par le théorème suivant.

Théorème 5.3.1. *Soient E et F deux espaces vectoriels normés et ℓ une application linéaire de E dans F. L'application ℓ est lipschitzienne si et seulement si il existe un ouvert U de E tel que ℓ soit bornée sur U, c'est-à-dire tel que*

$$\sup_{x \in U} \|\ell(x)\|_F < \infty.$$

Démonstration. Il est clair qu'il existe un ouvert sur lequel une application lipschitzienne est bornée (par exemple n'importe quelle boule). Réciproquement, considérons un ouvert U sur lequel l'application linéaire ℓ est bornée. Soit x_0 appartenant à U. L'ensemble $U - x_0$ est un ouvert (exercice: démontrez-le!) contenant l'origine. Par définition d'un ouvert, il existe une boule ouverte de centre 0 et de rayon α incluse dans $U - x_0$. On peut alors écrire

$$\begin{aligned} M \stackrel{\text{déf}}{=} \sup_{x \in B(0,\alpha)} \|\ell(x)\| &\leq \sup_{x \in U - x_0} \|\ell(x)\| \\ &\leq \sup_{y \in U} \|\ell(y - x_0)\| \\ &\leq \sup_{y \in U} \|\ell(y)\| + \|\ell(x_0)\|. \end{aligned}$$

Donc l'application ℓ est bornée sur une boule de centre 0 et de rayon α. Pour tout y de $E \backslash \{0\}$, on a

$$\frac{\alpha}{2\|y\|_E} y \in B(0, \alpha).$$

On en déduit que

$$\left\| \ell\left(\frac{\alpha}{2\|y\|_E} y\right) \right\| \leq M.$$

Par linéarité de ℓ et homogénéité de la norme, on trouve que

$$\forall y \in E\,, \; \|\ell(y)\|_F \leq \frac{2M}{\alpha} \|y\|_E.$$

Cette inégalité étant vraie pour $y = x - x_0$, la démonstration du théorème est achevée. □

Lorsque l'espace vectoriel de départ E est de dimension finie, les applications linéaires sont toujours continues.

Théorème 5.3.2. *Soient E et F deux espaces vectoriels normés. Si E est de dimension finie, alors toute application linéaire de E dans F est continue.*

Démonstration. D'après la proposition 5.2.3 page 54, toutes les normes sont équivalentes. Soit $\mathcal{B} = (\overrightarrow{e}_j)_{1\leq j\leq N}$ une base de E, on choisit comme norme $\|x\|_\infty = \sup_{1\leq j\leq N} |x_j|$ où les x_j désignent les coordonnées de x dans la base $\mathcal{B}$. Soit ℓ une application linéaire de E dans F. On a

$$\|\ell(x)\|_F = \Big\|\sum_{j=1}^{N} x_j \ell(\overrightarrow{e}_j)\Big\|_F \leq M\|x\|_\infty \quad \text{avec} \quad M \overset{\text{déf}}{=} \sum_{j=1}^{N} \|\ell(\overrightarrow{e}_j)\|_F. \tag{5.3}$$

D'où le résultat. □

5.4 Les espaces d'applications linéaires continues

Proposition 5.4.1. *Soient E et F deux espaces vectoriels normés; on désigne par $\mathcal{L}(E,F)$ l'ensemble des applications linéaires continues de E dans F. L'application définie par*

$$\|\ell\|_{\mathcal{L}(E,F)} \overset{\text{déf}}{=} \sup_{\|x\|_E\leq 1} \|\ell(x)\|_F$$

est une norme sur $\mathcal{L}(E,F)$.

Convention: Sauf mention expresse du contraire, on considérera toujours l'espace $\mathcal{L}(E,F)$ muni de la norme $\|\ell\|_{\mathcal{L}(E,F)}$.

Démonstration. Il est très facile de démontrer que $\mathcal{L}(E,F)$ est un espace vectoriel. Vérifions les trois propriétés définissant une norme. Supposons que $\|\ell\|_{\mathcal{L}(E,F)} = 0$. Ceci implique que $\ell(x) = 0$ pour tout x de la boule unité (c'est ainsi que l'on nomme usuellement la boule de centre 0 et de rayon 1). Alors, pour tout $x \neq 0$,

$$\ell\left(\frac{x}{2\|x\|_E}\right) = \frac{1}{2\|x\|_E}\ell(x) = 0$$

On en déduit que $\ell(x) = 0$ et donc que $\ell = 0$.

De plus, soient $\ell \in \mathcal{L}(E,F)$ et $\lambda \in \mathbb{K}$, on a $\|\lambda\ell(x)\|_F = |\lambda| \times \|\ell(x)\|_F$. D'où il résulte que

$$\sup_{\|x\|_E\leq 1} \|\lambda\ell(x)\|_F = |\lambda| \sup_{\|x\|_E\leq 1} \|\ell(x)\|_F.$$

Soient ℓ_1 et ℓ_2 deux éléments de $\mathcal{L}(E,F)$, on a

$$\|\ell_1(x) + \ell_2(x)\|_F \leq \|\ell_1(x)\|_F + \|\ell_2(x)\|_F$$

Ainsi donc, pour tout x de E de norme inférieure à 1, on a

$$\|\ell_1(x) + \ell_2(x)\|_F \leq \|\ell_1\|_{\mathcal{L}(E,F)} + \|\ell_2\|_{\mathcal{L}(E,F)}.$$

La borne supérieure étant le plus petit des majorants, on en déduit que

$$\|\ell_1 + \ell_2\|_{\mathcal{L}(E,F)} \leq \|\ell_1\|_{\mathcal{L}(E,F)} + \|\ell_2\|_{\mathcal{L}(E,F)}.$$

D'où la proposition. □

Remarques On a les deux relations suivantes (à démontrer en exercice), pour $\ell \in \mathcal{L}(E,F)$,

$$\|\ell\|_{\mathcal{L}(E,F)} = \sup_{\|x\|_E=1} \|\ell(x)\|_F = \sup_{x\neq 0} \frac{\|\ell(x)\|_F}{\|x\|_E} \cdot \tag{5.4}$$

Lorsque $E = F$, on désigne $\mathcal{L}(E,E)$ par $\mathcal{L}(E)$.

Proposition 5.4.2. *Soient E, F et G trois espaces vectoriels normés et (ℓ_1,ℓ_2) un élément de $\mathcal{L}(E,F)\times\mathcal{L}(F,G)$. La composée $\ell_2\circ\ell_1$ appartient à $\mathcal{L}(E,G)$ et*

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

Démonstration. Pour démontrer cela, il suffit d'observer que, grâce à (5.4), on a

$$\begin{aligned}\|\ell_2\circ\ell_1(x)\|_G &\leq \|\ell_2\|_{\mathcal{L}(F,G)}\|\ell_1(x)\|_F\\ &\leq \|\ell_2\|_{\mathcal{L}(F,G)}\|\ell_1\|_{\mathcal{L}(E,F)}\|x\|_E.\end{aligned}$$

Remarques Si $E = F = G$, on pose alors $\mathcal{L}(E) \stackrel{\text{déf}}{=} \mathcal{L}(E,E)$. Pour tout $(\ell_1,\ell_2)\in\mathcal{L}(E)^2$,

$$\|\ell_2\circ\ell_1\|_{\mathcal{L}(E)} \leq \|\ell_1\|_{\mathcal{L}(E)}\|\ell_2\|_{\mathcal{L}(F)}. \tag{5.5}$$

La proposition suivante va nous fournir un nouvel exemple très important d'espaces de Banach.

Proposition 5.4.3. *Soient E un espace normé et F un espace de Banach. Alors l'espace $\mathcal{L}(E,F)$ muni de la norme définie dans la proposition 5.4.1 est un espace de Banach.*

Démonstration. Considérons une suite de Cauchy $(\ell_n)_{n\in\mathbb{N}}$ de $\mathcal{L}(E,F)$. D'après la définition de la norme sur l'espace $\mathcal{L}(E,F)$, on en déduit que, pour tout x de E, la suite $(\ell_n(x))_{n\in\mathbb{N}}$ est une suite de Cauchy de F. L'espace F étant supposé complet, il existe, pour tout x de E, un élément de F, noté $\ell(x)$ tel que

$$\lim_{n\to\infty}\ell_n(x) = \ell(x).$$

Il suffit maintenant de démontrer que ℓ appartient à $\mathcal{L}(E,F)$ et que

$$\lim_{n\to\infty}\ell_n = \ell \quad \text{dans} \quad \mathcal{L}(E,F).$$

L'unicité de la limite assure très facilement que ℓ est une application linéaire. Comme la suite $(\ell_n)_{n\in\mathbb{N}}$ est de Cauchy, elle est bornée (proposition 2.2.7). Il existe donc une constante C telle que

$$\sup_{\substack{n\in\mathbb{N}\\ \|x\|_E\leq 1}} \|\ell_n(x)\|_F \leq C.$$

Par passage à la limite, on en déduit que

$$\sup_{\|x\|_E\leq 1} \|\ell(x)\|_F \leq C.$$

Donc, l'application linéaire ℓ est continue. Démontrons la convergence de la suite $(\ell_n)_{n\in\mathbb{N}}$ vers ℓ dans $\mathcal{L}(E,F)$. La suite $(\ell_n)_{n\in\mathbb{N}}$ étant de Cauchy, pour tout ε strictement positif, il existe un entier n_0 tel que l'on ait, pour tout entier $n\geq n_0$,

$$\sup_{\substack{\|x\|_E\leq 1\\ p\in\mathbb{N}}} \|\ell_n(x)-\ell_{n+p}(x)\|_F < \varepsilon.$$

Par passage à la limite lorsque p tend vers l'infini, on obtient

$$\sup_{\|x\|_E \leq 1} \|\ell_n(x) - \ell(x)\|_F \leq \varepsilon,$$

ce qui achève la démonstration de la proposition. $\square$

Exercice 5.4.4. *Soient E et F deux espaces normés tels que F soit complet. On considère une suite bornée $(\ell_n)_{n\in\mathbb{N}}$ dans $\mathcal{L}(E,F)$. On suppose qu'il existe un sous-espace vectoriel G de E, dense dans E et une application linéaire ℓ de G dans F telle que*

$$\forall x \in G\,,\ \lim_{n\to\infty} \ell_n(x) = \ell(x).$$

Démontrez que l'on peut prolonger ℓ en un élément $\widetilde{\ell}$ de $\mathcal{L}(E,F)$ et que

$$\forall x \in E\,,\ \lim_{n\to\infty} \ell_n(x) = \widetilde{\ell}(x).$$

Exercice 5.4.5. *Soient E et F deux espaces vectoriels normés tels que E soit un espace de Banach. Si $\imath$ est une isométrie de E dans F, alors $\imath(E)$ est fermé dans F.*

5.5 Le cas particulier de $\mathcal{L}(E)$

L'ensemble $\mathcal{L}(E)$ est un espace de Banach ; la composition de deux éléments ℓ_0, ℓ_1 définit un nouvel élément que l'on notera simplement $\ell_0\ell_1$; on notera $\ell^n \stackrel{\text{déf}}{=} \ell\circ\cdots\circ\ell$ (n facteurs). D'après ce qui précède on a l'inégalité

$$\|\ell_0\ell_1\|_{\mathcal{L}(E)} \leq \|\ell_0\|_{\mathcal{L}(E)}\|\ell_1\|_{\mathcal{L}(E)}, \tag{5.6}$$

ce que l'on résume en disant que $\mathcal{L}(E)$ est une algèbre normée. Remarquons que cette inégalité dit en particulier que l'application $\ell \mapsto \ell_0\ell$, qui est évidemment linéaire, est continue (elle appartient à $\mathcal{L}(\mathcal{L}(E))$!).

L'un des résultats de base sur les éléments inversibles de $\mathcal{L}(E)$ est le suivant.

Théorème 5.5.1. *Soit E un espace de Banach; l'ensemble des éléments de $\mathcal{L}(E)$ qui sont à distance strictement inférieure à* 1 *de* Id *sont inversibles dans $\mathcal{L}(E)$. Dit autrement,*

$$\forall \ell \in B_{\mathcal{L}(E)}(\mathrm{Id},1)\,,\ \exists!\, \ell^{-1}\in\mathcal{L}(E)\,/\ \ell\circ\ell^{-1} = \ell^{-1}\circ\ell = \mathrm{Id}\,.$$

Démonstration. Posons

$$S_N = \sum_{n=0}^{N} (-1)^n (\ell - \mathrm{Id})^n.$$

D'après la proposition 5.1.6 page 49, la suite $(S_N)_{N\in\mathbb{N}}$ converge vers un élément ℓ^{-1} de $\mathcal{L}(E)$ dès que $\|\ell - \mathrm{Id}\,\|_{\mathcal{L}(E)} < 1$. Par ailleurs, on a

$$\begin{aligned} S_N\ell &= S_N(\mathrm{Id} + \ell - \mathrm{Id}) \\ &= \mathrm{Id} - (-1)^{N+1}(\ell - \mathrm{Id})^{N+1} \\ &= \ell S_N. \end{aligned}$$

En passant à la limite dans l'égalité ci-dessus, on conclut la démonstration de ce théorème. $\square$

Nous allons donner un exemple d'application de ce théorème à la théorie des équations différentielles ordinaires linéaires.

Soit I un intervalle de $\mathbb{R}$, on désigne par $\mathcal{C}(I,\mathbb{R}^N)$ l'ensemble des fonctions continues de I dans $\mathbb{R}^N$. On se donne une fonction continue A de I dans $\mathcal{L}(\mathbb{R}^N)$ et l'on fixe un réel t_0 dans l'intervalle I. On définit alors, pour tout réel λ, l'espace F_λ des fonctions f de $\mathcal{C}(I,\mathbb{R}^N)$ telles que

$$\|f\|_\lambda \stackrel{\text{déf}}{=} \sup_{t\in I} e^{-\lambda \int_{[t_0,t]} \|A(t')\|_{\mathcal{L}(\mathbb{R}^N)}dt'} \|f(t)\| < \infty.$$

Nous laissons en exercice la démonstration du fait que l'espace F_λ muni de la norme $\|\cdot\|_\lambda$ est un espace de Banach. De plus, remarquons que toute fonction constante appartient à F_λ. Démontrons la proposition suivante.

Proposition 5.5.2. *Soit $\mathcal{A}$ l'application linéaire définie par*

$$\mathcal{A}\left\{\begin{array}{ccl} F_\lambda & \longrightarrow & F_\lambda \\ f & \longmapsto & \mathcal{A}(f):\ t\mapsto \displaystyle\int_{t_0}^t A(t')f(t')dt' \end{array}\right.$$

Alors, pour tout réel strictement positif λ, $\mathcal{A}$ appartient à $\mathcal{C}(E_\lambda)$ et vérifie

$$\|\mathcal{A}\|_{\mathcal{L}(E_\lambda)} \leq \frac{1}{\lambda}\cdot$$

Démonstration. La démonstration du fait que $\mathcal{A}$ soit linéaire est laissée en exercice au lecteur. Pour tout t dans I (on suppose $t\geq t_0$ pour simplifier l'écriture) , on a

$$\begin{aligned} e^{-\lambda \int_{[t_0,t]} \|A(t')\|_{\mathcal{L}(\mathbb{R}^N)}dt'} \Big\|\int_{t_0}^t A(t')f(t')ddt'\Big\| & \leq e^{-\lambda\int_{t_0}^t \|A(\tau)\|_{\mathcal{L}(\mathbb{R}^N)}d\tau} \int_{t_0}^t \|A(t')f(\tau)\|d\tau \\ & \leq e^{-\lambda\int_{t_0}^t \|A(\tau)\|_{\mathcal{L}(\mathbb{R}^N)}d\tau} \int_{t_0}^t \|A(\tau)\|_{\mathcal{L}(\mathbb{R}^N)}\|f(\tau)\|d\tau \\ & \leq \int_{t_0}^t e^{-\lambda\int_{t'}^t \|A(t'')\|_{\mathcal{L}(\mathbb{R}^N)}dt''}\|A(t')\|_{\mathcal{L}(\mathbb{R}^N)} \\ & \qquad\qquad \times e^{-\lambda\int_{t_0}^\tau \|A(t'')\|_{\mathcal{L}(\mathbb{R}^N)}dt''}\|f(t')\|dt'. \end{aligned}$$

Par définition de la norme $\cdot\|_\lambda$, on a

$$\begin{aligned} \|\mathcal{A}(f)\|_\lambda & \leq \|f\|_\lambda \sup_{t\in I\cap[t_0,\infty[} \int_{t_0}^t e^{-\lambda\int_\tau^t \|A(t'')\|_{\mathcal{L}(\mathbb{R}^N)}dt''}\|A(t')\|_{\mathcal{L}(\mathbb{R}^N)}dt' \\ & \leq \frac{1}{\lambda}\|f\|_\lambda. \end{aligned}$$

Ceci démontre la proposition. $\square$

D'après le théorème 5.5.1, pour tout x_0 de $\mathbb{R}^N$, il existe une unique fonction x de F_λ telle que

$$\forall t\in I\,,\quad x(t)-\mathcal{A}(x)(t)=x_0.$$

Comme la fonction x est continue, la fonction $\mathcal{A}(x)$ est dérivable et donc la fonction x aussi. On a ainsi démontré que pour tout x_0 de $\mathbb{R}^N$, il existe une unique fonction x sur I telle que

$$\dot{x}(t)=A(x(t)) \quad\text{et}\quad x(t_0)=x_0.$$

Corollaire 5.5.3. *L'ensemble $U(E)$ des éléments inversibles de $\mathcal{L}(E)$ est un ouvert et l'application* Inv *définie par*

$$\left\{\begin{array}{ccc} U(E) & \longrightarrow & U(E) \\ \ell & \longmapsto & \ell^{-1} \end{array}\right.$$

est continue.

Démonstration. Soit ℓ_0 un élément de $U(E)$. Considérons alors un élément ℓ de $\mathcal{L}(E)$ tel que

$$\|\ell - \ell_0\|_{\mathcal{L}(E)} < \frac{1}{\|\ell_0^{-1}\|_{\mathcal{L}(E)}}\cdot$$

D'après l'inégalité (5.5), il vient

$$\|\ell \circ \ell_0^{-1} - \mathrm{Id}\,\|_{\mathcal{L}(E)} < 1.$$

D'après le théorème précédent, $\ell \circ \ell_0^{-1}$ est inversible, donc ℓ aussi. L'ensemble $U(E)$ est donc ouvert.

Démontrons la continuité sur la boule (ouverte) de centre Id et de rayon 1. On peut écrire, pour deux éléments ℓ_1 et ℓ_2 de la boule de centre Id et de rayon α strictement inférieur à 1, que

$$\ell_1^{-1} - \ell_2^{-1} = \sum_{n=0}^{\infty}(-1)^n((\ell_1 - \mathrm{Id})^n - (\ell_2 - \mathrm{Id})^n).$$

Démontrons que, si a et b soont deux éléments de $\mathcal{L}(E)$, alors

$$\forall n \geq 2\,,\ \|a^n - b^n\|_{\mathcal{L}(E)} \leq n\|a-b\|_{\mathcal{L}(E)} \sum_{m=0}^{n-1} \|a\|_{\mathcal{L}(E)}^{n-1-m}\|b\|_{\mathcal{L}(E)}^m. \tag{5.7}$$

Commençons par le démontrer pour $n = 2$. On a

$$a^2 - b^2 = (a-b)a - b(b-a)$$

D'après (5.6), on a

$$\|a^2 - b^2\|_{\mathcal{L}(E)} \leq \|a-b\|_{\mathcal{L}(E)}(\|a\|_{\mathcal{L}(E)} + \|b\|_{\mathcal{L}(E)})$$

ce qui est exactement (5.7) dans le cas où $n = 2$. Supposons (5.7) pour un entier $n \geq 2$. On peut écrire que

$$a^{n+1} - b^{n+1} = (a-b)a^n + b(a^n - b^n).$$

En utilisant (5.6) et l'hypothèse de récurrence, on a

$$\begin{aligned} \|a^{n+1} - b^{n+1}\|_{\mathcal{L}(E)} &\leq \|(a-b)\|_{\mathcal{L}(E)}\|a\|_{\mathcal{L}(E)}^n + \|b\|_{\mathcal{L}(E)}\|a^n - b^n)\|_{\mathcal{L}(E)} \\ &\leq \|(a-b)\|_{\mathcal{L}(E)}\|a\|_{\mathcal{L}(E)}^n + \|b\|_{\mathcal{L}(E)} \sum_{m=0}^{n-1} \|a\|_{\mathcal{L}(E)}^{n-1-m}\|b\|_{\mathcal{L}(E)}^m \end{aligned}$$

ce qui est exactement (5.7) pour $n+1$.

Comme nous avons supposé que ℓ_1 et ℓ_2 appartenaient à la boule de centre Id et de rayon α, on en déduit que

$$\|(\ell_1 - \mathrm{Id})^n - (\ell - \mathrm{Id})^n\|_{\mathcal{L}(E)} \leq n\alpha^{n-1}\|\ell_1 - \ell_2\|_{\mathcal{L}(E)}$$

Comme α est supposé strictement inférieur à 1, on a

$$\|\ell_1^{-1} - \ell_2^{-1}\|_{\mathcal{L}(E)} \leq C\|\ell_1 - \ell_2\|_{\mathcal{L}(E)}.$$

Pour achever la démonstration du corollaire, la continuité étant une propriété locale, il suffit de démontrer, pour chaque $\ell_0 \in U(E)$, sur une boule ouverte centrée en ℓ_0. Supposons donc que ℓ_0 appartient à $U(E)$ et que $\ell \in B(\ell_0, \|\ell_0^{-1}\|^{-1})$. Écrivons

$$\ell\ell_0^{-1} - \mathrm{Id} = \ell\ell_0^{-1} - \ell_0\ell_0^{-1} = (\ell - \ell_0)\ell_0^{-1}.$$

Comme on a supposé que $\ell \in B(\ell_0, \|\ell_0^{-1}\|^{-1})$, on a

$$\|\ell\ell_0^{-1} - \mathrm{Id}\,\| \leq \|\ell - \ell_0\|\|\ell_0^{-1}\| < 1.$$

Ainsi donc $\ell\ell_0^{-1}$ est inversible et

$$\ell\left(\ell_0^{-1}((\ell\ell_0^{-1})^{-1})\right) = (\ell\ell_0^{-1})(\ell\ell_0^{-1})^{-1} = \mathrm{Id}\,.$$

Ceci permet d'écrire $\ell^{-1} = \ell_0^{-1}((\ell\ell_0^{-1})^{-1}$. Autrement dit, l'application $\ell \mapsto \ell^{-1}$, définie sur la boule $B(\ell_0, \|\ell_0^{-1}\|^{-1})$, peut s'obtenir en composant les applications $\ell \mapsto \ell\ell_0^{-1}$ (sur cette même boule), $u \mapsto u^{-1}$ (sur la boule $B(\mathrm{Id}, 1)$), et $v \mapsto \ell_0^{-1}v$. La première et la dernière sont des applications linéaires continues (voir le premier paragraphe de cette section 5.5), et la seconde est continue d'après ce qui précède. L'application $\ell \mapsto \ell^{-1}$ est donc sur la boule $B(\ell_0, \|\ell_0^{-1}\|^{-1})$. Le corollaire 5.5.3 est démontré. $\square$

ATTENTION ! En dimension finie, il n'existe qu'une seule façon, pour une application linéaire, de ne pas être inversible. En effet, soit ℓ une application linéaire de E dans E, l'espace E étant supposé être de dimension finie. L'algèbre linéaire élémentaire dit que

$$\ell \text{ bijective} \Leftrightarrow \ell \text{ surjective} \Leftrightarrow \ell \text{ injective}.$$

L'application réciproque ℓ^{-1} est linéaire donc continue puisque nous sommes en dimension finie.

Il en va tout autrement en dimension infinie. Par exemple, soit $E = \ell^\infty(\mathbb{N})$ muni de la norme

$$\|(x(n))_{n\in\mathbb{N}}\|_{\ell^\infty(\mathbb{N})} \stackrel{\text{déf}}{=} \sup_{n\in\mathbb{N}} |x(n)|.$$

C'est un espace de Banach. Soit ℓ_1 l'application linéaire définie par

$$\ell_1 \left\{ \begin{array}{ccl} E & \longrightarrow & E \\ (x(n))_{n\in\mathbb{N}} & \longmapsto & (y(n)_{n\in\mathbb{N}} \,/\, y(n) = x(n+1). \end{array} \right.$$

L'application ℓ_1 est surjective, mais pas injective. Considérons ℓ_2 définie par

$$\ell_2 \left\{ \begin{array}{ccl} E & \longrightarrow & E \\ (x(n))_{n\in\mathbb{N}} & \longmapsto & (y(n)_{n\in\mathbb{N}} \,/\, y(n) = x(n-1) \quad \text{si} \quad n \geq 1\,,\ 0 \quad \text{sinon}. \end{array} \right.$$

L'application ℓ_2 est une isométrie, c'est-à-dire que $\|\ell_2 x\| = \|x\|$, mais elle n'est pas surjective.

5.6 Le théorème de Baire

Le théorème suivant est une conséquence du théorème de Baire (théorème 2.2.3 page 26).

Théorème 5.6.1 (de Banach-Steinhaus)**.** *Soient E et F deux espaces normés et $(\ell_n)_{n\in\mathbb{N}}$ une suite de $\mathcal{L}(E,F)$. Supposons que E soit complet. Supposons que, pour tout x de E, la limite de la suite $(\ell_n(x))_{n\in\mathbb{N}}$ existe; désignons la par $\ell(x)$.*

Alors la suite $(\ell_n)_{n\in\mathbb{N}}$ est une suite bornée de $\mathcal{L}(E,F)$, ce qui implique en particulier que ℓ appartient à $\mathcal{L}(E,F)$. De plus, on a

$$\|\ell\|_{\mathcal{L}(E,F)} \leq \liminf_{n\to\infty} \|\ell_n\|_{\mathcal{L}(E,F)}.$$

Démonstration. La démonstration de ce théorème repose sur le lemme suivant.

Lemme 5.6.2. *Soient E et F deux espaces normés et $(\ell_n)_{n\in\mathbb{N}}$ une suite de $\mathcal{L}(E,F)$. Supposons que E soit complet. Supposons que, pour tout x de E, la suite $(\ell_n(x))_{n\in\mathbb{N}}$ soit une suite bornée de F.*

Alors la suite $(\ell_n)_{n\in\mathbb{N}}$ est une suite bornée de $\mathcal{L}(E,F)$.

Démonstration. Considérons les ensembles $F_{n,p}$ définis par

$$F_{n,p} = \{x \in E \,/\, \|\ell_n(x)\|_F \leq p\}.$$

Ces ensembles $F_{n,p}$ sont des fermés en tant qu'image réciproque de fermés par une application continue. Donc les ensembles F_p définis par

$$F_p \stackrel{\text{déf}}{=} \bigcap_{n\in\mathbb{N}} F_{n,p}$$

sont des ensembles fermés car ce sont des intersections de fermés. De plus, soit x un élément quelconque de E. La suite $(\ell_n(x))_{n\in\mathbb{N}}$ est supposée bornée. Donc, il existe un entier p tel que x appartienne à F_p. Cela signifie que la réunion de tous les F_p est l'espace E tout entier. D'après le corollaire 2.2.6 page 27 du théorème 2.2.3 de Baire, il existe un entier p_0 tel que $\overset{\circ}{F}_{p_0} \neq \emptyset$. Donc, il existe un point x_0 et un réel strictement positif α tel que $B(x_0,\alpha)$ soit inclus dans F_{p_0}. On a donc

$$\begin{aligned}
\sup_{\substack{n\in\mathbb{N}\\ \|x\|\leq\alpha}} \|\ell_n(x)\|_F &\leq \sup_{\substack{n\in\mathbb{N}\\ \|x\|\leq\alpha}} \|\ell_n(x+x_0)\|_F + \|\ell_n(x_0)\|_F\\
&\leq \sup_{\substack{n\in\mathbb{N}\\ x\in B(x_0,\alpha)}} \|\ell_n(x)\|_F + \|\ell_n(x_0)\|_F\\
&\leq 2p_0.
\end{aligned}$$

Ainsi, pour tout x de E de norme plus petite que 1, on peut écrire

$$\begin{aligned}
\|\ell_n(x)\|_F &\leq \alpha^{-1}\left\|\ell_n\left(\frac{x}{\alpha}\right)\right\|_F\\
&\leq \frac{2p_0}{\alpha}\cdot
\end{aligned}$$

Donc la suite $(\ell_n)_{n\in\mathbb{N}}$ est une suite bornée de $\mathcal{L}(E,F)$. □

Conclusion de la démonstration du théorème 5.6.1 Par passage à la limite, on obtient alors que,

$$\forall x \in B(x_0,\alpha)\,,\ \|\ell(x)\| \leq 2p_0.$$

Démontrons maintenant la majoration de la norme de ℓ. Soit x un élément de E de norme 1. On peut alors écrire

$$\begin{aligned}
\|\ell(x)\|_F &= \lim_{n\to\infty} \|\ell_n(x)\|_F\\
&= \liminf_{n\to\infty} \|\ell_n(x)\|_F\\
&\leq \liminf_{n\to\infty} \|\ell_n\|_{\mathcal{L}(E,F)} \quad (\text{car } \|x\|_E = 1).
\end{aligned}$$

Le théorème est ainsi complètement démontré. □

Remarque L'inégalité majorant la norme de ℓ peut être stricte comme le montre l'exemple suivant.

On considère l'espace $\ell^1(\mathbb{N})$ des suites à valeurs complexes absolument convergentes. C'est bien sûr un espace de Banach. On considère la suite de formes linéaires $(e_n)_{n\in\mathbb{N}}$ définie par

$$e_n((x_p)_{p\in\mathbb{N}}) \stackrel{\text{déf}}{=} x_n.$$

C'est un exercice facile de démontrer que

$$\forall x \in \ell^1(\mathbb{N})\,,\ \lim_{n\to\infty} e_n(x) = 0.$$

Il est très facile d'observer que $\|\ell_n\|_{\mathcal{L}(E,\mathbb{C})} = 1$.

Comme application du théorème de Baire, on peut aussi citer le théorème suivant, que nous admettrons.

Théorème 5.6.3 (de Banach)**.** *Soient E et F deux espaces de Banach et A un élément de $\mathcal{L}(E,F)$. Si A est bijective, alors $A^{-1} \in \mathcal{L}(F,E)$.*

Chapitre 6

Calcul différentiel: les bases

6.1 Différentielle et dérivées partielles

Commençons par rappeler la notion de dérivée d'une fonction d'une variable réelle.

Définition 6.1.1. *Soit f une fonction d'un intervalle ouvert I de $\mathbb{R}$ à valeurs dans un espace normé F. Soit t_0 un point de I, on dit que f est dérivable en t_0 si et seulement si*

$$\lim_{t \to t_0} \frac{f(t) - f(t_0)}{t - t_0}$$

existe. On l'appelle alors le vecteur dérivé de f au point t_0 ; ce vecteur est noté $f'(t_0)$.

Si la fonction f est dérivable en tout point t de l'intervalle I, on dit alors que la fonction f est dérivable sur I.

Si la fonction f' est continue, la fonction f est dit que classe C^1 sur I (on note $C^1(I)$ l'ensemble de toutes les fonctions de classe $C^1(I)$).

Remarquons qu'une fonction dérivable en un point est continue en ce point.

Nous allons étendre cette notion aux fonctions dites "de plusieurs variables" (sous entendues réelles) c'est-à-dire en fait aux fonctions d'une variable appartenant à un espace vectoriel normé pouvant être de dimension finie ou non. Plus précisement, nous considérerons des fonctions définies sur un ouvert Ω d'un espace vectoriel normé $(E, \|\cdot\|_E)$ à valeurs dans un espace vectoriel normé $(F, \|\cdot\|_F)$.

Définition 6.1.2. *Soient f une fonction de Ω dans E à valeurs dans F et a un point de Ω. On dit que la fonction f est différentiable au point a si et seulement si, il existe une application linéaire L continue de E dans F telle que, pour tout réel strictement positif ε, il existe un réel strictement positif α tel que la boule ouverte $B(a, \alpha)$ de centre a et de rayon α soit incluse dans Ω et tel que*

$$\forall \overrightarrow{h} \in B(0, \alpha)\,,\ \|f(a + \overrightarrow{h}) - f(a) - L \cdot \overrightarrow{h}\|_F < \varepsilon \|\overrightarrow{h}\|_E.$$

Si une fonction est différentiable en tout point d'un ouvert Ω, elle est dite "différentiable sur Ω".

Proposition 6.1.3. *Soit L une application linéaire conitnue de E dans F. Si*

$$\forall \varepsilon > 0\,,\ \exists \alpha > 0\,/\ \|\overrightarrow{h}\|_E < \alpha \Longrightarrow \|L \cdot \overrightarrow{h}\|_F < \varepsilon \|\overrightarrow{h}\|_E\,,$$

alors l'application linéaire L est nulle.

Démonstration. Elle utilise l'importante idée qui est celle d'homogénéïté. Soit $\overrightarrow{h}_0$ un vecteur non nul de E. Par hypothèse, pour tout ε strictement positif, il existe un réel strictement positif α tel que

$$\Big\|L\Big(\frac{\alpha}{2}\frac{\overrightarrow{h}_0}{\|\overrightarrow{h}_0\|_E}\Big)\Big\|_F < \varepsilon\frac{\alpha}{2}\cdot$$

D'après la linéarité de L et l'homogénéïté de la norme, ceci s'écrit

$$\forall\varepsilon > 0\,,\ \|L\overrightarrow{h}_0\|_F < \varepsilon\|\overrightarrow{h}_0\|_E.$$

Ceci étant vrai pour tout ε strictement positif, on a donc que $L\cdot\overrightarrow{h}_0 = 0$ et donc que l'application linéaire L est identiquement nulle. $\square$

Remarque La proposition 6.1.3 ci-dessus implique que l'application linéaire L de la définition ci-dessus est unique. On l'appelle la différentielle de f au point a et on le notera $Df(a)$.

Donnons tout d'abord quelques exemples (triviaux ou bien connus) de fonctions différentiables.

- Si $E = \mathbb{R}$, il ne s'agit que de la notion usuelle de dérivée. En effet, qu'est-ce qu'une application linéaire de $\mathbb{R}$ dans F? On peut la décrire par son image en 1. Plus précisement, soit I l'application linéaire définie par

$$I\left\{\begin{array}{ccc}\mathcal{L}(\mathbb{R};F) & \longrightarrow & F\\ A & \longmapsto & A(1).\end{array}\right.$$

  Il est très facile de démontrer que l'application linéaire I est une bijection.

- Les fonctions constantes $x \mapsto b \in F$ sont différentiables et leur différentielle est nulle.

- Les fonctions linéaires continues sont différentiables car $L(a+\overrightarrow{h}) = L(a) + L\cdot\overrightarrow{h}$

- Soit E un espace vectoriel de dimension finie et $(\overrightarrow{e}_j)_{1\leq j\leq N}$. On considère comme norme

$$N(x) \stackrel{\text{déf}}{=} \left(\sum_{j=1}^N x_j^2\right)^{\frac{1}{2}} \quad\text{avec}\quad x = \sum_{j=1}^N x_j\overrightarrow{e}_j.$$

  La fonction $f(x) \stackrel{\text{déf}}{=} N^2(x)$ est différentiable sur E et l'on a, pour tout point a de E,

$$Df(a)\cdot\overrightarrow{h} = 2(a|\overrightarrow{h}) \quad\text{avec}\quad (x|y) \stackrel{\text{déf}}{=} \sum_{j=1}^N x_jy_j.$$

  En effet, on a

$$N^2(a+\overrightarrow{h}) = N^2(a) + 2(a|\overrightarrow{h}) + N^2(\overrightarrow{h}).$$

Proposition 6.1.4. *Soit f une application de Ω dans $\mathbb{R}^N$ et a un point de Ω. Si f est différentiable au point a, alors f est continue au point a.*

Démonstration. Par définition de la différentiabilité, pour tout ε strictement positif, il existe un réel strictement positif α (que l'on peut supposer inférieur à $1/2$) tel que

$$\|\overrightarrow{h}\|_E < \alpha \Longrightarrow \|f(a+\overrightarrow{h}) - f(a) - Df(a)\cdot\overrightarrow{h}\|_F < \varepsilon\|\overrightarrow{h}\|_E.$$

L'inégalité triangulaire implique et la continuité de l'application linéaire L implique que

$$\begin{aligned}\|f(a+\overrightarrow{h})-f(a)\|_F &\leq \|Df(a)\cdot\overrightarrow{h}\|_F+\varepsilon\|\overrightarrow{h}\|_E\\ &\leq \|Df(a)\|_{\mathcal{L}(E,F)}\|\overrightarrow{h}\|_E+\frac{\varepsilon}{2}\cdot\end{aligned}$$

Ainsi donc

$$\|\overrightarrow{h}\|_E<\min\Big\{\alpha,\frac{\varepsilon}{2\|Df(a)\|_{\mathcal{L}(E,F)}+1}\Big\}\Longrightarrow\|f(a+\overrightarrow{h})-f(a)\|_F<\varepsilon.$$

ce qui conclut la démonstration. □

Remarque La réciproque est fausse. L'exemple typique est la fonction valeur absolue de $\mathbb{R}$ dans $\mathbb{R}$. Cette fonction est continue de $\mathbb{R}$ dans $\mathbb{R}$ mais n'est pas dérivable en 0 (démontrez-le!).

6.2 Dérivée directionnelle et dérivées partielles

La différentielle d'une fonction f prend en compte les variations de celle-ci dans toutes les directions de l'espace de départ. Nous n'allons prendre en compte ici que les variations de la fonction f dans une direction de l'espace de départ. Ceci repose la proposition suivante, simple mais fondamentale.

Proposition 6.2.1. *Soient Ω un ouvert de $(E,\|\cdot\|_E)$ et f une application de Ω dans $(F,\|\cdot\|_F)$ différentiable en un point a de Ω. Soient α_0 un réel strictement positif et $\overrightarrow{h}_0$ un vecteur non nul de E tels que le segment $]a-\alpha_0\overrightarrow{h}_0,a+\alpha_0\overrightarrow{h}_0[$ soit inclus dans Ω. Alors l'application $f_{a,\overrightarrow{h}_0}$ définie par*

$$\begin{cases}]\alpha_0,\alpha_0[& \longrightarrow & F\\ t & \longmapsto & f(a+t\overrightarrow{h}_0)\end{cases}$$

est dérivable en 0 et l'on a

$$\frac{d}{dt}f_{a,\overrightarrow{h}_0|t=0}=Df(a)\overrightarrow{h}_0.$$

Démonstration. Comme f est différentiable en a, on a

$$\forall\varepsilon>0\,,\ \exists\beta>0\,/\ \|\overrightarrow{h}\|_E<\beta\Longrightarrow\|f(a+\overrightarrow{h})-f(a)-Df(a)\overrightarrow{h}\|_F<\frac{\varepsilon}{\|\overrightarrow{h}_0\|_E}\cdot$$

En appliquant cette inégalité avec $\overrightarrow{h}=t\overrightarrow{h}_0$, on trouve que

$$\forall\varepsilon>0\,,\ \exists\beta>0\,/\ |t|<\frac{\beta}{\|\overrightarrow{h}_0\|_E}<\Longrightarrow\|f(a+t\overrightarrow{h}_0)-f(a)-tDf(a)\overrightarrow{h}_0\|_F<\varepsilon|t|$$

ce qui signifie exactement que $f_{a,\overrightarrow{h}_0}$ est dérivable en 0 et que

$$\frac{d}{dt}f_{a,\overrightarrow{h}_0|t=0}=Df(a)\overrightarrow{h}_0.$$

D'où la proposition. □

Plaçons nous maintenant dans le cas où l'espace de départ E est $\mathbb{R}^N$ La notion de dérivées partielles est liée au choix d'une base de l'espace de départ. Considérons $(\overrightarrow{e}_j)_{1\leq j\leq n}$ une base de $\mathbb{R}^N$. On choisit en général la base canonique de $\mathbb{R}^N$. Les dérivées partielles de la fonction f au point a relatives à la base $(\overrightarrow{e}_j)_{1\leq j\leq n}$ sont les vecteurs $L\cdot\overrightarrow{e}_j$ de $\mathbb{R}^N$.

Soit $\overrightarrow{h}$ un vecteur de $\mathbb{R}^N$. On peut le décomposer dans la base $(\overrightarrow{e}_j)_{1\leq j\leq n}$ en écrivant que

$$\overrightarrow{h}=\sum_{j=1}^n h^j\overrightarrow{e}_j.$$

On a alors

$$Df(a)\cdot\overrightarrow{h}=\sum_{j=1}^n h^jDf(a)\cdot\overrightarrow{e}_j. \tag{6.1}$$

Ceci s'écrit souvent d'une manière différente. Désignons par dx_j la forme linéaire j-ème coordonnée, c'est-à-dire la forme linéaire définie par

$$\langle dx_j,\overrightarrow{h}\rangle=h^j.$$

La famille $(dx_j)_{1\leq j\leq n}$ ainsi définie est la base duale de la base $(\overrightarrow{e}_j)_{1\leq j\leq n}$. La relation (6.1) ci-dessus peut s'écrire alors

$$Df(a)\cdot\overrightarrow{h}=\sum_{j=1}^n Df(a)\cdot\overrightarrow{e}_j\langle dx_j,\overrightarrow{h}\rangle. \tag{6.2}$$

Introduisons la notation, qui sera utilisée dans toute la suite,

$$\frac{\partial f}{\partial x_j}(a)\stackrel{\text{déf}}{=}Df(a)\cdot\overrightarrow{e}_j.$$

La relation (6.2) s'écrit alors

$$Df(a)=\sum_{j=1}^n\frac{\partial f}{\partial x_j}(a)dx_j.$$

Plaçons nous maintenant dans le cas où l'esapce d'arrivée F est $\mathbb{R}^M$. On peut dès lors utiliser les matrices . Soient $(\overrightarrow{e}_j)_{1\leq j\leq N}$ et $(\widetilde{e}_i)_{1\leq i\leq M}$ des bases de $\mathbb{R}^N$ et $\mathbb{R}^M$. La matrice de l'application linéaire $Df(a)$ dans ces bases est la matrice de terme général

$$A^i_j=\frac{\partial f^i}{\partial x_j}(a),$$

où f^i désigne la $j$ème composante de f dans la base $(\widetilde{e}_i)_{1\leq i\leq N}$. Le j ème vecteur colonne de la matrice $Df(a)$ sont le vecteur dérivé partiel $\dfrac{\partial f}{\partial x_j}(a)$.

ATTENTION La notion de dérivées partielles n'a de sens qu'une fois effectué le choix d'une base $(\overrightarrow{e}_j)_{1\leq j\leq N}$. La relation entre différentielle et dérivées partielles est très analogue à celle entre application linéaire et matrice.

Une remarque conclusive Comme le montre l'exemple suivant, posséder des dérivées partielles est une propriété beaucoup plus faible que d'être différentiable. Considérons la fonction de $\mathbb{R}^2$ dans $\mathbb{R}$ définie par

$$f(x,y)=\frac{xy}{x^2+y^2}\quad\text{si}\quad(x,y)\neq(0,0)\quad\text{et}\quad f(0,0)=0.$$

Cette fonction admet au point $(0,0)$ une dérivée suivant tout vecteur $\overrightarrow{h}_0$ non nul. Pourtant, cette fonction n'est pas différentiable en $(0,0)$ car elle n'y est pas continue. Cependant, nous avons le théorème suivant:

Théorème 6.2.2. *Soit f une fonction d'un ouvert Ω de $\mathbb{R}^N$ dans $(F, \|\cdot\|_F)$. Supposons que la fonction f admette des dérivées partielles en tout point de Ω. Soit a un point de Ω. Si les dérivées partielles de f sont des fonctions continues de Ω, alors la fonction f est différentiable en a.*

Démonstration. Elle nécessite l'inégalité des accroissements finis qui sera l'objet de la section 6.4. Elle sera faite en détail page 71. □

6.3 Composition d'applications différentiables

Le résultat principal de cette section est le théorème fondamental suivant.

Théorème 6.3.1. *Soient f une application d'un ouvert Ω d'un espace normé E dans un ouvert Ω' d'un espace normé F et g une application de Ω' dans un espace normé G et a un point de Ω. Si f est différentiable au point a et si g est différentiable au point $f(a)$, alors l'application $g \circ f$ est différentiable au point a et l'on a*

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

Avant de démontrer ce théorème, faisons quelques commentaires à son propos. La notion de fonction différentiable en un point peut être comprise de la manière suivante: une fonction différentiable en un point est une fonction qui est bien approximée près de ce point par une (unique) application linéaire. Ainsi donc, le théorème ci-dessus affirme que, si deux fonctions possèdent cette propriété, leur composée aussi et l'approximation de la composée est la composée des approximations, ce qui ne doit pas surprendre.

Démonstration du théorème 6.3.1. Posons

$$\Delta_a(\overrightarrow{h}) \stackrel{\text{déf}}{=} g(f(a+\overrightarrow{h})) - g(f(a)) - Dg(f(a)) \cdot (Df(a) \cdot \overrightarrow{h}.$$

Écrivons alors que $\Delta_a(\overrightarrow{h}) = \Delta_a^{(1)}(\overrightarrow{h}) + \Delta_a^{(2)}(\overrightarrow{h})$ avec

$$\begin{aligned} \Delta_a^{(1)}(\overrightarrow{h}) &\stackrel{\text{déf}}{=} g(f(a+\overrightarrow{h})) - g(f(a)) - Dg(f(a))(f(a+\overrightarrow{h}) - f(a)) \quad \text{et} \\ \Delta_a^{(2)}(\overrightarrow{h}) &\stackrel{\text{déf}}{=} Dg(f(a))\big(f(a+\overrightarrow{h}) - f(a) - Df(a) \cdot \overrightarrow{h}\big). \end{aligned}$$

Soit ε un réel positif quelconque. Il existe un réel α_0 strictement positif tel que

$$\|\overrightarrow{h}\|_E < \alpha_0 \Longrightarrow \big\|f(a+\overrightarrow{h}) - f(a) - Df(a) \cdot \overrightarrow{h}\big\|_F < \frac{\varepsilon \|\overrightarrow{h}\|_E}{2\|Dg(f(a))\|_{\mathcal{L}(F;G)} + 2} \cdot$$

Par définition de $\Delta_a^{(2)}$, on a

$$\|\overrightarrow{h}\|_E < \alpha_0 \Longrightarrow \|\Delta_a^{(2)}(\overrightarrow{h})\|_G < \frac{\varepsilon}{2} \|\overrightarrow{h}\|_E. \tag{6.3}$$

La fonction g étant différentiable en $f(a)$, il existe un réel strictement positif β tel que

$$\|\overrightarrow{k}\|_F < \beta \Longrightarrow \|g(f(a) + \overrightarrow{k}) - g(f(a) - Dg(f(a)) \cdot \overrightarrow{k}\|_G < \frac{\varepsilon \|\overrightarrow{k}\|_F}{2\|Df(a)\|_{\mathcal{L}(E;F)} + 2} \cdot \tag{6.4}$$

La fonction f étant différentiable, il existe un réel strictement positif α_1 tel que

$$\|\overrightarrow{h}\|_E < \alpha_1 \Longrightarrow \|f(a+\overrightarrow{h}) - f(a)\|_F < (\|Df(a)\|_{\mathcal{L}(E;F)} + 1)\|\overrightarrow{h}\|_E.$$

On déduit alors de (6.3) que

$$\|\overrightarrow{h}\|_E < \min\Big\{\alpha_0, \alpha_1, \frac{\beta}{\|Df(a)\|_{\mathcal{L}(E;F)} + 1}\Big\} \Longrightarrow \|\Delta_a(\overrightarrow{h})\|_G < \varepsilon\|\overrightarrow{h}\|_E.$$

D'où le théorème. □

Théorème 6.3.2. *Soient E et F deux espaces normés et f un homéomorphisme entre deux ouverts Ω_1 et Ω_2 de E et F respectivement. Soit a un point de Ω_1. Si la fonction f est différentiable au point a et si $Df(a)$ est une application linéaire inversible de E dans F, alors la fonction réciproque f^{-1} l'est au point $f(a)$ et l'on a*

$$Df^{-1}(f(a)) = (Df)^{-1}(a).$$

Démonstration. Ce théorème est très lié au théorème de composition. Remarquons que si f^{-1} est différentiable en $f(a)$ alors, comme $f^{-1} \circ f = \mathrm{Id}_E$, on a nécessairement

$$D(f^{-1})(f(a)) \circ Df(a) = \mathrm{Id}_E\,.$$

Posons

$$g(\overrightarrow{h}) \stackrel{\text{déf}}{=} (Df(a))^{-1}(f(a+\overrightarrow{h}) - f(a)).$$

Remarquons que g est un homéomorphisme d'un voisinage de 0 dans E sur un voisinage de 0, que d'après le théorème 6.3.1, g est une fonction différentiable en 0 et sa différentielle en ce point est Id_E. De plus, les fonctions f^{-1} et g^{-1} sont liées par les relations

$$f^{-1}(y) = a + g^{-1}\big(Df(a)^{-1}(y - f(a)\big) \quad \text{et} \quad g^{-1}(\overrightarrow{k}) = f^{-1}\big(f(a) + Df(a)\cdot\overrightarrow{k}\big) - a$$

Toujours d'après le théorème de composition, on est ramené à démontrer le théorème dans le cas où $E = F$, $a = f(a) = 0$ et $Df(a) = \mathrm{Id}_E$, c'est-à -dire que

$$\forall\varepsilon\,,\ \exists\alpha\,/\ \|\overrightarrow{k}\|_E < \alpha \Longrightarrow \|g^{-1}(\overrightarrow{k}) - \overrightarrow{k}\|_E < \varepsilon\|\overrightarrow{k}\|_E. \tag{6.5}$$

Pour ce faire, observons tout d'abord qu'il existe un réel α_0 tel que

$$\|\overrightarrow{h}\|_E < \alpha_0 \Longrightarrow \|g(\overrightarrow{h}) - \overrightarrow{h}\|_E < \frac{1}{2}\|\overrightarrow{h}\|_E.$$

L'inégalité triangulaire assure que, pour tout $\overrightarrow{h}$ dans la boule $B(0,\alpha_0)$ on a

$$\frac{1}{2}\|\overrightarrow{h}\|_E \leq \|g(\overrightarrow{h})\|_E \leq \frac{3}{2}\|\overrightarrow{h}\|_E. \tag{6.6}$$

Comme la fonction g est différentiable en 0 et que sa différentielle en ce point est Id_E, on a

$$\forall\varepsilon\,,\ \exists\alpha > 0\,,\ \|\overrightarrow{h}\|_E < \alpha \Longrightarrow \|g(\overrightarrow{h}) - \overrightarrow{h}\|_E < \frac{\varepsilon}{2}\|\overrightarrow{h}\|_E.$$

Si α est choisi assez petit, on peut appliquer cette inégalité avec $\overrightarrow{h} = g^{-1}(\overrightarrow{k})$. Il vient alors, d'après (6.6),

$$\forall\varepsilon\,,\ \exists\alpha > 0\,,\ \|g^{-1}(\overrightarrow{k})\|_E < \alpha \Longrightarrow \|\overrightarrow{k} - g^{-1}(\overrightarrow{k})\|_E < \frac{\varepsilon}{2}\|g^{-1}(\overrightarrow{k})\|_E < \varepsilon\|\overrightarrow{k}\|_E.$$

En utilisant à nouveau (6.6), on a

$$\|\overrightarrow{k}\|_E < \frac{\alpha}{2} \Longrightarrow \|g^{-1}(\overrightarrow{k})\|_E < \alpha.$$

Ainsi donc, on a

$$\|\overrightarrow{k}\|_E < \frac{\alpha}{2} \Longrightarrow \|\overrightarrow{k} - g^{-1}(\overrightarrow{k})\|_E < \varepsilon\|\overrightarrow{k}\|_E$$

et le théorème est démontré. □

6.4 L'inégalité des accroissements finis

Commençons par rappeler des résultats et des démonstrations très classiques. Nous allons maintenant supposer que l'espace vectoriel normé d'arrivée est $\mathbb{R}$.

Théorème 6.4.1. *Soit f un fonction dérivable d'un intervalle $I =]a,b[$ de $\mathbb{R}$ à valeurs dans $\mathbb{R}$, continue sur $[a,b]$. On suppose que $f(a) = f(b)$. Alors, il existe un point c de l'intervalle I tel que $f'(c) = 0$.*

Démonstration. Si la fonction f est constante, il n'y a rien à démontrer. Si f n'est pas constante, comme f est continue sur l'intervalle compact $[a,b]$, il existe un point c de I où f atteint son maximum ou son minimum. La dérivée est nulle en ce point. D'où le théorème. □

Corollaire 6.4.2. *Soit f une fonction continue sur un intervalle $[a,b]$ de $\mathbb{R}$ et dérivable sur $]a,b[$. Alors il existe c dans $]a,b[$ tel que*

$$f'(c) = \frac{f(b) - f(a)}{b - a} \cdot$$

Démonstration. Considérons la fonction

$$g(t) \stackrel{\text{déf}}{=} f(t) - f(a) - \frac{f(b) - f(a)}{b - a}(t - a)\,.$$

La fonction g est continue $[a,b]$ et dérivable sur $]a,b[$ et l'on a

$$g'(t) = f'(t) - \frac{f(b) - f(a)}{b - a} \cdot$$

De plus $g(a) = g(b) = 0$. Donc il existe un c de l'intervalle $]a,b[$ tel que $g'(c) = 0$. D'où le corollaire. □

Remarquons que ce corollaire entraine l'inégalité suivante: si la dérivée de f est bornée sur $]a,b[$, alors

$$|f(b) - f(a)| \leq \sup_{t\in]a,b[} |f'(t)|(b - a). \tag{6.7}$$

C'est cette inégalité qui se généralise pour les fonctions à valeurs dans un espace vectoriel normé quelconque.

Théorème 6.4.3. *Soit f (resp. g) une fonction continue d'intervalle $[a,b]$ de $\mathbb{R}$ à valeurs dans un espace normé $(F, \|\cdot\|_F)$ (resp. $\mathbb{R}$). Supposons que f et g soient dérivables sur $]a,b[$ et que*

$$\forall t \in]a,b[\,,\ \|f'(t)\|_F \leq g'(t).$$

Alors, on a

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

Démonstration. Soit ε un réel strictement positif. On pose

$$\varphi_\varepsilon(t) \stackrel{\text{déf}}{=} \|f(t)-f(a)\|_F - (g(t)-g(a)) - \varepsilon(t-a) - \varepsilon.$$

C'est un fonction continue de $[a,b]$ dans $\mathbb{R}$. Soit J_ε l'ensemble des t de $[a,b]$ tels que, pour tout $t' \leq t$, φ_ε soit négative. La fonction φ_ε étant continue, l'ensemble J_ε est fermé. (Exercice: démontrez-le!). Il est non vide car $\varphi_\varepsilon(a) = -\varepsilon$. La fonction φ_ε étant continue, il existe un réel strictement positif α_ε tel que, pour tout t dans $[a, a+\alpha_\varepsilon]$, on ait $\varphi_\varepsilon(t) \leq 0$. Soit t_0 un élément de $J_\varepsilon \setminus \{a\}$. Les fonctions f et g sont dérivables en t_0. Il existe donc un réel η_ε tel que

$$\forall t \in [t_0, t_0+\eta_\varepsilon[\,,\ \left\|\frac{f(t)-f(t_0)}{t-t_0}\right\|_F \leq \|f'(t_0)\|_F + \frac{\varepsilon}{2} \quad \text{et} \quad g'(t_0) \leq \frac{g(t)-g(t_0)}{t-t_0} + \frac{\varepsilon}{2}\cdot$$

Comme $t > t_0$ et que par hypothèse, $\|f'(t)\|_F \leq g'(t)$ on a

$$\forall t \in [t_0, t_0+\eta_\varepsilon[\,,\ \|f(t)-f(t_0)\|_F \leq g(t)-g(t_0)+\varepsilon(t-t_0). \tag{6.8}$$

Grâce à l'inégalité triangulaire, écrivons que, pour tout t de $[t_0, t_0+\eta_\varepsilon[$, on a

$$\begin{aligned}
\varphi_\varepsilon(t) &= \|f(t)-f(a)\|_F - (g(t)-g(a)) - \varepsilon(t-a) - \varepsilon \\
&\leq \|f(t) - \|f(t_0)\|_F + \|f(t_0)-f(a)\|_F - (g(t)-g(t_0)) \\
&\qquad\qquad - (g(t_0)-g(a)) - \varepsilon(t-t_0) - \varepsilon(t_0-a) - \varepsilon.
\end{aligned}$$

Par définition de φ_ε, on a

$$\begin{aligned}
\varphi_\varepsilon(t) &\leq \|f(t) - \|f(t_0)\|_F - (g(t)-g(t_0)) - \varepsilon(t-t_0) + \varphi_\varepsilon(t_0). \\
&\leq \|f(t) - \|f(t_0)\|_F - (g(t)-g(t_0)) - \varepsilon(t-t_0).
\end{aligned}$$

D'après (6.8), il vient $\varphi_\varepsilon(t) \leq 0$ et ce pour tout t dans $[t_0, t_0+\eta_\varepsilon[$. Comme par définition de J_ε, $]t_0-\eta_\varepsilon, t_0] \cap [a,b]$ est inclus dans J_ε, J_ε est ouvert. Comme il est non vide, $J_\varepsilon = [a,b]$ et donc, pour tout réel strictement positif ε, on a

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

En passant à la limite lorsque ε tend vers 0, on conclut la démonstration du théorème. □

Théorème 6.4.4. *Soit f une fonction différentiable sur un ouvert Ω d'un espace normé E à valeurs dans un espace normé F. Soient a et b deux points de Ω tels que le segment $[a,b]$ soit inclus dans Ω, c'est-à-dire que, pour tout t dans $[0,1]$, le point $(1-t)a+tb$ est dans Ω. On a alors*

$$\|f(b)-f(a)\|_F \leq \sup_{t\in[0,1]} \|Df((1-t)a+tb)\|_{\mathcal{L}(E,F)} \|b-a\|_E.$$

Démonstration. Elle repose sur l'inégalité des accroisements finis pour les fonctions d'une variable. Considérons en effet la fonction $F_{a,b}$ définie de l'intervalle $[0,1]$ dans $\mathbb{R}^N$ par

$$t \mapsto f(a+t(b-a)).$$

D'après la proposition 6.2.1, cette fonction est dérivable et sa dérivée au point t vaut $Df(a+t(b-a))\cdot(b-a)$. Donc l'inégalité des accroissements finis pour les fonctions d'une variable réelle implique que l'on a

$$\|F_{a,b}(1)-F_{a,b}(0)\| \leq \sup_{t\in[0,1]} \|F'_{a,b}(t)\|.$$

Mais, comme $F'_{a,b}(t) = Df(a+t(b-a))\cdot(b-a)$, le théorème est démontré. □

Cette inégalité est d'application très fréquente, et pas seulement dans ce cours. Donnons en tout de suite une application.

Théorème 6.4.5. *Soit f une fonction d'un ouvert Ω de $\mathbb{R}^N$ dans F. Supposons que la fonction f admette des dérivées partielles en tout point de Ω. Soit a un point de Ω. Si les dérivées partielles de f sont des fonctions continues de Ω, alors la fonction f est différentiable en a.*

Démonstration. Si $\overrightarrow{h} = (h_1, \cdots, h_n)$, on pose $\overrightarrow{h}_0 = 0$ et pour j dans $\{1, \cdots N\}$, $\overrightarrow{h}_j = (h_1, \cdots h_j, 0; \cdots 0)$. On a alors

$$f(a+\overrightarrow{h}) - f(a) - \sum_{j=1}^{N} \frac{\partial f}{\partial x_j}(a) h_j = \sum_{j=1}^{N} \Big(f(a+\overrightarrow{h}_j) - f(a+\overrightarrow{h}_{j-1}) - \frac{\partial f}{\partial x_j}(a) h_j \Big)$$

Pour j dans $\{1, \cdots N\}$, introduisons la fonction

$$g_j(t) = f(a + t(\overrightarrow{h}_j - \overrightarrow{h}_{j-1}) + \overrightarrow{h}_{j-1}) - t \frac{\partial f}{\partial x_j}(a) h_j.$$

La fonction g_j, pour $\|\overrightarrow{h}\|$ assez petit dérivable sur l'intervalle $]-1, 1[$ est dérivable et l'on a

$$g_j'(t) = \Big(\frac{\partial f}{\partial x_j}(a + t(\overrightarrow{h}_j - \overrightarrow{h}_{j-1})) - \frac{\partial f}{\partial x_j}(a) \Big) h_j$$

Le théorème 6.4.3 implique que

$$\|g_j(1) - g_j(0)\|_F \leq \sup_{t\in]0,1[} \Big\| \frac{\partial f}{\partial x_j}(a + t(\overrightarrow{h}_j - \overrightarrow{h}_{j-1})) - \frac{\partial f}{\partial x_j}(a) \Big\|_F \|\overrightarrow{h}\|.$$

La dérivée partielle $\dfrac{\partial f}{\partial x_j}$ étant continue en a, pour tout ε strictement positif, il existe α tel que, si $\|\overrightarrow{h}\| < \alpha$, alors

$$\forall t \in]0, 1[\ \|g_j(1) - g_j(0)\|_F < \frac{\varepsilon}{N} \|\overrightarrow{h}\|.$$

Comme

$$f(a+\overrightarrow{h}) - f(a) = \sum_{j=1}^{N} g_j(1) - g_j(0),$$

le théorème est démontré. □

On peut donner une version légèrement raffinée du théorème 6.4.4.

Proposition 6.4.6. *Sous les hypothèses du théorème 6.4.4 ci-dessus, on a*

$$\|f(b) - f(a) - Df(a)\cdot(b-a)\|_F \leq \sup_{t\in[0,1]} \|Df((1-t)a + tb) - Df(a)\|_{\mathcal{L}(E;F)} \|b-a\|_E.$$

Démonstration. Elle est très voisine de celle du théorème ci-dessus. Soit $g_{a,b}$ la fonction définie par

$$g_{a,b}(t) \stackrel{\text{déf}}{=} f(a + t(b-a)) - f(a) - tDf(a)\cdot(b-a).$$

Comme $g_{a,b}(0) = 0$ et $g_{a,b}(1) = f(b) - f(a) - Df(a)\cdot(b-a)$ et

$$g'_{a,b}(t) = Df(a + t(b-a))\cdot(b-a) - Df(a)\cdot(b-a)$$

on a le résultat en appliquant l'inégalité des accroissements finis pour les fonctions d'une variable réelle.

L'inégalité ci-dessus est évidemment plus intéressante lorsque la quantité

$$\sup_{t\in[0,1]} \|Df((1-t)a+tb) - Df(a)\|_{\mathcal{L}(E;F)}$$

est petite quand le point b est proche du point a. Ceci sera notamment le cas si l'application qui à x associe la différentielle de f au point x est continue au point a. D'où la définition suivante.

Définition 6.4.7. *Soit Ω un ouvert d'un espace normé E et f une fonction de Ω dans un espace normé F. On dit que la fonction f est de classe C^1 sur Ω si et seulement si la fonction f est différentiable en tout point de Ω et si l'application Df définie par*

$$Df\left\{\begin{array}{ccc} \Omega & \longrightarrow & \mathcal{L}(E;F) \\ x & \longmapsto & Df(x) \end{array}\right.$$

est continue en tout point de Ω.

Chapitre 7

Les différentielles d'ordre supérieur

7.1 Différentielle d'ordre deux: une première définition

On s'intéresse dans cette section aux propriétés de l'application Df apparue dans la définition 6.4.7 page 72 du chapitre précédent. On se place dans le cadre suivant. On considère une application f différentiable sur un ouvert Ω d'un espace normé E à valeurs dans un espace normé F.

Définition 7.1.1. *On dit que la fonction f est deux fois différentiable en un point a de l'ouvert Ω si elle est différentiable sur Ω et si et seulement si la fonction Df définie par*

$$Df \left\{ \begin{array}{ccc} \Omega & \longrightarrow & \mathcal{L}(E;F) \\ x & \longmapsto & Df(x) \end{array} \right.$$

est différentiable en a.

Derrière la simplicité de cette définition se cache une petite difficulté. La différentielle de Df en a est une application linéaire de E dans l'espace $\mathcal{L}(E;F)$ des applications linéaires continues de E dans F. Derrière la simplicité de cette définition se cache une petite difficulté.

Nous allons commencer par traiter le cas où l'espace de départ E est de dimension finie. Nous pourrons ainsi raisonner à l'aide du concept de dérivées partielles. Revisitons la définition ci-dessus à la lumière de cette notion. Comme l'espace vectoriel E est de dimension finie, on peut l'identifier à l'espace $\mathbb{R}^N$ muni de sa base canonique. La différentielle en un point a de Ω d'une fonction différentiable de Ω dans F s'écrit

$$Df(x) \cdot \overrightarrow{h} = \sum_{j=1}^{N} \frac{\partial f}{\partial x_j}(a)\, h^j.$$

Les vecteurs de F $\dfrac{\partial f}{\partial x_j}(a)$ représentent, pour chaque entier j, les images du j ème vecteur de la base canonique par l'application linéaire $Df(a)$. On identifie donc l'application linéaire $Df(a)$ avec le N-upplet de vecteurs de F

$$\left(\frac{\partial f}{\partial x_1}(a), \cdots, \frac{\partial f}{\partial x_N}(a)\right).$$

La définition 7.1.1 peut alors être formulée comme suit.

Définition 7.1.2. *On dit que la fonction f est deux fois différentiable en un point a de l'ouvert Ω de $\mathbb{R}^N$ si elle est différentiable sur Ω et si et seulement si la fonction définie par*

$$\begin{cases} \Omega & \longrightarrow & F^N \\ x & \longmapsto & \left(\dfrac{\partial f}{\partial x_1}(x), \cdots, \dfrac{\partial f}{\partial x_N}(x)\right) \end{cases}$$

est différentiable en a.

Nous allons maintenant donner une propriété fondamentale des applications deux fois différentiables.

La notion de différentielle seconde permet d'obtenir une meilleure approximation des fonctions.

Théorème 7.1.3. *Soit f une fonction sur un ouvert Ω de $\mathbb{R}^N$ dans F qui est différentiable sur Ω et 2 fois différentiable en un point a de Ω; on a alors*

$$f(a+\overrightarrow{h}) - f(a) - Df(a)\cdot\overrightarrow{h} - \frac{1}{2}\sum_{1\leq j,k\leq N} \frac{\partial}{\partial x_k}\left(\frac{\partial f}{\partial x_j}\right)(a)\, h^j\, h^k = o_2(\overrightarrow{h})$$

où $o_2(\overrightarrow{h})$ désigne une fonction telle que

$$\forall \varepsilon > 0\,,\ \exists \alpha \,/\ \|\overrightarrow{h}\| < \alpha \Longrightarrow \|o_2(\overrightarrow{h})\| \leq \varepsilon \|\overrightarrow{h}\|^2.$$

Démonstration. Soit F la fonction définie sur un intervalle $]\alpha_0, \alpha_0[$ par

$$\Delta(t) = f(a+t\overrightarrow{h}) - f(a) - tDf(a)\cdot\overrightarrow{h} - \frac{t^2}{2}\sum_{1\leq j,k\leq N} \frac{\partial}{\partial x_k}\left(\frac{\partial f}{\partial x_j}\right)(a)\, h^j\, h^k.$$

$$\begin{aligned} \Delta'(t) &= Df(a+t\overrightarrow{h})\cdot\overrightarrow{h} - Df(a)\cdot\overrightarrow{h} - t\sum_{1\leq j,k\leq N} \frac{\partial}{\partial x_k}\left(\frac{\partial f}{\partial x_j}\right)(a)\, h^j\, h^k \\ &= \sum_{j=1}^{N}\left(\frac{\partial f}{\partial x_j}(a+t\overrightarrow{h}) - \frac{\partial f}{\partial x_j}(a) - t\sum_{k=1}^{N}\frac{\partial}{\partial x_k}\left(\frac{\partial f}{\partial x_j}\right)(a)\, h^k\right) h^j \end{aligned}$$

D'après la définition 7.1.2, pour chaque j, la fonction G_j définie sur la boule ouverte de centre 0 et de rayon α_0 (assez petit) par

$$\overrightarrow{h} \longmapsto \frac{\partial f}{\partial x_j}(a+\overrightarrow{h}) - \frac{\partial f}{\partial x_j}(a) - \sum_{k=1}^{N}\frac{\partial}{\partial x_k}\left(\frac{\partial f}{\partial x_j}\right)(a)\, h^k$$

est différentiable en 0 et l'on a

$$DG_j'(0)\cdot\overrightarrow{h} = \sum_{k=1}^{N}\frac{\partial}{\partial x_k}\left(\frac{\partial f}{\partial x_j}\right)(a)\, h^k - \sum_{k=1}^{N}\frac{\partial}{\partial x_k}\left(\frac{\partial f}{\partial x_j}\right)(a)\, h^k = 0.$$

Ainsi donc, pour tout j dans $\{1, \cdots, N\}$ et pour tout ε strictement positif, il existe un réel strictement positif α_j tel que

$$\|\overrightarrow{h}\|_{\mathbb{R}^N} < \alpha_j \Longrightarrow \|G_j(t\overrightarrow{h})\|_F < \frac{\varepsilon}{N}\|\overrightarrow{h}\|_{\mathbb{R}^N}.$$

Il en résulte que si $\|\overrightarrow{h}\|_{\mathbb{R}^N}$ est strictement inférieur à $\alpha \stackrel{\text{déf}}{=} \min\{\alpha_1, \cdots, \alpha_N\}$, on a

$$\|\Delta'(t)\|_F < \frac{\varepsilon}{N}\|\overrightarrow{h}\|_{\mathbb{R}^N} \sum_{j=1}^N |\overrightarrow{h}^j| = \varepsilon \|\overrightarrow{h}\|^2_{\mathbb{R}^N}$$

si l'on a pris la norme $\|\cdot\|_\infty$ sur $\mathbb{R}^N$. L'inégalité des accrroissements finis assire alors que

$$\left| f(a+\overrightarrow{h}) - f(a) - Df(a)\cdot\overrightarrow{h} - \frac{1}{2}\sum_{1\leq j,k\leq N} \frac{\partial}{\partial x_k}\Big(\frac{\partial f}{\partial x_j}\Big)(a)\, h^j\, h^k \right| = |\Delta(1)-\Delta(0)| \leq \varepsilon \|\overrightarrow{h}\|^2_{\mathbb{R}^N}.$$

D'où le théorème. □

Nous allons maintenant étudier les dériv'ees partielles secondes. Une propriété importante est la symétrie au sens suivant.

Théorème 7.1.4. *Soit f une fonction d'un ouvert Ω d'un espace normé $\mathbb{R}^N$ dans $\mathbb{R}$ différentiable sur Ω. Si f est deux fois différentiable en un point a de Ω, alors on a, pour tout (ℓ, m) de $\{1, \cdots \mathrm{N}\}^2$,*

$$\frac{\partial}{\partial x_\ell}\Big(\frac{\partial f}{\partial x_m}\Big)(a) = \frac{\partial}{\partial x_m}\Big(\frac{\partial f}{\partial x_\ell}\Big)(a).$$

Démonstration. Elle utilise de manière décisive l'inégalité des accroissements finis. Posons

$$\Delta(\overrightarrow{h}, \overrightarrow{k}) \stackrel{\text{déf}}{=} f(a+\overrightarrow{h}+\overrightarrow{k}) + f(a) - f(a+\overrightarrow{h}) - f(a+\overrightarrow{k}) - B(\overrightarrow{k}, \overrightarrow{h}) \quad \text{avec}$$
$$B(\overrightarrow{k}, \overrightarrow{h}) \stackrel{\text{def}}{=} \sum_{1\leq \ell,m\leq N} \frac{\partial}{\partial x_m}\Big(\frac{\partial f}{\partial x_\ell}\Big)(a) k_m h_\ell.$$

Introduisons la fonction

$$g(t) \stackrel{\text{déf}}{=} f(a+t\overrightarrow{h}+\overrightarrow{k}) + f(a) - f(a+t\overrightarrow{h}) - tB(\overrightarrow{k}, \overrightarrow{h}).$$

On a

$$g'(t) = \big(Df(a+t\overrightarrow{h}+\overrightarrow{k}) - Df(a+t\overrightarrow{h})\big)(\overrightarrow{h}) - B(\overrightarrow{k}, \overrightarrow{h}).$$

Par définition de B, et d'après le lien entre différentielle et dérivée partielle, on en déduit que

$$g'(t) = \sum_{\ell=1}^N \bigg(\frac{\partial f}{\partial x_\ell} f(a+t\overrightarrow{h}+\overrightarrow{k}) - \frac{\partial f}{\partial x_\ell} f(a+t\overrightarrow{h}) - \sum_{m=1}^N \frac{\partial}{\partial x_m}\Big(\frac{\partial f}{\partial x_\ell}\Big)(a) k_m\bigg) h_\ell.$$

Les fonctions $\dfrac{\partial f}{\partial x_\ell}$ étant différentiable en a, pour tout ε strictement positif, il existe un réel strictement positif α tel que, si $\|\overrightarrow{h}\|_{\mathbb{R}^N} + \|\overrightarrow{k}\|_{\mathbb{R}^N} < \alpha$, alors

$$\left|\frac{\partial f}{\partial x_\ell}(a+t\overrightarrow{h}+\overrightarrow{k}) - \frac{\partial f}{\partial x_\ell}(a) - D\frac{\partial f}{\partial x_\ell}(a)(t\overrightarrow{h}+\overrightarrow{k})\right| < \varepsilon(\|\overrightarrow{h}\|_{\mathbb{R}^N} + \|\overrightarrow{k}\|_{\mathbb{R}^N}) \quad \text{et}$$
$$\left|\frac{\partial f}{\partial x_\ell}(a+t\overrightarrow{h}) - \frac{\partial f}{\partial x_\ell}(a) - D\frac{\partial f}{\partial x_\ell}(a)(t\overrightarrow{h})\right| < \varepsilon(\|\overrightarrow{h}\|_{\mathbb{R}^N} + \|\overrightarrow{k}\|_{\mathbb{R}^N}).$$

Par différence, on obtient que

$$\left|\Big(\frac{\partial f}{\partial x_\ell} f(a+t\overrightarrow{h}+\overrightarrow{k}) - \frac{\partial f}{\partial x_\ell} f(a+t\overrightarrow{h}) - D\frac{\partial f}{\partial x_\ell}(a)\overrightarrow{k}\right| < \varepsilon(\|\overrightarrow{h}\|_{\mathbb{R}^N} + \|\overrightarrow{k}\|_{\mathbb{R}^N}).$$

D'après le lien entre différentielle et dérivée partielle, on a

$$D\frac{\partial f}{\partial x_\ell}(a)\overrightarrow{k} = \sum_{m=1}^{N} \frac{\partial}{\partial x_m}\left(\frac{\partial f}{\partial x_\ell}\right)(a)k_m$$

Ainsi donc on trouve que, si $\|\overrightarrow{h}\|_{\mathbb{R}^N} + \|\overrightarrow{k}\|_{\mathbb{R}^N} < \alpha$, alors

$$|g'(t)| \leq \varepsilon\big(\|\overrightarrow{h}\|_{\mathbb{R}^N} + \|\overrightarrow{k}\|_{\mathbb{R}^N}\big)\sum_{\ell=1}^{N}|h^\ell|.$$

En pren ant sur $\mathbb{R}^N$ la norme "somme des valeurs absolues des coordonées, on trouve que si $\|\overrightarrow{h}\|_{\mathbb{R}^N} + \|\overrightarrow{k}\|_{\mathbb{R}^N} < \alpha$, alors

$$|g'(t)| \leq \varepsilon\big(\|\overrightarrow{h}\|_{\mathbb{R}^N} + \|\overrightarrow{k}\|_{\mathbb{R}^N}\big)\|\overrightarrow{h}\|_{\mathbb{R}^N}.$$

L'inégalité des accroissements finis implique que

$$|\Delta(\overrightarrow{h}, \overrightarrow{k})| \leq \varepsilon\big(\|\overrightarrow{h}\|_{\mathbb{R}^N} + \|\overrightarrow{k}\|_{\mathbb{R}^N}\big)^2 \tag{7.1}$$

Comme on a

$$B(\overrightarrow{k}, \overrightarrow{h}) - B(\overrightarrow{h}, \overrightarrow{k}) = \Delta(\overrightarrow{h}, \overrightarrow{k}) - \Delta(\overrightarrow{k}, \overrightarrow{h}),$$

on déduit de l'inégalité (7.1) que

$$\forall\varepsilon > 0\,,\ \exists\alpha > 0\,,\ \|\overrightarrow{h}\|_{\mathbb{R}^N} + \|\overrightarrow{k}\|_{\mathbb{R}^N} < \alpha \Longrightarrow \big|B(\overrightarrow{k}, \overrightarrow{h}) - B(\overrightarrow{h}, \overrightarrow{k})\big| < \varepsilon(\|\overrightarrow{h}\|_E + \|\overrightarrow{k}\|_E)^2.$$

Appliquons maintenant un argument d'homogénéïté. Pour tout ε strictement positif, il existe un réel α tel que, pour tout couple $(\overrightarrow{h}, \overrightarrow{k})$ de vecteurs de $\mathbb{R}^N$ non identiquement nul, on ait

$$\big|B(\overrightarrow{k}, \overrightarrow{h}) - B(\overrightarrow{h}, \overrightarrow{k})\big| < \varepsilon(\|\overrightarrow{h}_\alpha\|_{\mathbb{R}^N} + \|\overrightarrow{k}_\alpha\|_{\mathbb{R}^N})^2.$$

avec

$$\overrightarrow{h}_\alpha \stackrel{\text{déf}}{=} \frac{\alpha}{2}\frac{\overrightarrow{h}}{(\|\overrightarrow{h}\|_{\mathbb{R}^N} + \|\overrightarrow{k}\|_{\mathbb{R}^N})} \quad \text{et} \quad \overrightarrow{k}_\alpha \stackrel{\text{déf}}{=} \frac{\alpha}{2}\frac{\overrightarrow{k}}{(\|\overrightarrow{h}\|_{\mathbb{R}^N} + \|\overrightarrow{k}\|_{\mathbb{R}^N})}\cdot$$

Ceci implique que, pour tout ε, pour tout couple $(\overrightarrow{h}, \overrightarrow{k})$ de vecteurs de $\mathbb{R}^N$ non identiquement nul, on a

$$\big|B(\overrightarrow{k}, \overrightarrow{h}) - B(\overrightarrow{h}, \overrightarrow{k})\big| < 2\varepsilon(\|\overrightarrow{h}\|_{\mathbb{R}^N} + \|\overrightarrow{k}\|_{\mathbb{R}^N})^2.$$

En passant à la limite lorsque ε tend vers 0, on trouve que

$$B(\overrightarrow{k}, \overrightarrow{h}) = B(\overrightarrow{h}, \overrightarrow{k})$$

En choisissant pour $\overrightarrow{k}$ le m ème vecteur de la base canonique de $\mathbb{R}^N$ et pour $\overrightarrow{h}$ le ℓ ème, on trouve le résultat. □

Définition 7.1.5. *Soit f une fonction de Ω ouvert de $\mathbb{R}^N$ à valeurs dans $\mathbb{R}$ deux fois différentiable en un point a de Ω. On appelle différentielle seconde de f en a (elle est notée $D^2f(a)$) la forme quadratique définie sur $\mathbb{R}^N$ par*

$$D^2f(a)(\overrightarrow{h}, \overrightarrow{k}) \stackrel{\text{def}}{=} \sum_{1\leq\ell,m\leq N} \frac{\partial^2 f}{\partial x_\ell \partial x_m}(a)\, h_\ell k_m \quad \text{avec} \quad \frac{\partial^2 f}{\partial x_\ell \partial x_m}(a) \stackrel{\text{déf}}{=} \frac{\partial}{\partial x_\ell}\left(\frac{\partial f}{\partial x_m}\right)(a)$$

7.2 Une petite digression sur les formes quadratiques sur $\mathbb{R}^N$

Comme cette notion joue un rôle dans la section suivante, démontrons quelques propriétés des formes quadratiques sur $\mathbb{R}^N$. Une forme quadratique B est une application bilinéaire symétrique de $\mathbb{R}^N \times \mathbb{R}^N$ dans $\mathbb{R}$ c'est-à-dire que tout tout $\overrightarrow{k}$, les applications

$$\overrightarrow{h} \longmapsto B(\overrightarrow{h}, \overrightarrow{k}) \quad \text{et} \quad \overrightarrow{h} \longmapsto B(\overrightarrow{k}, \overrightarrow{h})$$

sont des applications linéaires de $\mathbb{R}^N$ dans $\mathbb{R}$ (c'est-à-dire des formes linéaires). C'est l'aspect bilinéaire. De plus, symétrique signifie que

$$\forall (\overrightarrow{h}, \overrightarrow{k}) \in \mathbb{R}^N \times \mathbb{R}^N\,,\ B(\overrightarrow{h}, \overrightarrow{k}) = B(\overrightarrow{k}, \overrightarrow{h}).$$

Si l'on cherche à écrire l'expression de B en fonction des coordonnés des vecteurs $\overrightarrow{h}$ et $\overrightarrow{k}$, on trouve que

$$B(\overrightarrow{h}, \overrightarrow{k}) = \sum_{1 \leq \ell, m \leq N} B_{\ell,m} h_\ell k_m \quad \text{avec} \quad B_{\ell,m} = B_{m,\ell}, \tag{7.2}$$

c'est-à-dire les formes quadratiques sur $\mathbb{R}^N$ s'identifie avec les matrices $N \times N$ symétriques. Remarquons qu'en tant que somme et produite de fonctions continues très élémentaires, les formes quadratiques sur $\mathbb{R}^N$ sont des applications continues de $s\,\mathbb{R}^N \times \mathbb{R}^N$ dans $\mathbb{R}$.

Remarquons aussi que si l'on considère la forme quadratique particulière

$$(\overrightarrow{h} \,|\, \overrightarrow{k})_e \stackrel{\text{déf}}{=} \sum_{\ell=1}^{N} h_\ell k_\ell$$

qui correspond au cas où la matrice B est l'identité, on a

$$(\overrightarrow{h} \,|\, \overrightarrow{h})_e = \|h\|_e^2 = \sum_{\ell=1}^{N} (h_\ell)^2.$$

De plus, on peut écrire que

$$B(\overrightarrow{h}, \overrightarrow{k}) = (B\overrightarrow{h} \,|\, \overrightarrow{k})_e \quad \text{avec} \quad (Bh)_\ell = \sum_{m=1}^{N} B_{\ell,m} h^m.$$

Une notion importante est la notion de forme quadratique non dégénérée.

Définition 7.2.1. *Soit B une forme quadratique sur $\mathbb{R}^N$. On dit que B est non dégénérée si et seulement si l'application*

$$\delta_B \left\{ \begin{array}{ccc} \mathbb{R}^N & \longrightarrow & (\mathbb{R}^N)' \\ \overrightarrow{k} & \longmapsto & B(\cdot, \overrightarrow{k}) \end{array} \right.$$

est une application linéaire bijective. On dit qu'une forme quadratique B est définie positive (resp. négative) si et seulement si

$$\forall \overrightarrow{h} \in \mathbb{R}^N \setminus \{0\}\,,\ B(\overrightarrow{h}, \overrightarrow{h}) > 0 \ (\textit{resp.} \ < 0).$$

Ceci est équivalent au fait que la matrice B soit inversible.

Théorème 7.2.2. *On considère $\mathbb{R}^N$ muni de la forme quadratique euclidienne, c'est-à-dire la forme quadratique définie par*

$$B_e(\overrightarrow{h}, \overrightarrow{k}) = \sum_{\ell=1}^{N} h_\ell k_\ell \quad si \quad \overrightarrow{h} = (h_1, \cdots h_N) \quad et \quad \overrightarrow{k} = (k_1, \cdots k_N).$$

Soit B une forme quadratique quelconque. Il existe alors une base orthonormée pour B_e, c'est-à-dire une base $(\overrightarrow{e}_\ell)_{1\leq \ell \leq N}$ telle que

$$\forall (\ell, m) \in \{1, \cdots N\}^2\,,\ B_e(\overrightarrow{e}_\ell, \overrightarrow{e}_m) = \delta_{\ell,m}$$

et telle que

$$B(\overrightarrow{e}_\ell, \overrightarrow{e}_m) = \lambda_\ell \delta_{\ell,m}.$$

Ce théorème dit que toute forme quadratique B est diagonalisable en base orthonormée.

Démonstration. Soit $\mathbb{S}^{N-1}$ la sphère unité $\mathbb{R}^N$, c'est-à-dire l'ensemble des points de $\mathbb{R}^N$ dont telle que $B_e(x,x) = 1$. C'est un compact de $\mathbb{R}^N$ (car un fermé comme image réciproque du fermé $\{1\}$ par l'application continue $x \mapsto \|x\|_e^2 = B_e(x,x)$). On considère

$$M_0 \stackrel{\text{déf}}{=} \max_{x\in\mathbb{S}^{N-1}} \big|B(x,x)\big|.$$

Si M_0 est nulle, alors, pour tout x de $\mathbb{R}^N \setminus \{0\}$,

$$B\Big(\frac{x}{\|x\|_e}, \frac{x}{\|x\|_e}\Big) = 0 \quad \text{et donc} \quad B(x,x) = 0$$

Comme on a

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

on a alors $B \equiv 0$. et le résultat est alors évident. Supposons que $B \not\equiv 0$. Quitte à changer B en $-B$ on peut supposer que M_0 est strictement positif. On a alors

$$\forall x \in \mathbb{S}^{N-1}\,,\ M_0 - B(x,x) \leq 0 \quad \text{et donc} \quad \forall x \in \mathbb{R}^N\,,\ M_0 - B\Big(\frac{x}{\|x\|_e}, \frac{x}{\|x\|_e}\Big) \leq 0.$$

On en déduit que

$$\forall x \in \mathbb{R}^N\,,\ M_0\|x\|_e^2 - B(x,x) \leq 0$$

Comme $\mathbb{S}^{N-1}$ est compact, il existe un point x_0 de $\mathbb{S}^{N-1}$ tel que $M_0 = B(x_0, x_0)$. Ainsi donc x_0 est un minimum

$$B_0 \left\{ \begin{array}{ccc} \mathbb{R}^N & \longrightarrow & \mathbb{R} \\ x & \longmapsto & M_0\|x\|_e^2 - B(x,x). \end{array} \right.$$

La différentielle de cette application B_0 s'annulle en x_0. On a donc

$$\forall \overrightarrow{h} \in \mathbb{R}^N\,,\ DB_0(x_0)\overrightarrow{h} = 2M_0(x_0|\overrightarrow{h}) - B(x_0, \overrightarrow{h}).$$

Ceci signifie que l'orthogonal au sens de la forme quadratique euclidienne et au sens de la forme quadratique B coï ncide. On conclue la démonstration par récurrence sur la dimension N de l'espace. □

7.3 Extrema et allure locale des fonctions

Démontrons tout d'abord une condition nécessaire fondamentale pour qu'une fonction diff 'erentiable admet ue extremum localen un point.

Théorème 7.3.1. *Soient f une fonction deux fois différentiable sur un ouvert Ω d'un espace vectoriel normé E à valeurs dans $\mathbb{R}$ et a un point de Ω. Supposons que le point a soit un minimum (resp. un maximum) local de la fonction f, c'est-à-dire qu'il existe un ouvert ω contenant a tel que, pour tout point x de ω, on ait $f(x) \geq f(a)$ (resp. $f(x) \leq f(a)$). Alors $Df(a) = 0$.*

Démonstration. Soit $\overrightarrow{h}$ un vecteur de E. Par définition d'un ouvert, il existe un réel strictement positif α tel que, si $|t| \leq \alpha$, alors $a + t\overrightarrow{h}$ appartient à ω. Bien sûr, le point 0 est un minimum local de la fonction F_a de $[-\alpha, \alpha]$ dans $\mathbb{R}$ définie par

$$F_a(t) \stackrel{\text{déf}}{=} f(a + t\overrightarrow{h}). \tag{7.3}$$

Donc $F_a'(0) = 0$. Or, on sait bien que $F_a'(t) = Df(a+t\overrightarrow{h}) \cdot \overrightarrow{h}$. Ainsi donc, on a $Df(a) \cdot \overrightarrow{h} = 0$; et ce pour tout vecteur $\overrightarrow{h}$ de E. D'où le résultat. □

La différentielle seconde permet de préciser cette condition nécessaire comme suit. La notion de différentielle seconde permet d'obtenir une meilleure approximation des fonctions.

Théorème 7.3.2. *Soient f une fonction d'un ouvert Ω de $\mathbb{R}^N$ à valeurs dans $\mathbb{R}$ et a un point de Ω. Supposons que le point a soit un minimum (resp. un maximum) local de la fonction f et que f soit deux fois différentiable en a. Alors*

$$Df(a) = 0 \quad \textit{et} \quad \forall \overrightarrow{h} \in E\,,\ D^2f(a)(\overrightarrow{h}, \overrightarrow{h}) \geq 0 \ \textit{(resp.} \leq 0\textit{)}.$$

Démonstration. Soit $\overrightarrow{h}$ un vecteur de E. On utilise la fonction F_a définie par (7.3). Pour tout t de l'intervalle $]-\alpha, \alpha[$, on a $F_a'(t) = Df(a + t\overrightarrow{h}) \cdot \overrightarrow{h}$. Comme la fonction f est deux fois différentiable, la fonction F_a' est différentiable sur l'intervalle $]-\alpha, \alpha[$ et l'on a

$$F_a''(t) = D^2f(a + t\overrightarrow{h})(\overrightarrow{h}, \overrightarrow{h}).$$

Si a est un minimum local de la fonction f, 0 est un minimum local de la fonction F_a. On sait que pour une fonction réelle F_a d'une variable réelle deux fois dérivable qui a un minimum local en 0 vérifie $F_a'(0) = 0$ et $F_a''(0) \geq 0$, ce qui assure le résultat. □

Nous allons démontrer maintenant une condition suffisante pour qu'un point soit un minimum (ou un maximum) local d'une fonction. À nouveau, dans le cas d'une fonction d'une variable réelle, on sait bien qu'il suffit que la dérivée seconde soit strictement positive pour un minimum (strictement négative pour un maximum).

Théorème 7.3.3. *Soient f une fonction d'un ouvert Ω de $\mathbb{R}^N$ à valeurs dans $\mathbb{R}$ et a un point de Ω. Supposons que f soit deux fois différentiable en a, que la différentielle de f au point a soit nulle et que la différentielle seconde au point a soit définie positive (resp. définie négative). Alors le point a est un minimum (resp. un maximum) local de la fonction f.*

Démonstration. Elle utilise une formule de Taylor-Young à l'ordre deux, c'est-à-dire le théorème 7.1.3. En effet, comme f est deux fois différentiable en point a et que $Df(a) = 0$, on a

$$f(a + \overrightarrow{h}) - f(a) - D^2f(a) \cdot (\overrightarrow{h}, \overrightarrow{h}) = o_2(\overrightarrow{h}). \tag{7.4}$$

La forme quadratique $D^2 f(a)$ est supposée définie positive. Donc, pour tout vecteur $\overrightarrow{h}$ de la sphère unité de $\mathbb{R}^N$, il existe $D^2 f(a)\cdot(\overrightarrow{h},\overrightarrow{h})$ est non nul. Nous sommes en dimension finie. Donc la sphère unité est compacte. La forme quadratique $D^2 f(a)$ est une fonction continue strictement positive sur ce compact. Donc elle est atteint son minimum; ce minimum est donc strictement positif. Il existe donc un réel strictement positif m tel que

$$\forall \overrightarrow{h} \in \mathbb{R}^N \ / \ \|\overrightarrow{h}\| = 1\,,\ D^2 f(a)\cdot(\overrightarrow{h},\overrightarrow{h}) \geq m.$$

Nous allons maintenant utiliser un argument d'homogénéîté. Soit $\overrightarrow{h}$ un vecteur non nul de $\mathbb{R}^N$. Le vecteur $\|\overrightarrow{h}\|^{-1}\overrightarrow{h}$ est de norme 1. Donc

$$D^2 f(a)\cdot\left(\frac{\overrightarrow{h}}{\|\overrightarrow{h}\|},\frac{\overrightarrow{h}}{\|\overrightarrow{h}\|}\right) \geq m\,;$$

d'où il résulte que $D^2 f(a)\cdot(\overrightarrow{h},\overrightarrow{h}) \geq m\|\overrightarrow{h}\|^2$. L'inégalité triangulaire appliquée à (7.4) implique que

$$f(a+\overrightarrow{h}) - f(a) \geq m\|\overrightarrow{h}\|^2 - \|o_2(\overrightarrow{h})\|.$$

Par définition de la fonction o_2, il existe un réel strictement positif α tel que

$$\forall \overrightarrow{h} \in B(0,\alpha)\,, \|o_2(\overrightarrow{h})\| \leq \frac{m}{2}\|\overrightarrow{h}\|^2.$$

Donc, pour tout vecteur $\overrightarrow{h}$ de la boule de centre 0 et de rayon α, on a

$$f(a+\overrightarrow{h}) - f(a) \geq \frac{m}{2}\|\overrightarrow{h}\|^2.$$

D'où le théorème. □

7.4 La notion de fonctions C^2 et la formule de Taylor avec reste intégral

Définition 7.4.1. *Soit f une application d'un ouvert Ω de $\mathbb{R}^N$ dans $\mathbb{R}$. On dit que f est de classe C^2 sur Ω si et seulement si la fonction f est deux fois différentiables en tout point de x de Ω et si les applications*

$$\left\{\begin{array}{ccc} \Omega & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \dfrac{\partial^2 f}{\partial x_j \partial x_k}(x) \end{array}\right.$$

sont continues.

Pour de telles fonctions, la formule de Taylor prend prendre une forme plus précise et plus globale que dans le théorème 7.1.3.

Théorème 7.4.2. *Soit f une application de classe C^2 d'un ouvert Ω de $\mathbb{R}^N$ dans $\mathbb{R}$. On considère deux points a et b de Ω tel que le segment $[a,b]$ soit inclus dans Ω. On a alors*

$$f(b) - f(a) = Df(a)(b-a) + \int_0^1 (1-t)D^2 f(a+t(b-a))(b-a,b-a)dt.$$

Démonstration. Observons que si g est une fonction deux fois continûment dérivable de l'intervalle $[0,1]$ à valeurs dans $\mathbb{R}$, on a, grâce à une intégration par parties,

$$\int_0^1 (1-t)g''(t)dt = -g'(0) + \int_0^1 g'(t)dt = -g'(0) + g(1) - g(0). \tag{7.5}$$

On applique ceci la fonction

$$g(t) \stackrel{\text{déf}}{=} f(a + t(b-a)).$$

Comme le segment $[a,b]$ est inclus dans Ω, la fonction g est différentiable et l'on a

$$g'(t) = Df(a + t(b-a)(b-a) = \sum_{j=1}^{N} \frac{\partial}{\partial x_j} f(a + t(b-a))(b-a)_j.$$

Le fait que l'on ait supposé que la fonction f soit de classe C^2 implique que la fonctiong' est dérivable et que

$$g''(t) = \sum_{j=1}^{N}\sum_{k=1}^{N} \frac{\partial^2 f}{\partial x_j \partial x_k}\big(a + t(b-a)\big)(b-a)_j(b-a)_k = D^2 f(a + t(b-a))(b-a, b-a).$$

La fonction g'' étant continue, il suffit d'appliquer la relation (7.5). □

7.5 La notion de fonction C^k

Le but de cette section est de donner une définition de la notion de fonction C^k et donner quelques propriétés de fonctions de classe C^k.

Définition 7.5.1. *Soit Ω un ouvert de $\mathbb{R}^N$ et k un entier. On dit que la fonction f est de classe C^k si elle admet des dérivées partielles d'ordre k continues sur Ω.*

On adettre que la composée de deux fonctions C^k est une fonction C^k. Ceci résutle de la formule de différentiation composé

$$\frac{\partial}{\partial x_j}(g \circ f)(x) = \sum_{k=1}^{N} \frac{\partial}{\partial y_k} g(f(x)) \frac{\partial f^k}{\partial x_j}(x)$$

et d'une récurrence omise.

Chapitre 8

L'inversion locale et les fonctions implicites

8.1 Le théorème d'inversion locale

La différentielle d'une fonction représente une approximation locale de cette fonction. C'est une question naturelle que de se demander si des propriétés faciles à lire sur l'approximation se transmettent à la fonction elle même. Par exemple, supposons que la différentielle de f en un point a soit inversible en tant qu'application linéaire. Est-ce que la fonction f sera inversible? Remarquons que l'inversibilté de la differéntielle en un point est une hypothèse assez forte: si les deux espaces sont de dimension finie, cela implique qu'elle est la même. Lorsque la dimension est infinie, c'est une hypothèse forte. Dans la pratique, on a très souvent $E = F$.

Lorsque $n = N = 1$, il est assez facile de répondre à la question. En effet, si f est une fonction C^1 de $]a, b[$ dans $\mathbb{R}$, l'inversibilité de la différentielle signifie que l'application dérivée ne s'annulle pas. Un résultat classique de la théorie des fonctions d'une variable réelle dit que cette application est une bijection de $]a, b[$ sur son image qui est alors un intervalle ouvert. De plus, l'application réciproque est aussi de classe C^1.

En dimension supérieure ou égale à deux, la question est plus délicate et n'a de sens que localement. Plus précisément, elle doit être posée de la manière suivante:

Existe-t-il deux ouverts Ω_1 et Ω_2 de $\mathbb{R}^N$ contenant respectivement a et $f(a)$ tels que f soit une bijection de Ω_1 dans Ω_2?

La réponse à cette question est positive et est contenue dans le théorème suivant. Ici, une hypothèse de dimension finie n'apporte aucune simplification dans la démonstration.

Théorème 8.1.1. *Soient Ω un ouvert d'un espace de Banach E et $f$une application de classe C^1 sur Ω à valeurs dans un espace de Banach F. Considérons un point a de Ω tel que $Df(a)$ soit un isomorphisme (c'est-à-dire une application linéaire continue inversible d'inverse continue) entre les deux espaces de Banach E et F. Il existe alors deux ouverts Ω_1 et Ω_2 de E et F respectivement contenant respectivement a et $f(a)$ tels que f soit un C^1-difféomorphisme Ω_1 sur Ω_2, c'est-à-dire une bijection C^1 d'inverse C^1)*

Démonstration. Comme dans la démonstration du théorème 8.1.5, on se ramenè au cas où $Df(a)$ est l'identité, $E = F$ et $f(0) = 0$. En effet, posons

$$g(\overrightarrow{h}) \stackrel{\text{déf}}{=} Df(a)^{-1}\big(f(a + \overrightarrow{h}) - f(a)\big).$$

La fonction g est différentiable sur l'ouvert Ω à valeurs dans E en tant que composée de fonctions différentiables et est de classe C^1 au point a en tant que composée d'une fonctions de classe C^1 au point a et d'une fonction de classe C^1 partout.

Calculons la différentielle de la fonction g au point 0. D'après la théorème de composition, on a

$$Dg(0) = Df(a)^{-1} \cdot Df(a) = \mathrm{Id}\,.$$

Démontrons maintenant le lemme suivant:

Lemme 8.1.2. *Soient Ω un ouvert de $\mathbb{R}^N$ et g une application de classe C^1 d'un ouvert Ω d'un espace de Banach E dans lui-même. Si de plus*

$$g(0) = 0 \quad et \quad Dg(0) = \mathrm{Id},$$

on a alors, en posant $\widetilde{g} \stackrel{\text{déf}}{=} g - \mathrm{Id}$

$$\forall \varepsilon > 0\,,\ \exists \alpha \,/\, \forall (x,y) \in B(0,\alpha) \times B(0,\alpha)\,,\ \|\widetilde{g}(x) - \widetilde{g}(y)\| \leq \varepsilon \|x - y\|.$$

Démonstration. Pour démontrer ceci, appliquons l'inégalité des accroissements finis sous la forme de la proposition 6.4.6. D'où

$$\forall (x,y) \in B(0,\alpha) \times B(0,\alpha)\,,\ \|\widetilde{g}(x) - \widetilde{g}(y)\| \leq \sup_{z \in B(0,\alpha)} \|Dg(z) - \mathrm{Id}\,\|_{\mathcal{L}(E)} \|x - y\|.$$

Comme la fonction Dg est supposée continue en 0, pour tout réel strictement positif ε, il existe un réel strictement positif α tel que

$$\forall (x,y) \in B(0,\alpha) \times B(0,\alpha)\,,\ \|\widetilde{g}(x) - \widetilde{g}(y)\| \leq \sup_{z \in B(0,\alpha)} \varepsilon \|x - y\|.$$

D'où le lemme. □

Remarque Lorsque ε est strictement plus petit que 1, pour tout $\overrightarrow{h}$ dans $B(0,\alpha)$, l'application $Dg(\overrightarrow{h})$ appartient à $\mathring{B}(\mathrm{Id}, 1)$ et est donc inversible.

Poursuite de la démonstration du théorème 8.1.1 Le théorème de point fixe de Picard (voir Théorème 2.2.1) assure l'existence d'un réel strictement positif α_0 tel que, pour tout $\overrightarrow{k}$ de $B_f(0,\alpha_0)$, (la boule fermée), il existe un unique $\overrightarrow{h}$ de $B_f(0,\alpha_0)$ tel que

$$g(\overrightarrow{h}) = \overrightarrow{h} + \widetilde{g}(\overrightarrow{h}) = \overrightarrow{k}\,.$$

L'image de g continue donc $B(0,\alpha_0)$. De plus, pour tout

$$\|g(\overrightarrow{h}_1) - g(\overrightarrow{h}_2)\|_E \leq \frac{3}{2} \|\overrightarrow{h}_1 - \overrightarrow{h}_2\|_E$$

De plus, on a

$$\|g(\overrightarrow{h}_1) - g(\overrightarrow{h}_2)\|_E \geq \|\overrightarrow{h}_1 - \overrightarrow{h}_2\|_E - \|\widetilde{g}(\overrightarrow{h}_1) - \widetilde{g}'(\overrightarrow{h}_2)\|_E \geq \frac{1}{2} \|\overrightarrow{h}_1 - \overrightarrow{h}_2\|_E.$$

Ainsi donc g est un homéomorphisme de $\mathring{B}(0,\alpha_0)$ sur $\mathring{B}(0,\alpha_0) \cap g^{-1}(\mathring{B}(0,\alpha_0))$. Vu que l'on a

$$f(a + \overrightarrow{h}) = f(a) + Df(a) \cdot g(\overrightarrow{h})$$

on en déduit que f est un homéomorphisme de Ω_1 sur Ω_2 avec

$$\Omega_1 \stackrel{\text{déf}}{=} a + \mathring{B}(0,\alpha_0) \cap g^{-1}(\mathring{B}(0,\alpha_0)) \quad \text{et} \quad f(a) + Df(a)(\mathring{B}(0,\alpha_0)).$$

Grâce à la remarque ci-dessus, on sait que, pour tout x dans Ω_1, l'application $Df(x)$ est inversible. L'application du théorème 8.1.5 conclue la démonstration. □

Donnons maintenant un corollaire très important de ce théorème.

Corollaire 8.1.3. *Soit f une application de classe C^1 d'un ouvert Ω de $\mathbb{R}^N$ dans $\mathbb{R}^N$ telle que, pour tout point x de Ω, la différentielle de f au point x soit une application linéaire inversible. L'ensemble $f(\Omega)$ est alors un ouvert.*

Démonstration. Considérons un point y de $f(\Omega)$ et démontrer qu'il existe une boule de $\mathbb{R}^N$ centrée en y et contenue dans $f(\Omega)$. Le théorème d'inversion locale nous dit qu'il existe deux ouverts Ω_1 et Ω_2 de $\mathbb{R}^N$ tels que Ω_1 contienne $f^{-1}(y)$ et Ω_2 contienne y et tel que f soit une bijection de Ω_1 sur Ω_2. Ceci signifie qu'il existe une boule ouverte de $\mathbb{R}^N$ centrée en y et incluse dans Ω_2; cette boule est donc telle que, pour tout y' dans cette boule, il existe un élément x' de Ω_1 tel que $f(x') = y'$; donc l'ensemble $f(\Omega)$ est ouvert. □

Nous allons maintenant introduire une des notions cruciales de ce cours: celle de difféomorphisme.

Définition 8.1.4. *Soient k un entier supérieur ou égal à 1 et f une application C^kd'un ouvert Ω de $\mathbb{R}^N$ dans $\mathbb{R}^N$. On dit que f est un difféomorphisme de classe C^k si et seulement si*

- *l'application f est injective;*
- *pour tout point x de l'ouvert Ω, la différentielle de f au point x est une application linéaire inversible;*
- *l'application réciproque f^{-1} définie par*
$$\left\{\begin{array}{ccl} f(\Omega) & \longrightarrow & \Omega \\ y & \longmapsto & f^{-1}(y) \quad \text{définie par} \quad f(f^{-1}(y)) = y. \end{array}\right.$$
est une fonction de classe C^k.

Cette définition appelle quelques commentaires. Tout d'abord, la troisième assertion de la définition ci-dessus n'a de sens que si l'ensemble $f(\Omega)$ est un ouvert. Comme la différentielle est inversible en tout point, c'est bien le cas d'après le corollaire 8.1.3 ci-dessus.

Ensuite, cette troisième assertion est en faite une conséquence du théorème de composition. En effet, nous avons le théorème suivant.

Théorème 8.1.5. *Soit f une fonction de classe C^k d'un ouvert Ω de $\mathbb{R}^N$ dans $\mathbb{R}^N$ qui est une bijection sur son image et qui est telle qu'en tout point x de Ω, la différentielle de f en x soit une application linéaire inversible; alors f^{-1} l'application réciproque de f est de classe C^k et l'on a*
$$Df^{-1}(y) = (Df)^{-1}(f^{-1}(y)).$$

Démonstration. Ce théorème est très lié au théorème de composition et la démonstration reprend les mêmes idées. Il suffit en fait de démontrer le lemme suivant

Pour démontrer l'intégralité du théorème 8.1.5, il faut démontrer que si la fonction f est de classe C^k, alors la fonction $D(f^{-1})$ l'est aussi. D'après le lemme **??** ci-dessus,
$$D(f^{-1}) = \mathrm{Inv} \circ Df \circ f^{-1},$$
où Inv est l'application définie par
$$\mathrm{Inv} \left\{\begin{array}{ccl} GL(\mathbb{R}^N) & \longrightarrow & GL(\mathbb{R}^N) \\ u & \longmapsto & u^{-1}, \end{array}\right.$$
où $GL(\mathbb{R}^N)$ désigne l'ensemble des applications linéaires inversibles de $\mathbb{R}^N$ dans $\mathbb{R}^N$.

Lemme 8.1.6. *L'application* Inv *est de classe* C^k *pour tout* k.

Démonstration. Pour démontrer ce lemme, il suffit de se souvenir que l'application Inv est différentiable et que

$$D(\mathrm{Inv})(u) \cdot \overrightarrow{h} = -(\mathrm{Inv} u) \circ \overrightarrow{h} \circ (\mathrm{Inv} u).$$

Une récurrence assure alors que si Inv est de classe C^k, alors DInv aussi en tant que composée de fonctions C^k.

Concluons maintenant la démonstration du théorème 8.1.5. Si f est de classe C^k, supposons démontré que l'application f^{-1} est de classe C^ℓ avec $\ell \leq k-1$. L'application $D(f^{-1})$ est alors une fonction de classe C^ℓ en tant que composée de fonctions de classe C^ℓ, donc la fonction f^{-1} est de classe $C^{\ell+1}$; d'où le théorème. □

8.2 Le théorème des fonctions implicites

Théorème 8.2.1. *Soient* Ω *un ouvert de* $\mathbb{R}^N$, p *un entier compris entre* 1 *et* $N-1$ *et* f *une fonction de classe* C^k *de* Ω *dans* $\mathbb{R}^{N-p}$. *On se donne des espaces vectoriels supplémentaires* E *et* F *dans* $\mathbb{R}^N$ *tels que la dimension de* E *soit égale à* p. *On considère un point* (a,b) *de* $E \times F$ *tels que* $X_0 = (a,b) \in \Omega$ *et* $f(X_0) = 0$. *et On suppose que* $Df(X_0)_{|F}$ *est une application linéaire bijective de* F *sur* $\mathbb{R}^{N-p}$.

Il existe alors un ouvert U *de* E *contenant* a *et un ouvert* ω *de* Ω *contenant* X_0 *et une fonction* g *de classe* C^k *sur* U *tels que*

$$\big\{(x,y) \in \omega \,/\, f(x,y) = 0\big\} = \big\{(x, g(x)),\ x \in U\big\}.$$

Avant de démontrer ce théorème, faisons quelques commentaires pour tâcher de comprendre sa signification.

Supposons que f soit une application linéaire. Désignons respectivement par p_E (resp. p_F) la projection sur E (resp. F) parallèlement à F (resp. E). Alors, on a

$$f = f_{|E} \circ p_E + f_{|F} \circ p_F.$$

Si l'application f restreinte à F est inversible, alors on a

$$f(x,y) = 0 \Longleftrightarrow f_{|F}\, y = -f_{|E}\, x.$$

Lorsque $N=2$ et $p=1$, c'est très simple à écrire car $f(x,y) = \alpha x + \beta y$. L'hypothèse sur $f_{|F}$ signifie simplement que $\beta \neq 0$. Dans ce cas bien sûr,

$$f(x,y) = 0 \Longleftrightarrow y = -\frac{\alpha}{\beta} x.$$

Démonstration du théorème 8.2.1. Elle consiste à se ramener à appliquer le théorème d'inversion locale. Considérons pour cela l'application f_1 définie de la manière suivante:

$$f_1 \left\{ \begin{array}{ccc} \Omega & \longrightarrow & E \times \mathbb{R}^{N-p} \\ (x,y) & \longmapsto & (x, f(x,y)). \end{array} \right.$$

Il est clair que l'application f_1 est de classe C^k. Calculons sa différentielle au point (a,b). On a

$$Df_1(a,b) \left\{ \begin{array}{ccc} E \times F & \to & E \times \mathbb{R}^{N-p} \\ (\overrightarrow{h}, \overrightarrow{k}) & \mapsto & (\overrightarrow{h}, Df(a,b)_{|E} \cdot \overrightarrow{h} + Df(a,b)_{|F} \cdot \overrightarrow{k}) \end{array} \right.$$

Cette application est inversible. En effet, supposons que $Df_1(a,b)(\overrightarrow{h}, \overrightarrow{k}) = 0$; cela implique que

$$\overrightarrow{h} = 0 \quad \text{et} \quad Df(a,b)_{|E} \cdot \overrightarrow{h} + Df(a,b)_{|F} \cdot \overrightarrow{k}) = 0.$$

Il en résulte que $\overrightarrow{h} = 0$ et donc que $Df(a,b)_{|F} \cdot \overrightarrow{k} = 0$. Mais comme $Df(a,b)_{|F}$ est inversible, on a $\overrightarrow{k} = 0$. Donc, l'application $Df_1(a,b)$ est inversible. D'après le théorème d'inversion locale, il existe deux ouverts Ω_1 et Ω_2 tels que f_1 soit un difféomorphisme de Ω_1 sur Ω_2 dont on désigne par g_1 l'inverse. Posons alors

$$g(x) = g_1(x,0).$$

Soit U l'ensemble des x de E tels qu'il existe un élément z de $\mathbb{R}^{N-p}$ tel que $(x,z) \in \Omega_2$. Par définition de g_1, on a

$$\{(x,y) \in \Omega_1 \,/\, f(x,y) = 0\} = \{g_1(x,0), (x,0) \in \Omega_2\} = \{(x, g(x))\,,\; x \in U\}.$$

$\square$

Chapitre 9

Les hypersurfaces de $\mathbb{R}^N$

9.1 Définitions des hypersurfaces de $\mathbb{R}^N$

Nous allons définir les hypersurfaces de $\mathbb{R}^N$ comme étant localement, l'ensemble des zéros d'une fonction régulière dont la différentielle est non nulle. Cette hypothèse sur la différentielle est essentielle, car sans cette restriction, la définition que nous venons d'exposer est totalement sans intérêt du fait du théorème suivant:

Définissons maintenant de manière précise le concept d'hypersurface.

Définition 9.1.1. *Soit S une partie de $\mathbb{R}^N$ et k un entier supérieur ou égal à 1; on dit que S est une hypersurface C^k de $\mathbb{R}^N$ si et seulement si pour tout point x_0 de S, il existe un ouvert U de $\mathbb{R}^N$ contenant x_0 et une fonction f de classe C^k sur ω, à valeurs réelles telle que*

- *on ait $f^{-1}(0) = S \cap U$,*
- *pour tout point x de $S \cap U$, on a $Df(x) \neq 0$.*

Avant de poursuivre, donnons toute de suite un exemple très simple, la sphère unité euclidienne notée $\mathbb{S}^{n-1}$. C'est l'ensemble des points $x = (x_1, \cdots, x_n)$ de $\mathbb{R}^N$ tels que

$$\|x\|^2 = \sum_{j=1}^{n} x_j^2 = 1.$$

Considérons l'ensemble $\omega = \mathbb{R}^N \setminus \{0\}$ et posons $f(x) \stackrel{\text{déf}}{=} \|x\|^2 - 1$. Par définition de la sphère, on a $\mathbb{S}^{n-1} \cap \omega = f^{-1}(0)$. De plus, d'après l'un des exemples donnés page 64, on a

$$Df(x) \cdot \overrightarrow{h} = 2(x|\overrightarrow{h}).$$

Donc, dès que $x \neq 0$, la différentielle $Df(x)$ est non nulle; donc les conditions de la définition sont bien satisfaites.

Nous allons donner tout de suite une propriété d'invariance des hypersurfaces.

Proposition 9.1.2. *Soit S une hypersurface C^k incluse dans un ouvert Ω de $\mathbb{R}^N$ et χ un C^k difféomorphisme de Ω sur Ω'. L'ensemble $\chi(S)$ est une hypersurface de Ω'.*

En effet, soit y_0 un point de $\chi(S)$. Il existe un point x_0 de S tel que $\chi(x_0) = y_0$. Par définition d'une hypersurface, il existe un ouvert ω contenant x_0 et une fonction f de classe C^k

de ω dans $\mathbb{R}$ tel que $S \cap \omega = f^{-1}(0)$. Posons $\widetilde{f} \stackrel{\text{déf}}{=} f \circ \chi^{-1}$. Cette fonction est une fonction de classe C^k sur l'ouvert $\widetilde{\omega} \stackrel{\text{déf}}{=} \chi(\omega)$. De plus, il est clair que $\chi(S) \cap \widetilde{\omega} = \widetilde{f}^{-1}(0)$. Enfin, d'après le théorème de dérivation composé 6.3.1 et le théorème de dérivation de l'application réciproque 8.1.5, on a $D\widetilde{f}(y_0) = Df(x_0) \cdot D\chi(x_0)^{-1}$ qui est bien sr non nulle puisque $Df(x_0)$ l'est et que l'application linéaire $D\chi(x_0)^{-1}$ est inversible.

Nous allons maintenant donner une définition équivalente des hypersurfaces sous la forme du théorème suivant.

Théorème 9.1.3. *Soit S une partie de $\mathbb{R}^N$. Cette partie S est une hypersurface C^k si et seulement si pour tout point x_0 de S, il existe un ouvert ω de $\mathbb{R}^N$ contenant 0 et un difféomorphisme φ de classe C^k de ω sur U, ouvert de $\mathbb{R}^N$ contenant 0 tel que*

$$\varphi(S \cap \omega) = U \cap \{y \in \mathbb{R}^N \ / y_1 = 0\}.$$

Démonstration. Pour démontrer ce théorème, supposons que, pour tout point x_0 de S, il existe un ouvert ω de $\mathbb{R}^N$ contenant 0 et un difféomorphisme φ de classe C^k de ω sur U, ouvert de $\mathbb{R}^N$ contenant 0 tel que l'image de $S \cap \omega$ par φ soit l'intersection de U avec l'hyperplan $x_1 = 0$. Posons alors $f = \varphi^1$, la première composante de φ dans $\mathbb{R}^N$. Il est clair que l'application f est de classe C^k et que $f^{-1}(0) = S \cap \omega$. Comme $Df(x) = D\varphi^1(x)$, la forme linéaire $Df(x)$ apparait comme la première ligne de la matrice $D\varphi(x)$, matrice qui est inversible puisque φ est supposé être un difféomorphisme. Donc $Df(x) \neq 0$.

Réciproquement, soit S une hypersurface de $\mathbb{R}^N$ et x_0 un point quelconque de S. On considère l'ouvert $\widetilde{\omega}$ et l'application f donnée par la définition 9.1.1. On peut supposer sans perte de généralité que

$$\frac{\partial f}{\partial x_1}(x_0) \neq 0.$$

On considère alors l'application φ définie par

$$\varphi \left\{ \begin{array}{ccc} \widetilde{\omega} & \rightarrow & \mathbb{R}^N \\ x & \mapsto & (f(x), \Pi(x - x_0)) \end{array} \right.$$

où Π désigne la projection sur les $N-1$ dernières dernières coordonnées, c'est-à-dire

$$\Pi(x_1, \cdots x_N) = (x_2, \cdots x_N).$$

Vérifions que $D\varphi(x_0)$ est inversible. En effet, la matrice de $D\varphi(x_0)$ (application linéaire de $\mathbb{R}^N$ dans $\mathbb{R}^N$ dans la base canonique de $\mathbb{R}^N$ est

$$\begin{pmatrix} \partial_1 f(x_0) & \partial_2 f(x_0) & \cdots & \cdots & \partial_N f(x_0) \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & \cdots & 0 & 1 & 0 \\ 0 & \cdots & 0 & 0 & 1 \end{pmatrix}$$

Le fait que $\partial_1 f(x_0) \neq 0$ entraine que $D\varphi(x_0)$ est inversible. Le théorème d'inversion locale implique l'existence d'un ouvert ω contenant x_0 (et inclus dans $\widetilde{\omega}$) tel que l'application φ soit un difféomorphisme sur son image $U = \varphi(\omega)$. Par construction, le difféomorphisme φ transforme l'hypersurface S en un hyperplan et le théorème est démontré. □

Cette procédure fondamentale est appelée redressement de l'hypersurface S.

Nous allons donner un exemple d'application de cette idée de redressement au travers de la proposition suivante.

Proposition 9.1.4. *Soit ω un ouvert de $\mathbb{R}^N$ et f une fonction de classe C^k sur ω telle que, pour tout point x, on ait*

$$f(x) = 0 \Longrightarrow Df(x) \neq 0.$$

Si g est une fonction de classe C^k s'annulant sur l'hypersurface $S \stackrel{\text{déf}}{=} f^{-1}(0)$, il existe alors une fonction $\widetilde{g}$ de classe C^{k-1} sur ω telle que $g = \widetilde{g}f$.

Démonstration. Nous n'allons démontrer cette proposition que localement. Soit un point x_0 de S, considérons alors le difféomorphisme φ et l'ouvert U construits à partir de f dans la démonstration du théorème 9.1.3. Posons $g_U \stackrel{\text{déf}}{=} g \circ \varphi^{-1}$. Cette fonction s'annulle sur l'intersection de l'hyperplan d'équation $y_1 = 0$ avec U. En écrivant une formule de Taylor à l'ordre 1, on a, lorsque y appartient à une boule ouverte de centre 0 incluse dans U,

$$g_U(y_1, y_2, \cdots, y_n) = g_U(0, y_2, \cdots, y_n) + y_1 \int_0^1 \frac{\partial g_U}{\partial y_1}(ty_1, y_2, \cdots, y_n)dt.$$

Comme la fonction g_U est nulle sur $y_1 = 0$, on a

$$g_U(y_1, \cdots, y_n) = y_1 \widetilde{g}_u(y) \quad \text{avec} \quad g_U(y_1, y_2, \cdots, y_n) \stackrel{\text{déf}}{=} \int_0^1 \frac{\partial g_U}{\partial y_1}(ty_1, y_2, \cdots, y_n)dt.$$

Vérifier que l'application $\widetilde{g}$ est de classe C^{k-1} est un exercice que nous laissons au lecteur. Comme on a, en posant $\pi_1(y) \stackrel{\text{déf}}{=} y_1$,

$$\begin{aligned} g &= g_U \circ \varphi \\ &= (\widetilde{g}_U \pi_1) \circ \varphi \\ &= (\widetilde{g}_U \circ \varphi)(\pi_1 \circ \varphi) \\ &= \widetilde{g} \circ f \quad \text{avec} \quad \widetilde{g} \stackrel{\text{déf}}{=} \widetilde{g}_U \circ \varphi. \end{aligned}$$

D'où la proposition puisque l'on sait que la composée de deux fonctions de classe C^{k-1} en est une. □

Il existe une autre façon de voir les hypersurfaces; elle consiste à les voir localement comme l'image par des applications suffisamment régulières d'ouvert de $\mathbb{R}^{N-1}$.

Théorème 9.1.5. *Soient S une partie de $\mathbb{R}^N$ et k un entier supérieur ou égal à 1; S est une hypersurface si et seulement si pour tout point x_0 de S, il existe un ouvert U_{N-1} de $\mathbb{R}^{N-1}$ contenant 0, un voisinage ouvert ω de x_0 et une fonction Γ de classe C^k de U_{N-1} dans $\mathbb{R}^N$ tels que:*

- *on ait $\Gamma(U_{N-1}) = S \cap \omega$,*
- *la fonction Γ soit injective sur l'ouvert U_{N-1},*
- *la fonction Γ^{-1} soit continue de $S \cap \omega$ sur U_{N-1},*
- *pour tout point x de U_{N-1}, la différentielle de Γ au point x soit injective.*

Démonstration. Elle est présentée à titre culturel . Pour démontrer ce théorème, nous allons utiliser la caractérisation des hypersurfaces donnée dans l'énoncé du théorème 9.1.3. Soit x_0 un point de S tel qu'il existe un ouvert U_{N-1} de $\mathbb{R}^{N-1}$ contenant 0, un voisinage ouvert ω de x_0 et une fonction Γ de classe C^k de U_{N-1} dans $\mathbb{R}^N$ tels que l'image de Γ soit $S\cap\omega$, telle que la fonction Γ soit injective sur l'ouvert U_{N-1}. Considérons un vecteur $\overrightarrow{v}_0$ de $\mathbb{R}^N$ n'appartenant pas à l'image de l'application linéaire $D\Gamma(0)$ (qui est de rang au plus $n-1$) et définissons alors l'application ψ par

$$\psi\left\{\begin{array}{ccc}\mathbb{R}\times U_{N-1} & \rightarrow & \mathbb{R}^N\\ (y_1,y') & \mapsto & \Gamma(y')+y_1\overrightarrow{v}_0.\end{array}\right.$$

Par construction $\psi(U\cap\{y_1=0\})\subset S\cap\omega$. Pour démontrer l'égalité, nous allons procéder par contraposition. Nous allons démontrer que si il existe une suite $(y'_n)_{n\in\mathbb{N}}$ tendant vers 0 et une suite $(\lambda_n)_{n\in\mathbb{N}}$ de réels non nuls telle que

$$\lim_{n\to\infty}\Gamma(y'_n)+\lambda_n\overrightarrow{v}_0=x_0$$

alors l'application linéaire $D\Gamma(0)$ n'est pas injective.

Il est clair que la suite $(\lambda_n)_{n\in\mathbb{N}}$ ci-dessus tend vers 0. Comme $S\cap\Omega=\Gamma(U_{n-1})$, il existe une suite $(y''_n)_{n\in\mathbb{N}}$ telle que

$$\Gamma(y'_n)+\lambda_n\overrightarrow{v}_0=g(y''_n).$$

Soit P_0 la projection sur $\mathrm{Im}D\Gamma(0)$ par rapport à un supplémentaire contenant le vecteur $\overrightarrow{v}_0$, on a

$$P_0(\Gamma(y'_n)-\Gamma(y''_n))=0$$

Le fait que l'application Γ^{-1} soit continue de $S\cap\omega$ dans U_{n-1} implique la suite $(y''_n)_{n\in\mathbb{N}}$ tend vers 0. Par définition de la différentielle et comme les suites $(y'_n)_{n\in\mathbb{N}}$ et $(y''_n)_{n\in\mathbb{N}}$ tendent vers 0, pour tout $\varepsilon>0$, il existe un entier n_0 tel que, pour tout $n\geq n_0$, on ait

$$\|D\Gamma(0)(y'_n-y''_n)\|\leq\varepsilon\|y'_n-y''_n\|.$$

Alors, pour tout $\varepsilon>0$, il existe un n_0 tel que, pour tout $n\geq n_0$, on ait

$$\|D\Gamma(0)\overrightarrow{z}_n\|\leq\varepsilon\quad\text{avec}\quad\overrightarrow{z}_n\stackrel{\text{déf}}{=}\frac{y'_n-y''_n}{\|y'_n-y''_n\|}.$$

La suite $(\overrightarrow{z}_n)_{n\in\mathbb{N}}$ est une suite de norme 1 de $\mathbb{R}^{N-1}$, on peut donc en extraire une sous-suite convergente vers un vecteur $\overrightarrow{z}_\infty$ de norme 1 tel que $D\Gamma(0)\overrightarrow{z}_\infty=0$, ce qui implique que $D\Gamma 0)$ n'est pas injective. Comme $D\Gamma(0)$est supposée injective, il existe un ouvert $\widetilde{U}$ de $\mathbb{R}^N$ contenant 0 tel que l'on ait $\psi(\widetilde{U}\cap\{y_1=0\})=S\cap\omega$.

Calculons maintenant la différentielle de ψ au point 0, puis démontrons qu'elle est inversible. On a

$$\begin{aligned}D\psi(0)\cdot\overrightarrow{h} &= D\psi(0)\cdot(h_1,\overrightarrow{h}')\\ &= D\Gamma(0)\cdot\overrightarrow{h}'+h_1\overrightarrow{v}_0.\end{aligned}$$

Comme $D\Gamma(0)$ est une application linéaire injective de $\mathbb{R}^{N-1}$ dans $\mathbb{R}^N$, son image est un espace vectoriel de dimension $N-1$. Comme $\overrightarrow{v}_0$ n'appartient pas à cet espace, l'image de $D\psi(0)$ est l'espace $\mathbb{R}^N$ tout entier. On peut donc appliquer le théorème d'inversion locale en 0. Il existe ainsi un ouvert U contenant 0, un ouvert ω contenant x_0 tel que ψ soit difféomorphisme de U sur ω.

Réciproquement, supposons que pour tout point x_0 de S, il existe un ouvert ω de $\mathbb{R}^N$ contenant 0 et un difféomorphisme φ de classe C^k de ω sur U, ouvert de $\mathbb{R}^N$ contenant 0 tel que l'image de $S\cap\omega$ par φ soit $U\cap\{y\,/y_1=0\}$. On définit alors

$$U_{n-1}\stackrel{\text{déf}}{=}\{y'\in\mathbb{R}^{N-1}\ /\ (0,y')\in U\}.$$

Le lecteur vérifiera à titre d'exercice que U_{n-1} est un ouvert de $\mathbb{R}^{N-1}$. On pose alors

$$\Gamma(y')=\varphi^{-1}(0,y').$$

L'application Γ vérifie par construction les propriétés voulues, d'où le théorème. $\square$

9.2 Espace tangent à une hypersurface de $\mathbb{R}^N$

Le but de cette section est la définition et l'étude de la notion de vecteurs tangents à une hypersurface.

Définition 9.2.1. *Soient S une hypersurface de $\mathbb{R}^N$ et x_0 un point de S. On dit qu'un vecteur $\overrightarrow{v}$ de $\mathbb{R}^N$ est tangent à S au point x_0 si et seulement si il existe une fonction γ définie sur un intervalle ouvert I de $\mathbb{R}$, contenant 0, de classe C^1 et telle que l'on ait:*

- *l'image de 0 par γ est le point x_0,*
- *la courbe γ soit tracée sur S, c'est-à-dire que $\gamma(I) \subset S$,*
- *le vecteur dérivé de γ en 0 est $\overrightarrow{v}$.*

On désigne par $T_{x_0}S$ cet ensemble.

Proposition 9.2.2. *Soit un point d'une hypersurface S de classe C^k incluse dans un ouvert Ω de $\mathbb{R}^N$ et χ un C^k difféomorphisme de Ω sur Ω'. Alors on a*

$$T_{\chi(x_0)}\chi(S) = D\chi(x_0)(T_{x_0}S).$$

Démonstration. Soit $\overrightarrow{v}$ un vecteur tangent à S au point x_0. Par définition, il existe un intervalle ouvert de $\mathbb{R}$ contenant 0 et une fonction γ de classe C^1 de I dans S telle que l'on ait

$$\gamma(0) = x_0 \quad \text{et} \quad \frac{d\gamma}{ds}(0) = \overrightarrow{v}.$$

La courbe définie par $\widetilde{\gamma} \stackrel{\text{déf}}{=} \chi \circ \gamma$ est une fonction de classe C^1 de I dans $\chi(S)$ telle que

$$\widetilde{\gamma}(0) = \chi(x_0) \quad \text{et} \quad \frac{d(\chi\circ\gamma)}{ds}(0) = D\chi(\gamma(0))\frac{d\gamma}{ds}(0) = D\chi(x_0)\overrightarrow{v}.$$

On a donc

$$T_{\chi(x_0)}\chi(S) \subset D\chi(x_0)(T_{x_0}S) \quad \text{et donc que} \quad D\chi(x_0)^{-1}(T_{\chi(x_0)}\chi(S)) \subset (T_{x_0}S).$$

Appliquons maintenant l'inclusion de avec χ^{-1}. Il en résulte alors que

$$T_{x_0}S \subset D\chi(x_0)^{-1}(T_{\chi(x_0)}S).$$

D'où l'égalité. □

Cette propriété d'invariance est très importante comme nous allons le voir tout de suite.

Théorème 9.2.3. *L'ensemble des vecteurs tangents à un point x_0 d'une hypersurface S est un hyperplan de $\mathbb{R}^N$.*

Démonstration. Nous allons utiliser la méthode du redressement. Soit f une application de classe C^1 sur un voisinage ouvert ω de x_0 tel que $S \cap \omega = f^{-1}(0)$, on considère alors le redressement φ utilisé dans la démonstration du théorème 9.1.3. La proposition 9.2.2 ci-dessus implique que

$$T_{x_0}S = D\varphi(x_0)^{-1}(T_0H).$$

Mais qu'est-ce que T_0H? Soit $\overrightarrow{w}$ un quelconque vecteur de l'hyperplan H, on considère la courbe γ définie par

$$\gamma(s) \stackrel{\text{déf}}{=} s\overrightarrow{w}$$

sur l'intervalle $]-a, a[$, le nombre réel a étant tel que la boule de centre 0 et de rayon $a\|\overrightarrow{w}\|$ soit incluse dans U. Il est clair que γ est une courbe tracée sur H et passant par 0 dont la dérivée en 0 est $\overrightarrow{w}$. Donc $T_0H = H$. Ainsi donc $T_{x_0}S = D\varphi(x_0)^{-1}H$. C'est donc un espace vectoriel de dimension $n-1$. □

Nous allons maintenant donner deux manières de calculer l'hyperplan tangent à une hypersurface de $\mathbb{R}^N$ par le biais du théorème suivant.

Théorème 9.2.4. *Soit x_0 un point d'une hypersurface S de classe C^1 de $\mathbb{R}^N$. Considérons une fonction f de classe C^1 au voisinage de x_0 telle que, au voisinage de x_0, on ait $S = f^{-1}(0)$. Alors*

$$T_{x_0}S = \ker Df(x_0).$$

Soient Γ une fonction de classe C^1 de U_{N-1}, voisinage ouvert de 0 dans $\mathbb{R}^{N-1}$ dans $\mathbb{R}^N$ et ω un ouvert de $\mathbb{R}^N$ contenant x_0 tels que Γ soit injective, pour tout x' de U_{N-1}, la différentielle de Γ en x' soit injective, telle que $\Gamma(U_{N-1}) = S \cap \omega$ et $\Gamma(0) = x_0$, alors

$$T_{x_0}S = \operatorname{Im} Dg(0)$$

Démonstration. Considérons un élément $\overrightarrow{v}$ de $T_{x_0}S$. Par définition, il existe une courbe γ tracée sur S, passant par x_0, telle que

$$\frac{d\gamma}{ds}(0) = \overrightarrow{v}.$$

Mais, comme γ est tracée sur S, on a $f \circ \gamma = 0$. On a donc

$$Df(x_0) \cdot \overrightarrow{v} = 0.$$

Ainsi donc avons nous démontré que

$$T_{x_0}S \subset \ker Df(x_0).$$

Ces deux ensembles sont deux espaces vectoriels de même dimension $n-1$; ils sont donc égaux.

Considérons maintenant un vecteur $\overrightarrow{w}$ de $\mathbb{R}^{N-1}$; on définit alors la courbe $\widetilde{\gamma}$ par

$$\widetilde{\gamma} \left\{ \begin{array}{rcl}]-a, a[& \rightarrow & \mathbb{R}^{N-1} \\ s & \mapsto & s\overrightarrow{w}, \end{array} \right.$$

où a est tel que $B(0, a\|\overrightarrow{w}\|) \subset U_{N-1}$. La courbe $\gamma \stackrel{\text{déf}}{=} g \circ \widetilde{\gamma}$ est une courbe tracée sur S, passant par x_0 et tet que

$$\frac{d\gamma}{ds}(0) = Dg(0) \cdot \overrightarrow{w}.$$

Donc, nous avons démontré l'inclusion de $\operatorname{Im} Dg(0)$ dans $T_{x_0}S$. À nouveau, ce sont deux espaces vectoriels de même dimension $n-1$; ils sont donc égaux; d'où le théorème. $\square$

Théorème 9.2.5. *Soit S une hypersurface de $\mathbb{R}^N \times \mathbb{R}$ de classe C^k. Soit (x_0, u_0) un point de Γ tel que*

$$T_{(x_0,u_0)}S \cap \{0\} \times \mathbb{R}_u = \{0\}.$$

Il existe alors un ouvert Ω de $\mathbb{R}^N \times \mathbb{R}$ contenant (x_0, u_0), un ouvert ω de $\mathbb{R}^N$ et une fonction u de classe C^k sur ω tels que

$$S \cap \Omega = \{(x, u(x)), x \in \omega\}.$$

Ce théorème, qui n'est qu'une formulation géométrique du théorème des fonctions implicites, signifie que si l'espace tangent à S au point (x_0, u_0) ne contient pas de direction "verticale", c'est-à-dire de directions du type $(\overbrace{0, \cdots, 0}^{N \text{fois}}, 1)$, alors, cette hypersurface S est localement le graphe d'une fonction.

Démonstration du théorème 9.2.5 Pour démontrer ce théorème, considérons maintenant une équation locale f de S. Comme l'espace tangent est le noyau de la différentielle de f, on en déduit que $\partial_u f$ est non nul. Le théorème des fonctions implicites assure alors que S est localement de la forme

$$\{(x, u(x)), x \in \omega\}.$$

9.3 Extrema d'une fonction sur une hypersurface

Dans cette courte section, nous allons nous intéresser au problème suivant: comment trouver les minima ou les maxima locaux de la restriction d'une fonction à une hypersurface?

Une condition nécesaire est donné par le théorème suivant:

Théorème 9.3.1. *Soient g une fonction de classe C^1 sur un ouvert Ω de $\mathbb{R}^N$ et S une hypersurface de classe C^1 de Ω. Considérons un point x_0 de S; si x_0 est un minimum (ou un maximum local), alors*

$$Dg(x_0)_{|T_{x_0}S} = 0.$$

Démonstration. Pour démontrer cette propriété, supposons que le point x_0 est un minimum local de la fonction g restreinte à S et considérons un vecteur $\overrightarrow{v}$ tangent à S au point x_0. Par définition, il existe alors une courbe γ tracée sur S, passant par x_0 et telle que

$$\frac{d\gamma}{ds}(0) = \overrightarrow{v}.$$

Le point x_0 est un minimum local de la fonction g restreinte à S, donc, en particulier, le point 0 est un minimum local de

$$t \mapsto g \circ \gamma.$$

Ainsi donc, on a

$$\frac{d\,(g \circ \gamma)}{ds}(0) = 0.$$

Mais le théorème de composition dit que

$$\frac{d\,(g \circ \gamma)}{ds}(0) = Dg(x_0) \cdot \overrightarrow{v}.$$

D'où le théorème. □

On peut en déduire très facilement le corollaire suivant:

Corollaire 9.3.2. *Soient g une fonction de classe C^1 sur un ouvert Ω de $\mathbb{R}^N$ et S une hypersurface de classe C^1 de Ω. Considérons un point x_0 de S et une fonction f de classe C^1 sur un voisinage ω de x_0 telle que $S = f^{-1}(0) \cap \omega$; si x_0 est un minimum (ou un maximum local), alors, il existe un réel λ_0 tel que*

$$Dg(x_0) = \lambda_0 Df(x_0).$$

Remarque Le réel λ_0 est parfois appelé mutiplicateur de Lagrange.

Nous allons illustrer l'intérêt d'un tel résultat sur un exemple.

Proposition 9.3.3. *Soient S une hypersurface de $\mathbb{R}^N$ et (a,b) un couple de points de $\mathbb{R}^N \setminus S$; on considère alors la fonction g définie par*

$$g(x) \stackrel{\text{déf}}{=} \|x-a\| + \|x-b\|,$$

où $\|\cdot\|$ désigne la norme euclidienne sur $\mathbb{R}^N$. Si un point x_0 de S est un extremum local de $g_{|S}$, alors on a

$$\frac{x_0-a}{\|x_0-a\|} + \frac{x_0-b}{\|x_0-b\|} \in (T_{x_0}S)^{\perp}.$$

Démonstration. Il suffit d'observer que, en tout point x de $\mathbb{R}^N \setminus \{a,b\}$, on a

$$Dg(x)\cdot \overrightarrow{h} = \frac{(x-a|\overrightarrow{h})}{\|x_0-a\|} + \frac{(x-b|\overrightarrow{h})}{\|x-b\|}\cdot$$

Le fait que, pour tout $\overrightarrow{h} \in T_{x_0}S$, on doit avoir $Dg(x_0)\cdot\overrightarrow{h} = 0$ donne immédiatement le résultat. □

Remarques On peut ainsi établir les lois de la réflexion en partant du principe dit "variationnel" qui soutient que tout mouvement se fait le long d'extremale d'une fontionnelle bien choisie.

Nous allons maintenant étudier l'équivalent de la proposition 7.3.3 dans le cadre de la restriction à une hypersurface. Le résultat est le suivant.

Théorème 9.3.4. *Soient f une fonction C^2 d'un ouvert Ω de $\mathbb{R}^N$ et a un point d'une hypersuface S de classe C^2 de Ω.*

Si a est un minimum local, alors si g est une l'equation de S près de a, il existe un réel λ_0 tel que

$$Df(a) = \lambda_0 Dg(a) \quad \text{et} \quad \forall \overrightarrow{h} \in T_aS\,,\ (D^2f(a) - \lambda_0 D^2g(a))(\overrightarrow{h},\overrightarrow{h}) \geq 0.$$

Si g est une l'equation de S près de a, et s'il existe un réel λ_0 tel que

$$Df(a) = \lambda_0 Dg(a) \quad \text{et} \quad \forall \overrightarrow{h} \in T_aS\,,\ (D^2f(a) - \lambda_0 D^2g(a))(\overrightarrow{h},\overrightarrow{h}) \geq c_0\|\overrightarrow{h}\|^2,$$

alors a est un minimum local.

Remarque La quantité pertinente pour étudier la variation seconde de la restriction d'une fonction f à une hypersurface n'est pas la restriction de la différentielle seconde à l'espace tangent. Ceci est illustré par l'exemple suivant.

On considère sur $\mathbb{R}^2$ la fonction

$$f(x_1,x_2) = x_2 + x_1^2 + x_2^2$$

et les deux surfaces S_1 et S_2 définies respectivement par $g_1(x_1,x_2) = x_2 - x_1^2$ et $g_2(x,y) = x_2 + 4x_1^2$. On a

$$Df(0,0)(h_1,h_2) = h_2\,,\ Dg_1(0,0)(h_1,h_2) = Dg_2(0,0)(h_1,h_2) = h_2 \text{ et}$$
$$D^2f(0,0)((h_1,h_2),(h_1,h_2)) = h_1^2 + h_2^2.$$

Mais si $(h_1,h_2) \in S_1$, on a $f(h_1,h_2) = 2x_1^2 + x_1^4$ et si $(x_1,x_2) \in S_2$, on a $f(x_1,x_2) = -3x_1^2 + 16x_1^4$ et donc $(0,0)$ est un minimum local de $f_{|S_1}$ et un maximum local de $f_{|S_2}$

Démonstration du théorème 9.3.4 L'idée consiste à dire que, comme $g_{|S} \equiv 0$, on a

$$f_{|S} = (f - \lambda_0 g)_{|S}.$$

Or $D(f - \lambda_0 g)(a) = 0$ en tant que forme linéaire de $\mathbb{R}^N$. On utilise alors le thérème 9.1.5. Soit Γ une application de classe C^2 d'un ouvert U de $\mathbb{R}^{N-1}$ dans $\mathbb{R}^N$ donnée par le théorème 9.1.5. Comme Γ est un homéomorphisme de U sur $S \cap \omega$ (où ω est un ouvert de $\mathbb{R}^N$ inclus dans Ω, a est un minimum local de $(f - \lambda_0 g)_{|S}$ si et seulement si 0 est un minimum local de $\widetilde{f} \stackrel{\text{déf}}{=} (f - \lambda_0 g) \circ \Gamma$. Comme, d'après le théorème de composition 6.3.1

$$D\widetilde{f}(0) = D(f - \lambda_0 g)(a) \circ D\Gamma(0) = 0,$$

le lemme **??** implique que

$$\begin{aligned} D^2(\widetilde{f})(a) \cdot (\overrightarrow{h}, \overrightarrow{k}) &= D^2(f - \lambda_0 g)(a) \cdot (D\Gamma(0) \cdot \overrightarrow{h}, D\Gamma(0) \cdot \overrightarrow{k}) \\ &\qquad + D(f - \lambda_0 g)(a)) \cdot (D^2\Gamma(0) \cdot (\overrightarrow{h}, \overrightarrow{k})) \\ &= D^2(f - \lambda_0 g)(a) \cdot (D\Gamma(0) \cdot \overrightarrow{h}, D\Gamma(0) \cdot \overrightarrow{k}). \end{aligned}$$

Le fait que $D\Gamma(0)$ soit un isomorphisme de $\mathbb{R}^{N-1}$ sur $T_a S$ permet de conclure la démonstration du théorème. □

9.4 Un théorème sur les zéros d'une fonction régulière

À titre culturel et tout à fait hors programme, pour montrer qu'il est nécessaire de faire des hypothèses sur l'équation d'une hypersurface, nous présentons le théorème suivant.

Théorème 9.4.1 (de Whitney)**.** *Soit F une partie fermée de $\mathbb{R}^N$. Il existe une fonction φ de classe C^∞ de $\mathbb{R}^N$ dans $\mathbb{R}$ telle que F soit exactement l'ensemble des zéros de la fonction φ.*

La démonstration de ce théorème, dont l'énoncé est un peu surprenant, n'est présentée ici qu'à titre culturel et sa lecture peut être omise sans inconvénient pour la suite du cours.

La construction de la fonction φ est assez explicite. On rappelle que la distance d'un point x à un ensemble A, est définie par

$$d(x, A) \stackrel{\text{déf}}{=} \inf_{a \in A} d(x, a).$$

De plus, si l'ensemble A est fermé, alors x est élément de A si et seulement si la distance de x à A est nulle. Désignons par F_p l'ensemble des points x de $\mathbb{R}^N$ tels que la distance de x à F soit comprise entre 2^{-p-1} et 2^{-p+2} lorsque p est un entier positif et par F_{-1} l'ensemble des points x de $\mathbb{R}^N$ à distance plus grande 1 de F.

On considère alors la fonction caractéristique de F_p que l'on note $\mathbf{1}_{F_p}$ et on la régularise par convolution. Pour ce faire, on considère une fonction χ de classe C^∞, positive, d'intégrale 1 et qui est nulle en dehors de la boule unité de $\mathbb{R}^N$ (exercice : construire explicitement une telle fonction); on pose alors

$$\varphi_p(x) \stackrel{\text{déf}}{=} 2^{(p+2)n} \chi(2^{p+2} \cdot) \star \mathbf{1}_{F_p} \quad \text{pour} \quad p \geq -1.$$

Nous allons démontrer que, si p est un entier positif, alors la fonction φ_p est nulle en dehors de

$$\{x \in \mathbb{R}^N \ / \ d(x, F_p) \in 2^{-p}[1/4, 17/4]\}$$

et que la fonction φ_p vaut 1 sur

$$\widetilde{F}_p \stackrel{\text{déf}}{=} \{x \in \mathbb{R}^N \ / \ d(x, F_p) \in 2^{-p}[1, 2]\}.$$

Par définition de la convolution, on a

$$\varphi_p(x) = 2^{(p+2)n} \int_{F_p} \chi(2^{p+2}(x-y))dy.$$

La fonction χ étant positive, la fonction φ_p sera non nulle au point x uniquement s'il existe un point y de F_p et un point z de la boule de centre 0 et de rayon $2^{-(p+2)}$ tel que $x = y + z$. Soit alors $\widetilde{z}$ un point quelconque de F. L'inégalité triangulaire implique que

$$\|y - \widetilde{z}\| - \|z\| \leq \|x - \widetilde{z}\| \leq \|z\| + \|y - \widetilde{z}\|.$$

Comme z appartient à $B(0, 2^{-(p+2)})$, on en déduit que

$$\|y - \widetilde{z}\| - 2^{-(p+2)} \leq \|x - \widetilde{z}\| \leq 2^{-(p+2)} + \|y - \widetilde{z}\|.$$

La borne inférieure étant un minorant, on a

$$d(y, F) - 2^{-(p+2)} \leq \|x - \widetilde{z}\| \leq 2^{-(p+2)} + \|y - \widetilde{z}\|.$$

La borne inférieure étant le plus grand des minorants, on a

$$d(y, F) - 2^{-(p+2)} \leq d(x, F) \leq 2^{-(p+2)} + \|y - \widetilde{z}\|,$$

puis

$$d(y, F) - 2^{-(p+2)} \leq d(x, F) \leq 2^{-(p+2)} + d(y, F).$$

Comme y appartient à F_p, on en déduit que

$$2^{-(p+1)} - 2^{-(p+2)} \leq d(x, F) \leq 2^{-p+2} + 2^{-(p+2)}$$

Donc si x est tel que $\varphi_p(x)$ est non nul alors $d(x, F) \in 2^{-p}[1/4, 17/4]$. Supposons maintenant que $x \in \widetilde{F}_p$ et considérons un couple $(y, z) \in B(x, 2^{-(p+2)}) \times F$; on a alors

$$\|x - \widetilde{z}\| - \|x - y\| \leq \|y - \widetilde{z}\| \leq \|x - \widetilde{z}\| + \|x - y\|.$$

Comme y appartient à $B(x, 2^{-(p+2)})$, on en déduit que

$$\|x - \widetilde{z}\| - 2^{-(p+2)} \leq \|y - \widetilde{z}\| \leq \|x - \widetilde{z}\| + 2^{-(p+2)}.$$

Par passages successifs à la borne inférieure, on obtient

$$d(x, F) - 2^{-(p+2)} \leq d(y, F) \leq d(x, F) + 2^{-(p+2)}.$$

Comme x est supposé appartenir à $\widetilde{F}_p$, on a

$$2^p \left(1 - \frac{1}{4}\right) \leq d(y, F) \leq 2^p \left(2 + \frac{1}{4}\right) \cdot$$

Donc y appartient à F_p et donc, si $x \in \widetilde{F}_p$, on a

$$\varphi_p(x) = 2^{(p+2)n} \int_{\mathbb{R}^N} \chi(2^{p+2}(x-y))dy = 1.$$

Il s'agit maintenant de recoller tout cela, c'est-à-dire de sommer les fonctions φ_p en les multipliant par des coefficents convenables de telle sorte la somme de la série vérifie les propriétés voulues. Posons

$$\varphi \stackrel{\text{déf}}{=} \sum_{p \geq -1} e^{-2^p} \varphi_p.$$

Il est clair que cette série est une série de fonctions normalement convergente dans l'espace $C_b(\mathbb{R}^N; \mathbb{R})$. Il s'agit maintenant de démontrer que la fonction φ vérifie les propriétés voulues. Tout d'abord, la fonction est nulle sur l'ensemble F car pour tout p, la fonction φ_p est nulle sur F. Considérons maintenant un point x de $\mathbb{R}^N$ n'appartenant à F. Il existe un entier p_0 tel que $d(x, F)$ soit comprise entre 2^{-p_0} et 2^{-p_0+1}; autrement dit, il existe un entier p_0 tel que $x \in F_{p_0}$. Donc $\varphi_{p_0}(x)$ vaut 1. Comme toutes les fonctions φ_p sont positives, la fonction φ est bien strictement positive au point x.

Il nous reste à vérifier que la fonction φ est de classe C^∞. On a

$$D^k \varphi_p(x) = 2^{(p+2)(n+k)} \int_{\mathbb{R}^N} D^k \chi(2^{p+2}(x-y)\mathbf{1}(y)dy.$$

Mais la série de terme général $e^{-2^p} 2^{(p+2)k}$ est convergente; d'où le théorème. $\square$